THE UNIVERSITY OF DUBLIN

Material Selection for Spin-Transfer-Torque Magnetic Random Access Memories: a High-Throughput approach

by

Emanuele Bosoni

A thesis submitted in partial fulfillment for the degree of Doctor of Philosophy in the

Trinity College Dublin

School of Physics

2020

Declaration of authorship

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work. I was involved in a number of collaborations, and where it is appropriate my collaborators are acknowledged for their contributions. I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

Date _____

Signature _____

Acknowledgements

I am grateful to a number of people who have helped me to realize this thesis and, more in general, have supported me in these PhD years.

First, I would like to thank my supervisor, Professor Stefano Sanvito, for his guidance, for his patience in listening to my doubts and for his encouragement in pursuing new research paths in an independent manner. Many thanks also to Stefania Negro, who has kindly helped me with many administrative and financial matters throughout these years. Thanks to the AiiDA team in Lausanne, mostly in the person of Dr Giovanni Pizzi; he never missed an occasion to invite and sponsor me in order to participate at AiiDA workshops and coding weeks. I am also very grateful for the collaborative experience of the AiiDA-Siesta plugin, led by Professor Alberto Garcia and Mr Vladimir Dikan.

This PhD has been made an enjoyable experience by the members of the Computational Spintronics group, past and present, I'm in debt to them all. I start from Dr Mario Zič, Dr Jacopo Simoni, Dr Carlo Motta and Mr Filiberto Biolcati Rinaldi. They welcomed me in the group in the best way possible and their support was fundamental for me in the first year of the PhD. Lasts (in the timeline of my PhD) come the new-students crew (Mike, Matteo, Hugo, Luke, Nina, Laura, Michelangelo) and Dr Stella Skiadopoulou, an amazing mix that gave me so much energy during the writing period. To experience their company was a great pleasure. I was lucky to have colleagues on my side for the entire (or almost) duration of the project: Dr Matthew Ellis, Dr Rajarshi Tiwari, Dr Aswenee Kumar, Dr Sabine Körbel, Dr Yanhui Zhang and soon-to-be-doctors Anais Colibaba and Urversh Patil. Thanks for the fruitful discussions, for the kindness inside and outside the working environment and for the great trips to conferences. Overall they gifted me with some amazing memories. Extra special mention goes to the people that, from just colleagues, became my core friends in Dublin. Freshly-doctor Mario Galante, my PhD buddy, I can not imagine a better person as my supporting actor in this PhD movie. In the many struggles we went through together, he was the first to encourage and help me, always. A true fiend. Mr James Nelson, the Irish friend everybody would like to have, his help with machine learning was fundamental, but his influence on me went way beyond science. More-than-doctor (someone says Sir) Alessandro Lunghi, the most amazing researcher, the hardest worker, still always happy to listen to me. Thanks for all the enriching discussions, only few about science.

Many more beautiful people I met in these four years in Dublin, I'm so grateful for them all. Each one was an important part of my life and helped me to keep a positive mood. I would love to write a paragraph for each of them, but this is not supposed to become my biography. I mention few groups. All the CRANN people and the other PhD students of my class (Sonia, Jens, Romy and Brian above all). The senior-invigilators group and the TCD exam office that offered me a great environment for an extra-research job. The good-few housemates I had, I thank them all, except one. In particular, I love to mention my dearest friend Janne, that discovered with me the beauty of Ireland, and

Silvia, so caring towards me in the crazy last period of writing. The DU Judo club, a place I called home for four years. I dedicated time and energy to it and I was repaid with amazing experiences and friendships: Daniel, Ailbhe, Mary and my dearest Èlia. Without forgetting our sensei Donal, a true life-teacher. The Donnybrook-parish group, with Rose, Sowmya, Anita, Paul, Mauricio and Cris, and many more; all there to remind me how beautiful is a friendship rooted in faith. I include here also Benedetta, with whom a share the joy of being restless seeker of meaning in this life.

Finally, I would like to thank my family and friends back in Italy for their continuous support. A support that does not need many words, but that I feel present more than anything.

This work would not have been possible without the financial support from Science Foundation of Ireland through the project "Atomistic simulators for magnetic memories design: MMDesign“. The computational resources have been provided by the Trinity Center for High Performance Computing (TCHPC) and by the Irish Centre for High-End Computing (ICHEC).

Abstract

A Spin-Transfer-Torque Magnetic Random Access Memory (STT-MRAM) is a complex multilayered structure, whose sensitive element is formed by two ferromagnetic layers sandwiching an insulator, acting as a tunnel barrier. The operational principles of the memory rely on two physical effects: the tunnel magneto-resistance (the total resistance of the device depends on the relative orientation of the magnetization vectors of the two magnetic layers) and the spin-transfer-torque effect (the possibility of inducing magnetization reversal by means of a polarized electric current). The STT-MRAM technology shows promising features for the creation of a new class of memories that combines non-volatility with access speed close to that of a Dynamic Random Access Memory. At present, the Fe-Co/MgO materials set gives the best performance, but it is not clear whether further improvement will be possible. Fundamental research is needed to lead the design of new efficient junctions. Even though interfaces play a relevant role in this technology, some peculiar material-related features of the various components are essential for the device operation. Looking for these properties would lead to a selection of suitable ferromagnets and insulators to be combined for efficient STT-MRAM structures. This PhD thesis presents a detailed description of the STT-MRAM technology and a strategy for the selection of promising materials is proposed. The main aim is the creation of a database containing material properties of interest for the technology, calculated within the Density Functional Theory (DFT) framework. In order to extend the study to a large set of compounds, the calculations are performed in an high-throughput fashion. An *a-posteriori* analysis of the information in the database drives the identification of promising candidates for STT-MRAMs. The high-throughput approach is a fairly new methodology in computational material science. In this work, alongside with the scientific investigation, much attention is devoted to the presentation of challenges, benefits and strategies related to the task of performing DFT high-throughput simulations. The main results of the thesis are the identifications of Mn_1Ni_1 as a good candidate for the role of ferromagnet in an STT-MRAM and the formalization of a theory (and relative implementation) for the calculation of the Complex Band Structure of an insulator that is suitable for the high-throughput methodology. This latter material property is of extreme importance in the STT-MRAM technology, but finds applications also in other situations where the description of quantum tunneling is required.

Contents

Introduction **1**

1 STT-MRAM **5**
1.1 Operational principles of the memory 6
1.2 Tunnel Magneto-Resistance and Symmetry Filtering Effect 8
1.3 Spin-Transfer-Torque Effect 10
1.3.1 Toy model 11
1.3.2 Modeling *ab-initio* 14
1.3.3 Phenomenological models 16
1.4 Magnetic Dynamics 17
1.5 State-of-the-art materials combination and challenges 20
1.6 Material selection based on bulk simulations 21

2 Methods **25**
2.1 Density Functional Theory Fundamentals 25
2.2 Kohn-Sham formalism for magnetic systems 29
2.2.1 Collinear spins 31
2.2.2 Non-collinear spins 32
2.3 Pseudopotential Theory 33
2.3.1 Non-linear core correction 35
2.3.2 Relativistic Pseudopotentials 36
2.3.3 Practical aspects 37
2.4 DFT for solids: the Bloch theorem 38
2.4.1 Bloch's theory 38
2.4.2 Plane-wave basis set 41
2.4.3 Local, non-orthogonal basis set for extended systems 41
2.5 SIESTA 44
2.6 Group theory and band symmetries 45
2.7 Band gap in DFT and Atomic Self Interaction Correction 48
2.7.1 The problem of the gap in approximated DFT 49
2.7.2 Strongly correlated systems and DFT+U 49
2.7.3 The self-interaction problem and the ASIC 50
2.8 High-throughput approach 53

2.8.1 High-throughput DFT Simulations: AiiDA . . . 53

3 High-throughput DFT **55**
3.1 Optimized pseudopotentials . . . 56
3.1.1 All-electron calculations: FHI-AIMS . . . 58
3.1.2 Identification of optimal UPF PPs by means of QE calculations . . . 61
3.1.3 Optimized pseudopotatntial in PSF format . . . 65
3.1.4 Pseudopotentials with self-interaction correction . . . 68
3.2 ASIC . . . 70
3.2.1 Experimental data collection and High-throughput workflow . . . 71
3.2.2 Gaps analysis . . . 72
3.2.3 Selection of λ . . . 74

4 Ferromagnet selection: bands symmetries **79**
4.1 Preselection based on the Curie temperature . . . 80
4.2 Bands and symmetries . . . 82
4.3 Results discussion . . . 87

5 Insulator selection: the Complex Band Structure **89**
5.1 Complex band structure definition . . . 90
5.2 Complex band structure calculations: the Transfer Matrix Method . . . 91
5.3 Analytical properties of Complex Band Structure . . . 99
5.3.1 Simple analytical models . . . 100
5.3.2 Properties of the CBS in the case of orthonormal basis set . . . 103
5.3.3 Overlap and flat bands . . . 107
5.4 Complex bands of alkaline oxides/sulphides . . . 110

Conclusions **117**

A Tunneling transport for a finite potential barrier **121**

B Derivations Chapter 2 **125**
B.1 Definition of ρ in second quantization . . . 125
B.2 Introducing the Bloch theorem for local orbitals . . . 125

C Band gap experimental data **127**

Bibliograpy **142**

Introduction

In the last 50 years, magnetic memories have had a leading role in the non-volatile memory industry, providing cheap and reliable data storage capacity and fulfilling most of the needs in the data storage sector. From supercomputer centers to daily use laptops, Hard Disk Drives (HDDs) are still today the most used technology for massive data storage.

The current reading/recording protocol of a HDD involves a magnetic field in order to change the magnetization of ferromagnetic grains and it is a technology only viable down to information densities of about 1 Tb/in^2. Then the superparamagnetic limit will be reached: smaller grains may self-reverse (electro-thermal instability) unless their magnetic strength is increased (harder magnets), but all known write head materials are at the moment unable to generate a magnetic field sufficiently strong to write such an hard medium [1, 2]. A new protocol is then needed to maintain the competitiveness of HDDs with respect to products such as flash-memory-based solid-state drives.

Together with HDDs, a second large technological sector is currently going through a paradigm shift: Dynamic Random Access Memories (DRAMs). The present technology is based on silicon capacitors, where the information is stored in the charging state. Scalability problems are the main issue also in this sector. Moreover here there is less space for improvements because the production costs are already quite high for a technology targeting mass distribution. Quoting from the International Technology Road-map for Semiconductor 2.0 (2015 report): "DRAM products are approaching fundamental limitations as scaling DRAM capacitors is becoming very difficult in 2D structures. It is expected that these limits will be reached by 2024 and after this year DRAM technology will saturate at the 32Gbit level unless some major breakthrough will occur" [3].

In this context, there is space for an attractive technological achievement: the creation of a new class of memory (called Storage Class Memory), that can replace DRAMs introducing non-volatility. In this way no memory refresh would be needed and this would greatly simplify the system design and would reduce overheads. The challenge is then to combine non-volatility with memory access times comparable to nowadays DRAMs.

Various technologies such as Resistive RAM [4] and Phase Change Memories [5] (see recent announcement for 3D XPoint [6]) are claiming to be good candidates for the first real Storage Class Memory. Also the magnetic memory sector has its own candidate: the Spin-Transfer-Torque Magnetic Random Access Memory (STT-MRAM).

The phenomenology that led to the development of all the magnetic recording technologies has its origin in Mott's early work on spin dependent conduction in ferromagnets

when, for the first time, the idea that the spin state of electrons may influence their mobility was postulated [7]. Decisive was in the 1980s the advent of deposition techniques, such as molecular beam epitaxy, that made possible the fabrication of thin films down to the nanometers scale; in fact new magneto-resistance phenomena appears when the layer thicknesses is comparable to the mean free path of electrons. The most significant discovery was the fact that the electrical resistance of a stack composed by two thin layers is different when the magnetizations of the two films are in a parallel or in a anti-parallel alignment. This effect is commonly known as Giant Magneto-Resistance (GMR) and was simultaneously discovered by Albert Fert and Peter Grünberg [8, 9], reason for which they were jointly awarded the Nobel price for physics in 2007. The different resistance between the parallel and the anti-parallel configuration is the core operational principle of the MRAM as the two configurations represent the "0" and "1" state of the memory. The magnitude of the difference in resistance between the two states is the measure of the reading performances of the memory. In this sense improvements were seen including an insulating spacer between the two ferromagnetic layers, forming the so called Magnetic Tunnel Junction (MTJ) structure. In early MRAMs, an external magnetic field was used to switch between parallel and anti-parallel configurations, with the related scalability problems connected to the superparamagnetic limit described at the beginning of the chapter. A new writing protocol was introduced thanks to the STT effect, predicted by Berger and Slonczewski in 1996 [10, 11], that allows the switching by means of a current density only.

The STT-MRAM is the target technology of the study carried out in this thesis and Chapter 1 is entirely devoted to explain the MTJ structure, the operational principles of the memory, the technological state of the development and the challenges to face in order to improve the performances. STT-MRAM is a relatively mature technology, in fact the first commercial products came into the market in 2016 [12]. However, cost and information density are concerns that still need to be address in order to make STT-MRAMs suitable for mass production. The first product on the market had 256 Mb capacity and it was not a low-power device. Since then little enanchements have been seen [13]: there is space for further improvements and this a major point of discussion in the present thesis.

The complexity of the memory stack did not allow to experiment much on the material side and the development was limited essentially to the study of MgO/FeCo(B) junctions that is the state-of-the-art materials set for the STT-MRAMs now on the market. The possibility of a completely new combination of materials with competitive performances is still an open question and the work here presented wants to gather valuable informations to address that question.

Simulations can say a lot on the topic. An integrated approach of Density Functional Theory (DFT) and Spin Dynamics simulations [14] allows us to understand the spin-current distribution, the STT effect and the magnetic response of the device at the nanoscale (locally). As a consequence, a complete description of the current-driven magnetization reversal process, at the core of STT-MRAMs, could be achieved with great benefits for the research aimed to improve the technology. However, the possibility of applying these tools to a vast range of material combinations is limited by the computational cost of

the calculations. It is then important to choose materials combinations suitable for STT-MRAMs in a systematic way by looking at features not of the junction itself, but of the constituent materials. Even though interfaces play a relevant role in the technology, still some peculiar materials-related features are essential for the device operation. Looking for these properties would lead to a pre-selection of suitable ferromagnets and insulators to be combined for efficient STT-MRAM structures. DFT offers an extremely cheap way to obtain these information. This PhD work deals exactly with the task of gathering DFT-calculated properties of a vast range of compounds and help the pre-selection of good materials combinations on which to perform more accurate simulations.

The study is carried out in an high-throughput way, meaning following the idea of building a large database of interesting calculated properties for as many materials as possible and let the data guide the choice of good candidates. It is in fact a *a-posteriori* analysis to hopefully indicate the existence of materials with the desired characteristics. The high-throughput approach in computational materials science is a relatively new investigation method and it presents some technical challenges compared to standard simulations. It requires, in fact, the setting up of a framework able to automatically prepare, submit and retrieve the results of a large number of calculations. Moreover storing the data in a database and keeping the provenance of each single node of the database is essential. In this work the discussion on benefits and challenges of the high-throughput approach and the development of strategies and actual tools indispensable for an effective use of the methodology has as much importance as the production of scientifically relevant data. The totality of Chapter 3 and part of Chapter 2 are discussing the topic. To the best of our knowledge, this work represents the first high-throughput study ever made employing the DFT code SIESTA.

The scientific results of the thesis, besides the methodological developement, are divided into two chapters, each one of them related to one of the two relevant constituents of an MTJ: the ferromagnet (Chapter 4) and the insulator (Chapter 5).

The study brought us also to a deep analysis of the Complex Band Structure (CBS) of a material. Some technical difficulties in the calculations of CBS when expanding the wave functions on an non-orthogonal basis set are taken as a stimuli to deepen our knowledge regarding the analytical properties of the CBS and, more in general, of the complex multi-valued function $E(\mathbf{k})$ representing the dispersion relation of a material when $\mathbf{k}$ is allowed to be complex. Our study details the mathematical foundations for the understanding of a theory of the CBS in presence of non-orthogonal basis sets and offers a computational tool to calculate the CBS suitable for high-throughput studies. This latter part of the work (also presented in Chapter 5) find its applicability way beyond the study of the STT-MRAM technology.

Chapter 1

STT-MRAM

This chapter is devoted to the presentation of all the phenomenology at the core of the Spin Transfer-Torque Magnetic Random Access Memory (STT-MRAM) technology. Particular attention is paid to the description of the operational principles of the memory and of the physical properties required by the materials combination at the core of the device, meaning the ferromagnet/insulator junction. The concepts of ferromagnets and insulators are assumed to be known by the reader, even though the understanding of the microscopic origin of magnetism and of the band gaps in insulators are not an easy take. Good books on the first topic are references [15, 16], on the second we recommend the reference [17]. There are three main ingredients to understand the STT technology and these are treated in separate sections: the Tunnel Magneto-Resistance (TMR - Section 1.2), the STT effect (Section 1.3) and basic concepts of magnetic dynamics (Section 1.4). Once the background is set, the state-of-the art materials combination used in prototype STT-MRAM devices is analyzed, together with a list of challenges that one need to overcome to improve the technology. Finally, the central question motivating the entire project is discussed: what insights can we have from the study of bulk properties only to lead to the selection of new materials combinations for STT-MRAMs? The theory in the first four sections of this chapter shows clearly that there are some fundamental physical properties that the insulator/ferromagnet combination needs to have, independently on the interface or the stack design. These features must be the first target for a pre-selection of new materials for STT-MRAMs, but they are often difficult to access experimentally. On the contrary, they can be inexpensively calculated in the context of Density Functional Theory (DFT) and this opens up the possibility to approach the problem in an high-throughput way, namely we can systematically study entire categories of ferromagnets and insulators and let the data guide the discovery of new promising combinations. It is important to mention that a combination of DFT and Green's function method allows the direct calculation of the TMR and the STT effect. However this can not be done in an high-throughput way because of the high computational costs involved. No theoretical work on promising STT-MRAM materials would be complete without the calculations just mentioned, however this PhD work sets itself a step before. The high-throughput analysis of bulk properties wants to be the very first building block for a more extensive theoretical work, meaning

to provide a first guess of crystal structures with promising features on which to perform more accurate simulations. The ultimate goal is to shed light on the possibility to move away from the current materials set and to embrace completely new research paths for the improvement of STT-MRAMs. For many years now, the complex fabrication process of STT-MRAM devices and the high computational costs of STT calculations, forced researchers to experiment very little on the materials side. The possibility of a completely new, more performing, junctions is still an open question and this chapter wants to explain how our work can contribute to the answer.

1.1 Operational principles of the memory

A generic MRAM is a spin valve, *i.e.* a complex multilayer, whose sensitive element is formed by two ferromagnetic layers sandwiching an insulator, acting as tunnel barrier. This material combination is often referred in literature with the name Magnetic Tunnel Junction (MTJ). The first magnetic layer (FM1 in Figure 1.1) spin-polarizes the electrical current, while the second (FM2) acts as an analyzer. The total resistance of the device depends on the relative orientation of the magnetization vectors of the two magnetic layers (TMR effect). The resistance is usually maximized for anti-parallel orientation (R_{AP} in Figure 1.1) and minimized for parallel one (R_P) [18], however situations where the opposite happens have been recorded [19]. The reading procedure in the memory is then easily achieved by comparing the resistance at low bias in the two magnetic configurations. Larger is the difference in the resistance [the TMR ratio $= (R_{AP} - R_P)/R_P$], higher is the sensitivity of the sensor. Data areal densities in excess of 1 Tb/in^2 require TMR ratios in excess of 200% at room temperature in order to clearly identify the 0 and 1 state. At present only the Co-Fe/MgO materials set is capable of achieving such target, but it is not clear whether further improvement will be possible. A closer look at the TMR effect and its microscopic origin is in Section 1.2.

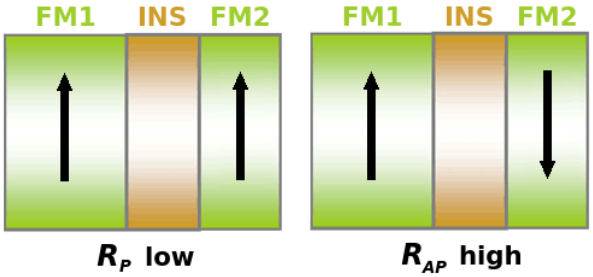


Figure 1.1: Sensitive element of an STT-MRAM in parallel configuration (left) and anti-parallel configuration (right). The resistivity (R) of the parallel set-up is usually lower than that of the anti-parallel one. Picture adapted from reference [20].

The described reading procedure is common to all the MRAMs. The peculiar revolution of STT-MRAMs is in the writing procedure that exploits the the so called Spin Transfer-Torque (STT) effect [21, 10, 11]. A spin-polarized electrical current passing through a magnet can exchange angular momentum with the static magnetization and exert a torque, which in turn can drive magnetization reversal. Such torque is proportional to the current

density, hence the STT effect is fully down scalable. This is very much an improvement respect to conventional MRAMs, where an induced magnetic field is employed for the writing procedure, leading to considerable additional energy costs and scalability problems. STT-MRAMs are thus MRAM junctions operating also in a reverse mode, *i.e.* where the spin-polarization of the current generated by the polarizer is used to manipulate the magnetic state of the analyzer by rotating its magnetization (Figure 1.2). Section 1.3 goes in more detail about the origin and phenomenology of the STT effect.

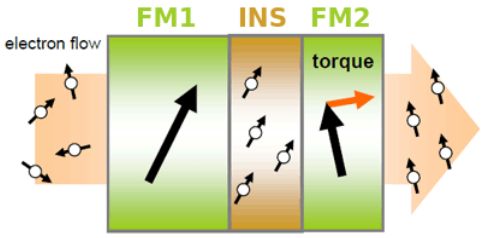


Figure 1.2: Schematic representation of the STT effect. The first magnetic layer (FM1, polarizer) spin polarizes the current. This current exerts a torque on the second ferromagnet (FM2, analyzer). Picture adapted from reference [20].

The STT effect is present on FM1 as much as on FM2. In order to ensure that FM1 keeps its polarizer role during the writing procedure, it is necessary to pin its magnetization. The way to achieve this task is an important aspect of the technology, however it is not relevant for the scope of this project. In the following we will assume the magnetization of FM1 fixed without further details and we will focus our attention on the second ferromagnet.

Suppose to have the magnetization pointing up for FM1, then a current coming from the left as in Figure 1.2 will be spin-up polarized. If the junction is in anti-parallel configuration (FM2 magnetization pointing down), the current exert a torque on FM2, whose effect is to bring it parallel to FM1. On the other hand, in a parallel configuration, it is a current flowing in opposite direction to flip FM2. The unpolarized current coming from the right side of Figure 1.2 exchanges momentum with the free layer FM2 and while spin-up electrons go through the junction, the spin-down component is reflected by FM1 adding torque, whose effect is to rotate FM2 anti-parallel respect to FM1. This completes the description of the operational principles of the memory.

When acting on the magnetization of FM2, the STT effect is in competition with other effects and the magnetization reversal is achieved only for a high enough current density. It is common use to define "switching current" the minimum current needed to trigger the magnetization flip. The value for the switching current of a particular junction is the central quantity needed to evaluate the writing performance of the memory and, as mentioned before, it depends on other factors beside the strength of the STT effect. A theoretical understanding can be achieved by looking at the magnetic dynamics in the Landau-Lifshitz-Gilbert (LLG) formalism [22, 23], as explained in Section 1.4. MTJs with the lowest possible switching current might seem ideal for the STT-MRAMs, however, the thermal stability of the magnetization is another important factor. Section 1.4 will clarify the competition between this two phenomena. We can in fact anticipate that the main

challenge of the STT-MRAM technology is the possibility to achieve high thermal stability in reduced dimensions, low switching current and high TMR ratio all at the same time.

Final important mention is a consideration on the geometry of the MTJ for STT-MRAMs. In Figure 1.1 and Figure 1.2 we presented the situation where the current is flowing perpendicular to the magnetization of the ferromagnets. Moreover the materials constituting the stack are in thin film shape. For this reason it is common to call this configuration "in-plane". This configuration was especially chosen for the good thermal stability, but it is not advantageous for energy efficient switching (see Section 1.4). Research is now moving to the so called "out-of-plane" (or "perpendicular") geometry, where the current is parallel to magnetization and the required switching current is lower. A schematic representation of the two geometries is in Figure 1.3.

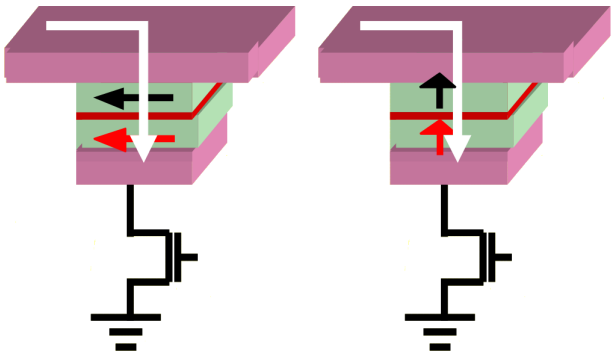


Figure 1.3: Schematic representation of the two possible STT-MRAM geometries. On the left the "in-plane" geometry, on the right the "out-of-plane". The white arrow indicates the current flow, the red arrow is the magnetization of the polarizer, the black one indicates the magnetization of the free layer. Picture from [24].

1.2 Tunnel Magneto-Resistance and Symmetry Filtering Effect

A preliminary understanding of the origin of the TMR can be obtained just by looking at the spin-dependent density of state of the two ferromagnets (Figure 1.4). An electron incident into a ferromagnet has different scattering probability, whether its spin is up or down with respect to the local direction of magnetization. As a consequence, the current passing through this material will be spin polarized. When the two ferromagnets are in the same magnetic configuration (*i.e.* their magnetization vectors are parallel to each other), they will both have high probability to let one spin type (let us say down) to pass. Thus the down spin polarized current will short out the junction resulting in a fairly low resistance. If the two ferromagnets are in an anti-parallel configuration, both spin up and down will be scattered strongly in either one layer or the other. In this picture, the role of the insulator is only to decouple the two ferromagnets and set the energy reference.

This simple model, however, is not enough to explain the extremely large TMR ratios recorded in Fe/MgO/Fe junctions [26]. The peculiar electronic structure of both the magnet and the insulator has been suggested to be the main origin of the high TMR ratios

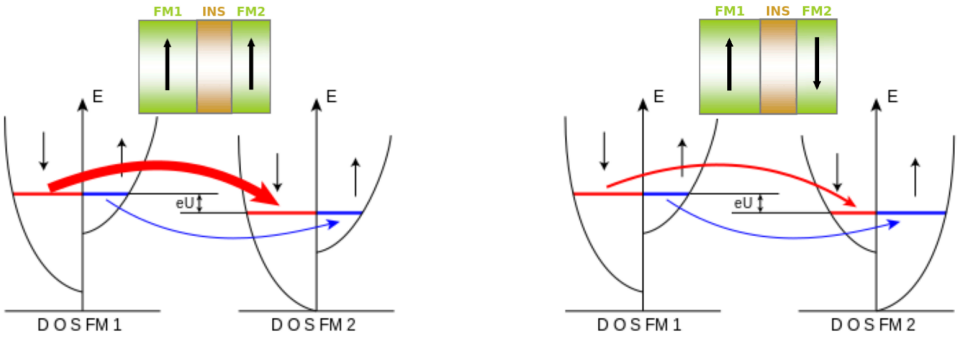


Figure 1.4: Spin filtering effect in a ferromagnetic junction. In the parallel configuration (left-hand side panel), spin down electrons have more states available for scattering in both ferromagnets. In presence of a small electric potential U, a down spin polarized current (red arrow) will short the junction out. In the anti-parallel configuration (right-hand side) both spin up and spin down are scattered strongly in either one layer or the other. Picture adapted from reference [25].

[27, 28, 29]. When the Fe/MgO/Fe junction is epitaxially grown along the [100] crystalline orientation a symmetry filtering effect takes place: the electron transport across the junction is determined by the symmetry matching of the wave function of the magnet with that of the insulator. In fact, the conservation of the transverse component of the electron wave vector imposes that, during the tunneling process, a wave function conserves its rotational symmetry in the plane perpendicular to the direction of propagation of the electrons. For Fe/MgO the key feature is that in MgO the electronic states with Δ_1 symmetry decay at a slower rate than those presenting other symmetries. As a consequence, for thick enough barriers, only Δ_1 electrons tunnel through the MgO with appreciable amplitude and luckily in Fe only majority spins occupy the Δ_1 states at the Fermi level. In other words, the electrons that dominate the transport in the tunneling process, have states to couple with only for majority spin. As a consequence the current in the junction becomes fully spin-polarized when the ferromagnets are in the parallel configuration, while in the anti-parallel configuration the current is substantially zero.

The symmetry filtering process just described can be understood by looking at the Complex Band Structure (CBS) of MgO (panel (b) of Figure 1.5) and the spin-dependent band structure of Fe (panels (c) and (d) of Figure 1.5). The CBS of a material describes for each energy the allowed complex wave vectors. Since the imaginary part of a wave vector, κ, determines the evanescent states of a crystal, the CBS has an important role in tunneling transport. In fact, the damping coefficient of the wave-function is proportional to κ in the insulator. For instance, in the simple model of the potential barrier of thickness a, it can be proven that the transmission coefficient is proportional to $\exp(-2\kappa a)$; this is presented in Appendix A. Each complex band has a symmetry. From Figure 1.5(b) it can be seen that Δ_1 states present the smallest κ. Therefore these states decay at a slower rate compared to others and they dominate the transport. Figure 1.5(c) and Figure 1.5(d) show the symmetry of the bands in Fe. As mentioned before, no state with Δ_1 symmetry are present in the minority channel, leading to the spin symmetry filtering.

From this picture it is clear that the performance of a junction depends on both the

magnet and the barrier and both need to be engineered at the same time. It is also clear that the CBS of the insulator plays a relevant role, together with the symmetry of the wave functions at the Fermi level of the ferromagnet.

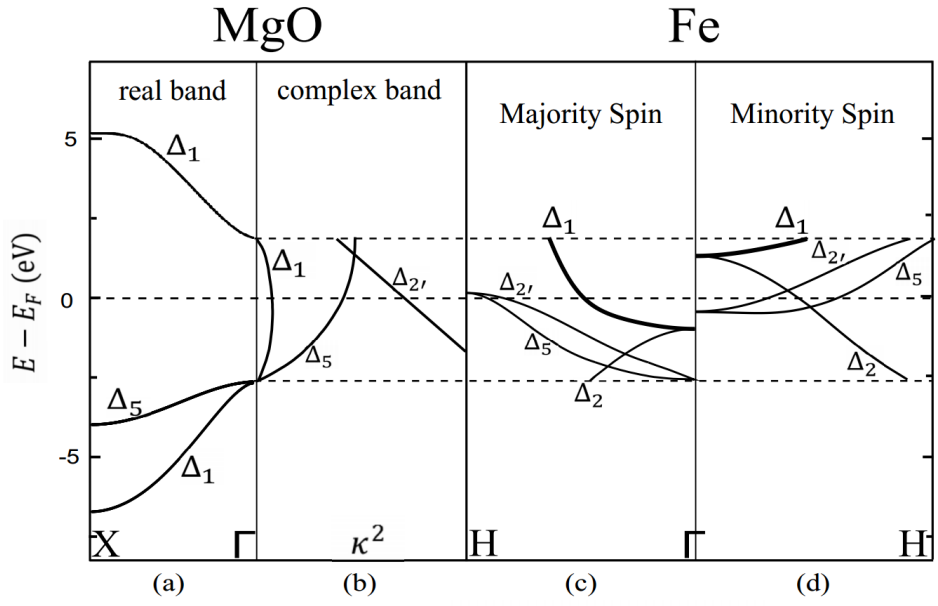


Figure 1.5: (a) Real (from X to Γ) and (b) complex (at Γ) band structure of MgO and (c-d) spin dependent band structure of Fe. The imaginary part of the wave vector, κ^2, of the complex band structure is proportional to the damping coefficient of the wave functions. The Δ_1 symmetry wave functions dominate the transport. Δ_1 states are available only for majority spin in Fe at the Fermi Energy (E_F). Picture taken from chapter 4 of reference [30].

A direct calculation of the TMR ratio,

$$TMR_{ratio} = \frac{R_{AP} - R_P}{R_P} = \frac{I_{AP} - I_P}{I_P}, \tag{1.1}$$

in DFT is possible with the addition of a quantum transport theory that allows to estimate the electric current of the entire junction in the parallel (I_P) and anti-parallel (I_{AP}) configuration. This requires, for example, the use of Green's function methods in order to calculate the transmission probability of the two configurations at a quantum level and consequently the currents.

As a final remark, we note that the TMR_{ratio} depends on the applied voltage [26] and it is maximized for low bias.

1.3 Spin-Transfer-Torque Effect

When approaching the study of the STT effect, the quantity of interest is the strength of the torque acting on the magnetization of the analyzer (FM2 in Fig 1.2) due to the spin polarized current. The origin of this torque can be understood as a simple consequence of angular momentum conservation. When a current flows through a ferromagnet, the electrons acquire spin polarization. The formation of the spin-polarized current is the result of the torque that the ferromagnet magnetization exerts on the current. Momentum

conservation establishes that an opposite and equal torque will act on the ferromagnet due to the current. This is the STT effect.

1.3.1 Toy model

In order to have few more insights about the microscopic origin of STT and its qualitative behavior, we report the relevant part of the study of a toy model done by Stiles and Zangwill [31]. The exercise consists in looking at the scattering of one single spin-polarized electron going from a non magnetic material (NMM) to a ferromagnet (FM) with uniform magnetization in direction $\hat{\mathbf{z}}$ (Figure 1.6). The surface is planar with no defects and perpendicular to $\hat{\mathbf{x}}$. The polarization of the electron is arbitrary and it is defined by the angles θ and ϕ with respect to the direction $\hat{\mathbf{z}}$ (see Figure 1.6).

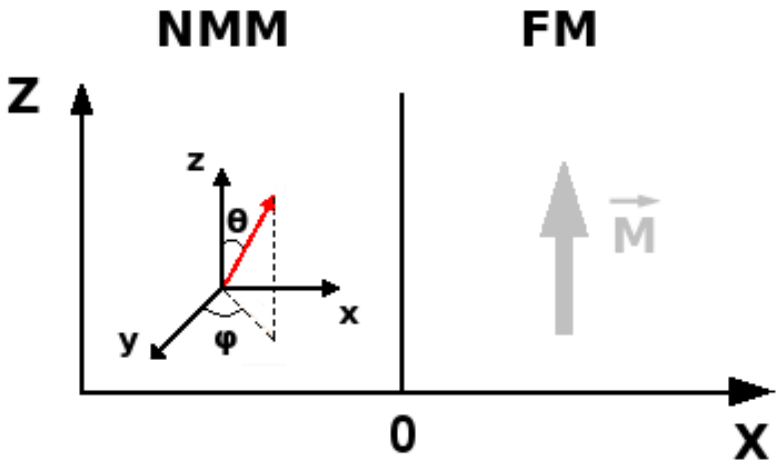


Figure 1.6: Interface between a non magnetic material (NMM) and a ferromagnet (FM). In gray we represent the uniform magnetization of the FM, that points in $\hat{\mathbf{z}}$ direction. The red arrow is the spin polarization of an incident electron, whose direction is defined by the angles θ and ψ.

The incident wave function is linear combination of up and down states and, assuming a free-electron description is adequate, it writes:

$$\psi_{in} = [\cos(\theta/2)e^{-i\phi/2}\,|\uparrow\rangle + \sin(\theta/2)e^{i\phi/2}\,|\downarrow\rangle]e^{ik_x x}e^{i\mathbf{qR}}, \tag{1.2}$$

with $\mathbf{r} = (x, \mathbf{R})$ and the wave vector $\mathbf{k} = (k_x, \mathbf{q})$ satisfying $|\mathbf{k}|^2\hbar/2m = E_F$ (Fermi energy) We define in the same way the reflected and transmitted wave functions:

$$\psi_{re} = [\cos(\theta/2)e^{-i\phi/2}R_\uparrow\,|\uparrow\rangle + \sin(\theta/2)e^{i\phi/2}R_\downarrow\,|\downarrow\rangle]e^{ik_x x}e^{i\mathbf{qR}}, \tag{1.3}$$

$$\psi_{tr} = [\cos(\theta/2)e^{-i\phi/2}T_\uparrow e^{ik_x^\uparrow x}\,|\uparrow\rangle + \sin(\theta/2)e^{i\phi/2}T_\downarrow e^{ik_x^\downarrow x}\,|\downarrow\rangle]e^{i\mathbf{qR}}. \tag{1.4}$$

While ψ_{in} and ψ_{re} are defined in the NMM, ψ_{tr} belongs to the FM. This explains why the two spin components propagate with two different wave vectors, k_x^σ. In fact their kinetic energy depends on the exchange potential energy in the ferromagnet. The component of the wave vector parallel to the transport direction, $\mathbf{q}$, is conserved during the scattering process. The scalars $R_\uparrow$, $R_\downarrow$, $T_\uparrow$ and $T_\downarrow$ are the reflection and transmission coefficients for

the spin up and down components. They are **q**-dependent but this is not relevant for the moment.

The electron propagates along $\hat{\mathbf{x}}$ with arbitrary spin polarization. It is interesting to calculate the current density along the $\hat{\mathbf{x}}$ direction,

$$j_x = \frac{-i\hbar}{m}\mathfrak{Re}\left[\psi^*(\mathbf{r})\frac{\partial}{\partial x}\psi(\mathbf{r})\right], \tag{1.5}$$

but even more interesting is the spin current density,

$$\mathbf{Q_x} = \frac{\hbar}{2}\mathfrak{Re}\left[\psi^*(\mathbf{r})\boldsymbol{\sigma}\frac{\partial}{\partial x}\psi(\mathbf{r})\right], \tag{1.6}$$

where $\boldsymbol{\sigma}$ is the Pauli matrices vector and $\mathbf{Q_x}$ is a vector in spin space. In fact $\mathbf{Q}$ is a tensor quantity, whose components Q_{ij} have the left index labeling the real space coordinates, while the right one labels the spin space. We are interested in electrons flowing along $\hat{\mathbf{x}}$ so in Eq. (1.6) we report only the x component in real space. When this is calculated for the incident, reflected and transmitted wave function, Eq. (1.6) gives:

$$\begin{aligned} Q^{in}_{xx} =& \frac{\hbar}{2} j^{in}_x \sin(\theta)\cos(\phi), \\ Q^{in}_{xz} =& \frac{\hbar}{2} j^{in}_x \cos(\theta), \end{aligned} \tag{1.7}$$

$$\begin{aligned} Q^{re}_{xx} =& -\frac{\hbar}{4} j^{in}_x \sin(\theta)\mathfrak{Re}[R^*_\uparrow R_\downarrow e^{i\phi}], \\ Q^{re}_{xz} =& -\frac{\hbar}{2} j^{in}_x [\cos^2(\theta/2)|R_\uparrow|^2 - \sin^2(\theta/2)|R_\downarrow|^2], \end{aligned} \tag{1.8}$$

$$\begin{aligned} Q^{tr}_{xx} =& \frac{\hbar}{4}\frac{k^\uparrow_x + k^\downarrow_x}{2k_x} j^{in}_x \sin(\theta)\mathfrak{Re}[T^*_\uparrow T_\downarrow e^{i\phi} e^{i(k^\downarrow_x - k^\uparrow_x)x}], \\ Q^{tr}_{xz} =& \frac{\hbar}{2k_x} j^{in}_x [k^\uparrow_x \cos^2(\theta/2)|T_\uparrow|^2 - k^\downarrow_x \sin^2(\theta/2)|T_\downarrow|^2]. \end{aligned} \tag{1.9}$$

Here, we omit the expression for the spin density polarized along $\hat{\mathbf{y}}$ as it is not necessary for the discussion. Its behavior is similar to that along $\hat{\mathbf{x}}$.

First we notice that $Q^{in}_{xz} = Q^{tr}_{xz} - Q^{re}_{xz}$ (it can be demonstrated by using the fact that $k^\sigma|T_\sigma|^2/k + |R_\sigma|^2 = 1$), while the same can not be said for the x component. For this latter direction of polarization there is a discontinuity in the spin current density passing from the NMM to the FM. This spin-current deficiency must be absorbed by the magnetization of the ferromagnet, a fact that defines the origin of the spin transfer torque. The concept just described is central in the STT theory and it was demonstrated in many ways [32, 21]: every time there is a discontinuity in the spin current density, a torque is generated. This situation is often associated to the idea that the torque rises anytime there is spin accumulation. In our model we can define the torques:

$$\mathbf{N^c} = (\mathbf{Q^{in}_x} - \mathbf{Q^{tr}_x} + \mathbf{Q^{re}_x})A, \tag{1.10}$$

where $\mathbf{N^c}$ is a vector in spin space and A is the area of the interface.

We saw that in the $\hat{\mathbf{z}}$ direction there is no spin accumulation, therefore there is no torque. This fact establishes that, if the polarization of the current is parallel/anti-parallel to the magnetization of the ferromagnet, there is no torque. Such observation seems to clash with the design of the memory described in Section 1.1, where one starts from parallel/anti-parallel configuration of FM1 e FM2 (the polarization of the current is due to the polarizer FM1). In reality, at finite temperature, the magnetizations vectors are never exactly parallel: thermal fluctuations are present and they provide the necessary mis-orientation to initiate the phenomenon.

By using Eq. (1.10) an explicit expression for the torque along $\hat{\mathbf{x}}$ can be derived and used to continue the discussion:

$$N_x^c = \frac{\hbar}{2} j_x^{in} \sin(\theta) \left[\cos(\phi) - \frac{k_x^{\uparrow} + k_x^{\downarrow}}{4k_x} \mathfrak{Re}[T_{\uparrow}^* T_{\downarrow} e^{i\phi} e^{i(k_x^{\downarrow} - k_x^{\uparrow})x}] - \frac{1}{2} \mathfrak{Re}[R_{\uparrow}^* R_{\downarrow} e^{i\phi}] \right]. \tag{1.11}$$

Let us summarize what we have learnt by looking at the formulas:

- Considering that θ and ϕ define the polarization of the incident electron with respect to the magnetization vector of the ferromagnet, as already mentioned, $\theta = 0$ and $\phi = 0$ leads to zero torque. A polarization at 90 degrees instead gives the maximum value of the torque.

- Spin-dependent transmission and reflection is necessary for the torque to develop. If $R_{\uparrow} = R_{\downarrow}$, $T_{\uparrow} = T_{\downarrow}$ and $k_x^{\uparrow} = k_x^{\downarrow}$ there is no discontinuity in the spin current and no torque. This could be seen directly from the wave functions in Eq. (1.3) and Eq. (1.4) that would have the same linear combination of spin vectors of that in Eq. (1.2) in case of spin independent transmission and reflection coefficients. However it can be proven also plugging $T_{\uparrow} = T_{\downarrow} = \tilde{T}$, $R_{\uparrow} = R_{\downarrow} = \tilde{R}$ and $k_x^{\uparrow} = k_x^{\downarrow} = k_x$ in Eq. (1.11) and remembering that $\tilde{T}^2 + \tilde{R}^2 = 1$.

- The torque depends linearly on the incident current density j_x^{in}.

- There is a spacial dependence of the torque originating from $\exp i(k_x^{\downarrow} - k_x^{\uparrow})x$. In the discussion so far it is implicit that the effect of the torque is relevant in the vicinity of the interface between the NNM and the FM as it is the surface discontinuity that gives rise to the spin accumulation. However, the spatial dependence of the torque seems to suggest that this acts in an oscillatory way in the entire space $x > 0$. This needs to be discussed in more details. The toy system studied so far is very simple, just one electron is considered. An extension to a distribution of electrons is presented in reference [31]. The main difference is the fact that the integration in $\mathbf{q}$ space brings to the self cancellation of various oscillatory contributions. It can be proven that a $1/x$ decay is expected [33]. Therefore for a distribution of electrons we obtain:

$$Q_{xx}^{tr} \mapsto \alpha \frac{\sin(k_x^{\uparrow}(E_F) - k_x^{\downarrow}(E_F))x}{(k_x^{\uparrow}(E_F) - k_x^{\downarrow}(E_F))x} \qquad \text{for} \quad x \mapsto \infty \tag{1.12}$$

and N_x^c follows the same dependence. As such, the main conclusion is that a spacial

dependence regulated by $k_x^{\uparrow}(E_F) - k_x^{\downarrow}(E_F)$ is predicted by the model, while other effects should be limited to the close vicinity of the surface.

- Pre-factor of the oscillatory part is the product of spin up and down transmission coefficients, that then influences the amplitude of the damped oscillations. In particular, if either $T_{\uparrow}$ or $T_{\downarrow}$ are zero, the torque is confined at the interface.

If one excludes the presence of other external torques and ignores spin-flip processes, the expression for the torque derived here is essentially the only contribution entering the continuity equation for the magnetization. There is, however, a small caveat. Here we have always discussed the current spin density and its discontinuity as a change in the spin angular momentum of the electrons, this means that the conservation of angular momentum has an effect on the total angular momentum of the ferromagnet. Angular momentum and magnetic moment are connected by $g\mu_B/\hbar$, where g is the g-factor of the ferromagnet and μ_B is the Bohr magneton. If this is applied to the magnetic moment ($\mathbf{m}$) of the ferromagnet, we will obtain:

$$\frac{\partial \mathbf{m}}{\partial t} = \frac{g\mu_B}{\hbar}\mathbf{N^c}. \tag{1.13}$$

Moreover the magnetization, $\mathbf{M}$ is $\mathbf{m}/V$, where V is the volume of a small portion of ferromagnet just after the interface with the non magnetic material. Therefore we obtain:

$$\frac{\partial \mathbf{M}}{\partial t} = \frac{g\mu_B}{\hbar V}\mathbf{N^c} \tag{1.14}$$

More details about the limits of validity of the formulas introduced above can be found in the review [21].

1.3.2 Modeling *ab-initio*

In real materials one expects the magnetic moment not to be uniform in space. Hence the torque is also expected to act differently for each point in space. Moreover, several other complications such as the symmetry of the wave function should enter in the discussion when approaching the study of an MTJ. Nevertheless, many insights we got from the very simple model described above have been proven correct at least for simulations of spin transfer torque in a few transition metals and MgO junctions.

For example in Figure 1.7 we present a calculation of the atom resolved torkance (torque in the limit of zero bias) for three MTJs (TM/MgO/TM with TM=Fe,Co,Ni in the body centered cubic structure) with interfaces perpendicular to the [100] direction of the crystal and where the polarization of the current is kept at 90° with respect to the magnetization ($\phi = 0$, $\theta = 90$ in the toy model of the previous subsection). The calculation has been performed with the code SMEAGOL [34], that combines DFT and the non-equilibrium Green's function formalism to calculate transport properties and the STT effect. For Fe and Co the STT effect is confined in the vicinity of the interface. This is in agreement with the simple model since in these two junctions the transmission of the minority channel is practically zero ($T_{\downarrow} \sim 0$) due to the symmetry filtering effect explained in Section 1.2. As

the spacial dependent contribution to the torque scales as $T_{\uparrow}^{*}T_{\downarrow}$ [see Eq. 1.11] one expects it to be small for Fe and Co. In the case of Ni, instead, a clear oscillatory behavior is calculated for the torque. In fact, by looking at the electronic bands of bulk Ni (Figure 1.8), the Δ_1 states are present at the Fermi energy for both minority and majority spin channels, leading to a non-zero value of $T_{\uparrow}^{*}T_{\downarrow}$. According to the model, the frequency of the oscillation is determined by $k_x^{\uparrow}(E_F) - k_x^{\downarrow}(E_F)$ and from Figure 1.8 we estimated $k_x^{\uparrow}(E_F) - k_x^{\downarrow}(E_F) = 0.45$ Å^{-1}. A fit with a function of the form in Eq. (1.12) to the data in Figure 1.7 reveals a frequency of 0.19 Å^{-1}. Only values of the torkance with position > 4 Å were considered in the fit in order to exclude interfacial effects. The mismatch between the value obtained with the fit and the value of $k_x^{\uparrow}(E_F) - k_x^{\downarrow}(E_F)$ from Figure 1.8 could be due to the different position of the Fermi energy inside the junction with respect to that of pure bulk Ni. In fact, considering in Figure 1.8 an energy $E - E_F = 0.4$ eV is sufficient to have $k_x^{\uparrow}(E_F) - k_x^{\downarrow}(E_F) = 0.19$ Å^{-1}, the same value obtained in the fit.

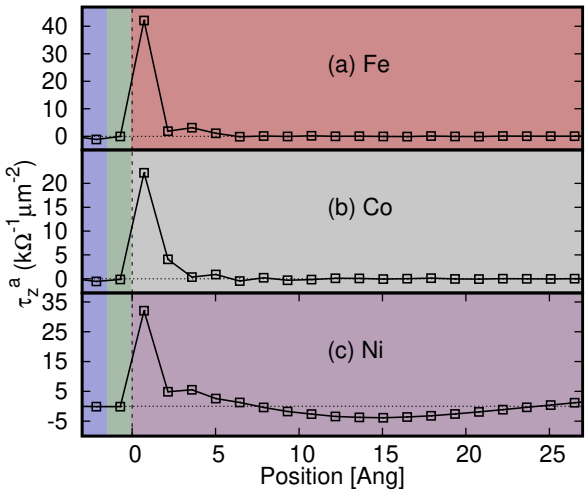


Figure 1.7: Torkance of the MTJs TM/MgO/TM, with a) TM=Fe, b) TM=Co, c) TM=Ni. All the transition metals are taken with a body centered cubic structure. The blue areas indicates O atoms, the green Mg atoms. The torkance, in unit of $\mu_B/2$ and area, has been calculated through SMEAGOL, courtesy of Mario Galante.

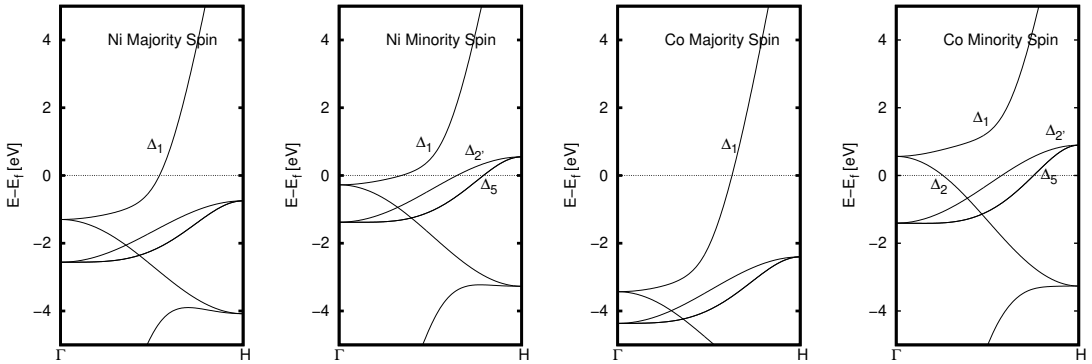


Figure 1.8: Band structure along the high symmetry direction Γ−H for body centered cubic Ni and Co. The band structure of Fe is similar to the one of Co and it is reported in Figure 1.5.

It is not the scope of this subsection to make a quantitative discussion of the STT

in these junctions, that should consider for example the fact that the moment of atoms at the interface is different with respect to the one of bulk atoms. The aim is to prove the reliability of the information gathered with the toy model in subsection 1.3.1 for the interpretation of *ab-initio* calculations performed on simple systems. These information are used in the following of the thesis.

Here we presented STT calculations from the code SMEAGOL. For the sake of completeness, it is important to mention that other theories to obtain the torque *ab-initio* are present in the literature [32, 35].

1.3.3 Phenomenological models

Macroscopic phenomenological expressions for the torque are also in use in the community and they agree to some extent to the theory so far outlined. For instance an expression for the torque acting on the magnetization of the analyzer in a MTJ [24] is:

$$\frac{\partial \mathbf{M_2}}{\partial t} = \frac{g\mu_B}{\hbar l}\frac{I}{e}P_{spin}(\hat{\mathbf{m}}_\mathbf{2} \times (\hat{\mathbf{m}}_\mathbf{2} \times \hat{\mathbf{m}}_\mathbf{1})) + \tau_{op}(\hat{\mathbf{m}}_\mathbf{2} \times \hat{\mathbf{m}}_\mathbf{1}). \tag{1.15}$$

Here we are using a notation that relates to Figure 1.2. $\mathbf{M_2}$ is the magnetization of FM2, the layer in the MTJ whose magnetization is free to rotate. $\hat{\mathbf{m}}_\mathbf{1}$ is the direction of the magnetization of FM1, the polarizer. Assuming that the magnitude of $\mathbf{M_2}$ does not change in time, it is possible to divide Eq. (1.15) for its module, M_2, and obtain an equation for the versor $\hat{\mathbf{m}}_\mathbf{2} = \mathbf{M_2}/M_2$:

$$\frac{\partial \hat{\mathbf{m}}_\mathbf{2}}{\partial t} = \frac{g\mu_B}{\hbar l M_2}\frac{I}{e}P_{spin}(\hat{\mathbf{m}}_\mathbf{2} \times (\hat{\mathbf{m}}_\mathbf{2} \times \hat{\mathbf{m}}_\mathbf{1})) + \tau_{op}(\hat{\mathbf{m}}_\mathbf{2} \times \hat{\mathbf{m}}_\mathbf{1}). \tag{1.16}$$

The scalar I is the total current, l is the thickness of the layer and P_{spin} is the polarization of the incident current. This latter quantity essentially measures the fraction of electrons that are effectively polarized by FM1 and can be evaluated as,

$$P_{spin} = \frac{D^{\uparrow}_{FM1}(E_F) - D^{\downarrow}_{FM1}(E_F)}{D^{\uparrow}_{FM1}(E_F) + D^{\downarrow}_{FM1}(E_F)}, \tag{1.17}$$

where $D^{\sigma}_{FM1}(E_F)$ is the density of state at the Fermi energy for spin σ. The angular dependence is now recasted in two components. The first one describes the torque acting in the plane defined by $\hat{\mathbf{m}}_\mathbf{1}$ (direction of the spin polarized current, imposed by the magnetization of FM1) and $\hat{\mathbf{m}}_\mathbf{2}$ (direction of the magnetization in FM2) and it is often referred to as "in-plane torque". The second components makes $\hat{\mathbf{m}}_\mathbf{2}$ to precess around $\hat{\mathbf{m}}_\mathbf{1}$ and it's called "out-of-plane torque". For a better understanding we can relate the two components to the toy model of subsection 1.3.1 when we restrict ourself to $\phi = 0$. This means that the polarization of the incident spin is in the plane defined by $\hat{\mathbf{x}}$ and $\hat{\mathbf{z}}$. In that case the "in-plane" part corresponds exactly to the N_x^c of Eq. (1.11), while the "out-of-plane" would be N_y^c, which was not explicitly derived. The "out-of-plane" component of the torque plays a minor role in the STT technology as clarified in the next section. For that reason we did not explicitly wrote any pre-factor in Eq. (1.16). The macroscopic expression of the

STT is useful since it provides a better understanding of the quantities to control in a real experiment, however it does not describe, for example, the small scale space dependence of the STT.

1.4 Magnetic Dynamics

Once the spin transfer torque has been generated thanks to the spin current, a study of the magnetic dynamics is needed to understand whether magnetization reversal will effectively take place. As in the previous section, we assume that the magnitude of the magnetization does not change in time and space. It is possible to write the equation of motion for the magnetization versor $\hat{\mathbf{m}}_\mathbf{2}$ in the Landau-Lifshitz-Gilbert formalism [22, 23] as

$$\frac{\partial \hat{\mathbf{m}}_\mathbf{2}}{\partial t} = -\gamma[\hat{\mathbf{m}}_\mathbf{2} \times \mathbf{H}_\mathbf{eff}] + \alpha[\hat{\mathbf{m}}_\mathbf{2} \times \frac{\partial \hat{\mathbf{m}}_\mathbf{2}}{\partial t}] + \tau_{ip} I(\hat{\mathbf{m}}_\mathbf{2} \times (\hat{\mathbf{m}}_\mathbf{2} \times \hat{\mathbf{m}}_\mathbf{1})). \tag{1.18}$$

Again the notation refers to Figure 1.2 and the definitions of $\hat{\mathbf{m}}_\mathbf{1}$ and $\hat{\mathbf{m}}_\mathbf{2}$ were given in the previous section. We recognize in the last term of Eq. 1.18 the "in-plane" torque. The level of theory at which we calculate τ_{ip} is now unimportant. For instance, it could be a spatial average from a very sophisticated DFT calculations, or the very rough estimate of the pre-factor in Eq. (1.16). We just keep explicit the linear dependence with the current I as it will be useful afterwards. The second term on the right-side of Eq. (1.18) is a phenomenological damping term called Gilbert damping and α is the parameter quantifying its magnitude. The factor γ is the gyromagnetic ratio for an elecron spin. The effective field, $\mathbf{H}_\mathbf{eff}$, is the instantaneous local field and comprieses many contributions [23]. In absence of an external magnetic field, the most important contribution is the anisotropy field, defined as:

$$\mathbf{H}_\mathbf{eff} = \mathbf{H}_\mathbf{a} = \frac{2k_u}{M_s}\mathbf{e}. \tag{1.19}$$

Here k_u is the anisotropy constant (usually expressed in units of $\mathrm{J/m^3}$) for a field pointing in direction $\mathbf{e}$ (uniaxial anisotropy) and M_s is the saturation magnetization (in units of Tesla). Another possible contribution to $\mathbf{H}_\mathbf{eff}$ is the so called demagnetizing field, the magnetic field generated by the magnetization itself. Few more insights on this contribution follow at the end of this section.

In the absence on any torque or damping the magnetization versor would precess around $\mathbf{H}_\mathbf{eff}$. However the Gilbert damping is necessary in the theory in order to account for energy losses deriving for instance from the interaction of the magnetization with the lattice. The damping is in competition with the Spin-Transfer-Torque. In fact, it can be proven that $\hat{\mathbf{m}}_\mathbf{2} \times \partial\hat{\mathbf{m}}_\mathbf{2}/\partial t$ has the same direction of $\hat{\mathbf{m}}_\mathbf{2} \times (\hat{\mathbf{m}}_\mathbf{2} \times \hat{\mathbf{m}}_\mathbf{1})$, but opposite sign [36].

The out-of-plane contribution to the torque discussed in subsection 1.3.3 is not included in Eq. (1.18). Because at the beginning of the dynamics $\hat{\mathbf{m}}_\mathbf{1}$ is parallel to $\mathbf{H}_\mathbf{eff}$, its effect would be similar to the one of $\mathbf{H}_\mathbf{eff}$, however smaller in magnitude. Since its relevance for the following of the discussion is marginal, it was not included in Eq.(1.18).

From Eq. (1.18) it is clear that the effective magnetization reversal depends not only on the STT, but on all the other factors entering the dynamic. The ultimate quantity of

interest is the switching current, *i.e.* the critical current intensity needed for the STT to overcome the damping and induce magnetization reversal. An explicit expression for the switching current can be derived by considering that the largest energetic barrier to overcome is when the two magnetic layer are in (almost) parallel/anti-parallel configuration. In other words, we are considering the situation when a small angle separates $\hat{\mathbf{m}}_\mathbf{1}$ and $\hat{\mathbf{m}}_\mathbf{2}$. In that case $\mathbf{H_{eff}}$ is parallel to $\hat{\mathbf{m}}_\mathbf{1}$ and one can approximate:

$$\frac{\partial \hat{\mathbf{m}}_\mathbf{2}}{\partial t} = -\gamma[\hat{\mathbf{m}}_\mathbf{2} \times \mathbf{H_{eff}}] + \alpha[\hat{\mathbf{m}}_\mathbf{2} \times (-\gamma[\hat{\mathbf{m}}_\mathbf{2} \times \mathbf{H_{eff}}])] + \tau_{ip} I(\hat{\mathbf{m}}_\mathbf{2} \times (\hat{\mathbf{m}}_\mathbf{2} \times \frac{\mathbf{H_{eff}}}{|\mathbf{H_{eff}}|})). \quad (1.20)$$

The switching current, I_s, is then the minimum current required to make the STT overcoming the Gilbert damping, namely:

$$I_s = \frac{\gamma \alpha |\mathbf{H_{eff}}|}{\tau_{ip}} \quad (1.21)$$

The quantity γ is a constant, not related to the choice of the material set. The switching current depends on the ratio between the Gilbert damping parameter and the strength of the torque, but also the absolute value of the effective field matters. One would say that the magnetic hardness of the ferromagnet counts.

Notably, the main contribution to $\mathbf{H_{eff}}$, meaning the anisotropy, is responsible for the thermal stability of the system. It sets the hight of the energetic barrier to overcome to switch from a parallel to an anti-parallel configuration. It is common to define the stability ratio:

$$E_b = \frac{k_u V}{k_B T} \quad (1.22)$$

as a measure of the thermal stability of the junction, being $\tau = \tau_0 \exp(E_b)$ the relaxation time for the switching probability, V the volume of the ferromagnet, T the temperature and k_B the Boltzmann constant. A barrier E_b in excess of 60 is needed to give a thermal stability for more than 10 years at room temperature. Increasing the anisotropy is the obvious way to increase E_b, however, according to Eq. (1.21) and Eq. (1.19), it leads to a higher switching current.

It is now worth analyzing in more details the origin of the anisotropy. In fact, it is the reason why the "perpendicular" geometry of an STT-MRAM is favorable respect to the "in-plane" one. Every bulk material has an intrinsic anisotropy called magnetocristalline anisotropy [15]. However in MTJs other contributions are relevant. In the "in-plane" geometry, the so called shape anisotropy, k_{sp}, is the most important. The thin-film shape of the ferromagnets tends to keep the magnetization in the film plane. Moreover, the films are built in disc shape with an elliptical form. The direction along which the disk is elongated is the preferred direction for the magnetization. Therefore, in the "in-plane" geometry, the shape of the sample is determining the relevant anisotropy that defines the thermal stability. In formula:

$$E_b = \frac{k_{sp} V}{k_B T} \quad (1.23)$$

On the other hand, it can be proved that for a thin film, when the magnetization is in

plane, the demagnetisation energy is minimized, but the magnetization vector needs to go out of plane for magnetization reversal. In this situation, an extra current is needed to account for the energy penalty of going away from the situation of minimum demagnetizing energy. The expression in Eq. (1.21) must include a demagnetizing field contribution, H_d:

$$I_{ip} = \frac{\gamma\alpha}{\tau_{ip}} \left(\frac{2k_{sp}}{M_s} + H_d \right) \tag{1.24}$$

The quantity H_d increases the critical current, but it does not contribute to the energy barrier for the thermal stability. The situation is different for a "perpendicular" geometry. In the case of CoFeB, for instance, for films with a thickness below 2 nm, the anisotropy spontaneusly turns out of plane due to the interfacial effects (interfacial anisotropy - k_{in}) [37, 38]. This allows stability of the magnetization in direction perpendicular to the plane of the thin film. In this situation, interfacial effects determine the thermal stability, not the sample geometry as in the "in-plane" case. We write:

$$E_b = \frac{k_{in}V}{k_BT} \tag{1.25}$$

A demagnetizing field is present here in the stable situation, but it is already overcame by the interfacial anysotropy and it does not add any extra term to the current. Therefore in the "out-of-plane" geometry one has:

$$I_{oop} = \frac{\gamma\alpha(2k_{in})}{\tau_{ip}M_s}. \tag{1.26}$$

To conclude, comparing the above expressions for the switching currents and thermal stability ratio, in the "perpendicular" situation, the current and E_b are both proportional to k_{in}. In the "in-plane" geometry, instead, an extra current due to the demagnetizing field is present. The phenomenology just described explains why the "perpendicular" geometry is the one of choice for STT-MRAMs and how the role of the anisotropy is important is for the entire technology. We also learnt the origin of the competition between high thermal stability and low switching current.

Experimental findings show that for small pillar geometries, the reduction of the lateral dimension of the memory brings to a rapid reduction of the interfacial anisotropy (the volume V enter explicitly in the expression for E_b, but also k_{in} depends on the dimensions and shape of the ferromagnets), which means a reduction of both the switching current and thermal stability. Keeping a high thermal stability in low dimension is one of the main challenge for STT-based devices.

The picture presented in the section is a macroscopic view of the problem, dealing with the dynamics of the total magnetization vector of the ferromagnet. More sophisticated studies could be performed by looking at spacial dependent theories. For example an equation similar to Eq. (1.18) can be written for the spin moment of each atom in the system. In that case an expression of the STT acting on each atom can be used and the effective field could include atomic effects like an explicit treatment of the exchange

interaction between atoms or other long range magnetostatic interactions. The switching would start from the interface and a strong exchange coupling between atoms would help the spin moments to rotate uniformly. However, a strong exchange on the surface layer would hamper the switching. The situation is, in general, complicated; the switching dynamics can be simulated using semi-empirical models describing atomic spin dynamics. One implementation is contained in the code VAMPIRE [39]. Having reliable input parameters is essential for all the semi-empirical models and *ab-initio* theories are a powerful tool when experimental data are not available. In the particular case of STT-MRAMs, we saw (subsection 1.3.2) how the STT strength and his spacial dependence can be obtained from *ab-initio* theories. Importantly, also the Gilbert damping [40], k_u [41] and the exchange parameters [42] are quantities that can be computed at the DFT level.

1.5 State-of-the-art materials combination and challenges

The exceptional spin filtering performed by MgO barriers matched with bcc Fe layers made Fe/MgO/Fe the first material composition for the fabrication of high TMR MTJs. Further theoretical studies predicted even higher TMR ratios using FeCo instead of pure iron. The reason for this TMR enhancement is due to the fact that in FeCo not only the Δ_1 states are absent for minority spin at the Fermi energy (like in Fe), but the Δ_1 states are also the only ones at the Fermi level for majority spin (the same situation encountered in pure Co, Figure 1.8). In other words, compared to Fe, there is no parasitic tunneling of Δ_2' and Δ_5 symmetry electrons that decreases the resistance of the anti-parallel configuration. Experimental studies [43] showed how the alloy $Fe_{0.75}Co_{0.25}$ has the best TMR performances. Furthermore, the significantly higher Curie temperature of Co (1388 K compared to 1043 K of Fe) implies an enhanced stability of the magnetic properties against thermal fluctuations in FeCo, making the FeCo/MgO set the first real candidate for prototypical STT-MRAMs.

The high-crystallinity of the junction is a fundamental requirement to achieve symmetry filtering and high TMR ratios. Even the formation of interfacial resonant states could enhance tunneling of spin-down electrons and hamper the memory performance. Moreover possible dislocations in the growth process are a real threat for the memory endurance because they could lead to trapped electrons and, in turn, mechanical stress due to the strong electrostatic interaction with the metallic electrodes. The creation of a multi-layered structures that can be grown keeping the ferromagnet/insulator interfaces as clean as possible is a challenge in itself. State-of-the-art FeCo/MgO based-MTJs are deposited in an amorphous phase, to then obtain crystallization and epitaxy through annealing. This is possible thanks to the high stability of the MgO rock salt like structure [18] and the addition of boron atoms to Co_x Fe_{1-x} that aid the crystallization of the electrodes. However CoFeB has lower TMR performances compared to FeCo. It is important to include boron in the first stage of the fabrication, but then try to attract it away from the tunnel barrier during the crystallization process. This is achieved by placing materials that are effective at attracting B (Ta,Ru,W) on the opposite side of the ferromagnets.

TMR ratios beyond 200% were observed in $(CoFe)_{0.8}B_{0.2}/MgO/(CoFe)_{0.8}B_{0.2}$ MTJs [44] and a switching current of about 10^6 A/cm^2 was registered, although in samples with a lower TMR ratio and large dimensions (area of 80·240 nm^2 equivalent to pillars of 160 nm diameter) [45].

The achievements described so far concern the "in-plane" STT geometry. The research field got revitalized in 2008 when Nakayama *et al.* demonstrated the fabrication and switching properties of FeCoB/MgO MTJs with perpendicular anisotropy [37], obtaining switching currents of about 5 10^6 A/cm^2 and good thermal stability for relatively large junctions (diameter 130 nm). This seemed promising for a new boost in the technology. Similar studies on all metallic junctions also proved the advantages of 'out-of-plane" geometries [46]. Further improvements were observed by Ikeda *et al.* that reported the performances of a junction where the perpendicular anisotropy was so strong to ensure the pinning of the reference layer without any extra coupling with other magnetic structures [38]. The interfacial anisotropy is in fact dependent on the thickness of the film and smaller is the film, stronger is the anisotropy. An extra-thin film of CoFeB (1 nm) showed sufficient anisotropy to be used as a reference layer, while the second ferromagnet (1.7 nm) acts as free layer. The ultra-simplified design of the memory meant a great deal for the technology, however the 45 nm diameter wide memory of Ikeda *et al.* was not sufficiently thermally stable to accomplish the desired combination of properties needed for an efficient memory. Switching properties down to junction of 11 nm in diameter were reported by Sato and coworkers [47], but again the thermal stability issues rise for very small devices. In that study, E_b of 70 was possible only for devices with 30 nm diameter or larger. In that range of dimensions the switching current is not smaller than 10^6 A/cm^2.

Despite the issues just mentioned, STT-MRAMs based on perpendicular MTJs are now a commercial product. For a complete review of the current memories on the market and their features we suggest the website `https://www.mram-info.com/stt-mram`. The most recent commercial product is at the 45 nm technological node and it is not a low consumption device compared to DRAM memories. Cost and density are concerns that still need to be address in order to make STT-MRAMs suitable for mass production.

1.6 Material selection based on bulk simulations

Despite the success in the creation of a commercialized product, there is still interest in the improvement of the STT-MRAM technology to push scalability and power consumptions to better standards. Section 1.5 went into details in the description of the current material set for the memory and it is clear how basically only the MgO/FeCo(B) combination was considered during the technology development. The aim of this thesis is to calculate properties of a large variety of materials and discuss the possibility of having a new combination able to prove itself as a good candidate for a further enhancement in the memory performance. The rich theory for the STT technology described in the first sections of this chapter leads to some observations about the characteristics that the insulator and the ferromagnet need to have, independently on the stack design; their description is the focus

of this section and it constitute the motivation of the entire PhD thesis.

We start with a summary of the principal requirements for an efficient STT memory:

1. As the read-out of STT-MRAMs is via the TMR effect, the TMR ratio should be large at room temperature. This means that the junction should have an efficient symmetry filtering as explained in Section 1.2. Such feature is indeed dependent on the band matching, but a preliminary requirement is that the insulator selects a single symmetry and that the ferromagnet shows symmetry-dependent spin polarization. As the working conditions for a STT spin-valve requires the introduction of a current, a good TMR ratio should be maintained at high voltages. Wider is the range of energies where the spin symmetry filtering is present, larger the voltage range at which high TMR is expected to be.

2. It is not clear whether it is better to have a torque that extends into the bulk or it is sufficient to have torque at the interface and relay on the exchange interaction between atoms to switch the entire ferromagnet. However, we know from Eq. (1.11) that if either the majority or minority transmission is zero, the torque is confined in the vicinity of the interface. This means that a perfect symmetry filtering effect implies also a torque with no spacial dependence.

3. The anisotropy is a key factor in the entire technology. The free magnetic layer (analyzer, FM2 in Figure 1.2) should be magnetically soft enough to rotate in a current density but hard enough to be thermally stable. The two effects are in completion: a direct manipulation of the anisotropy in the ferromagnet not necessarily leads to better memories. It is important to look at other ways to diminish the switching current, bringing us to the next two points.

4. The device total impedance should be small so that the current density can be large at moderate voltages, maintaining though high TMR ratios. In an MTJ, the overall resistance is determined by the damping coefficient of the evanescent wave function in the insulator.

5. The Gilbert damping parameter of the analyzer should be small.

6. The role of the polarizer (FM1 in Figure 1.2) is to ensure highest possible spin polarization of the incoming current. A large difference in the electronic density of state around the Fermi energy between spin up and spin down states is desirable (see Eq. (1.17)).

With the above conceptual points in mind, an attempt to list few properties of interest for the technology can be undertaken:

- The CBS of the insulator is a descriptor for the damping coefficient of the electrons wave functions crossing the barrier (see Section 1.2). It regulates the symmetry filtering and also the overall impedance of the device. Also the symmetry of the complex bands is important.

- The band structure of the ferromagnets around the Fermi energy (with the corresponding symmetries) is the other necessary piece needed to understand the spin symmetry filtering and it also influences the spacial dependence of the torque (see Eq. (1.11)).
- The spin-dependent electronic density of state of the polarizer is important to evaluate the efficiency of the junction to spin polarize the electric current.
- The reduction of the Gilbert damping parameter, α, is an appealing line of studies, especially considering that Co and Fe show relatively big values of α compared, for example, to Heusler alloys.
- The interfacial anisotropy of the analyzer is a key factor as it regulates both the switching current and the thermal stability.

The aim of this PhD work is to investigate the existence of a completely new combination of materials with promising features for improved STT-MRAMs. The study aims to consider as many materials as possible, therefore it should be based on properties that, although valuable for the technology, can be inexpensive calculated.

The interfacial anisotropy clearly depends on the electronic details of the interface and a systematic study for a large set of interfaces would require an enormous amount of time. The investigation must be limited to bulk properties only.

The Gilbert damping parameter is not easily accessible from *ab-initio* calculations. Its microscopic origin comes from many different contributions and a simple formulation for its estimate is still lacking. Few more details on the topic are presented in Chapter 4, however the Gilbert damping is not extensively studied in this PhD work.

Regarding the third property of our list, namely the spin-dependent electronic density of state, it is an information already available on the web. Projects of high-throughput analysis of materials property, like AFLOW (`http://aflowlib.org`) or Materials Project (`https://materialsproject.org`), provide on-line interfaces listing DFT calculated properties of a large number of compounds. Among these properties, the electronic density of state is always present.

The focus of our study falls then on the first two points of the list above. Band structure calculations of a series of high Curie temperature ferromagnets are presented in Chapter 4 with the relative symmetry analysis for the bands at the Fermi level. The CBS is the focus of Chapter 5 that starts with some theoretical considerations before presenting an attempt of systematic analysis of CBS on a set of binary compounds. To have these information would be enough to identify at least materials combinations with the desirable electronic features for an efficient TMR effect, giving, in addition, information about the overall resistance of the device.

Before presenting the results, the methodology used in the work and some preliminary consideration about the high-throughput approach to the problem are discussed in the next two chapters.

Chapter 2

Methods

The foundations of the computational methods used in this PhD project are rooted in the Density Functional Theory (DFT), one of the most popular quantum mechanical approaches to the study of matter. This chapter is almost entirely devoted to DFT, with the purpose to give a comprehensive review of the fundamental theorems at its core (Section 2.1), to formalize its extension to the study of magnetic systems (Sec. 2.2) and of extended systems (Section 2.4) and to explain the advantages and technical details of the introduction of pseudopotentials-PPs (Section 2.3). Basic concepts of group theory applied to crystal structures are the focus of Section 2.6. This topic is not necessary related to DFT, however, its only purpose in this PhD work is to help the identification of the symmetries of the DFT Kohn-Sham eigenvectors. Limitations of DFT are also discussed and the Atomic Self Interaction Correction (ASIC) approach is reviewed in Sec. 2.7. Moreover, a detailed look at the DFT implementation contained in the code SIESTA is presented in this chapter (Section 2.5). Even though this is not the only DFT code used in this work, it is definitely the most relevant as all the ASIC theory and the Complex Band Structure (CBS) theory are built around the SIESTA implementation. Because of our original contribution to the topic, we felt to report the CBS theory in a chapter containing results, namely Chapter 5. All the other theoretical foundations of this work are outlined in the present chapter. We do not include any proof of the theory presented, however the relevant literature is referenced along the way. Some original derivations are in Appendix B. A large part of the DFT calculations performed in the thesis, are carried out in an high-throughput fashion. We conclude the chapter with a general overview of the high-throughput approach to science and how this translates in the field of computational science. A discussion about the advantages, challenges and tools related to the task of performing high-throughput DFT calculations is included.

2.1 Density Functional Theory Fundamentals

The success of DFT is mainly due to the fact that it allows to calculate the ground-state properties of a portion of condensed matter in an efficient way, mapping the complex many-body problem into a set of single particle equations. The understanding of this extremely

powerful approach to quantum mechanics is the subject of the first section of this chapter.

First of all, only the electronic problem is taken into consideration because the adiabatic approximation (Born-Oppenheimer) is considered valid. The electrons mass is way smaller than the nuclei mass and for this reason the electrons follow instantaneously the motion of the nuclei. As a consequence the nuclei positions are parameters in the study of the electron problem.

For a system of N electrons, the following Schrödinger equation describes the physics:

$$(\hat{T} + \hat{V}_{ext} + \hat{W})\Psi(\mathbf{r_1}, \mathbf{r_2}, ..., \mathbf{r_N}) = E\Psi(\mathbf{r_1}, \mathbf{r_2}, ..., \mathbf{r_N}), \tag{2.1}$$

where Ψ is the many-body wave function (combination of slater determinants with the correct antisymmetrization imposed by the Pauli principle) and

$$\hat{T} = -\sum_{i}^{N} \frac{\hbar^2 \nabla_i^2}{2m}, \qquad \hat{W} = \sum_{i<j} \frac{e^2}{|\mathbf{r_i} - \mathbf{r_j}|}, \tag{2.2}$$

are respectively the kinetic energy and the electron-electron interaction operators (m mass of the electron and e its charge), while $\hat{V}_{ext}$ is a generic single particle, $\hat{V}_{ext} = \sum_i^N v(\mathbf{r_i})$, external potential, including the interaction with the nuclei. The fundamental quantity in DFT is the so-called single particle electronic density defined as:

$$\rho(\mathbf{r}) = N \int d^3 r_2 \int d^3 r_3 ... \int d^3 r_N \Psi^*(\mathbf{r}, \mathbf{r_2}, ..., \mathbf{r_N}) \Psi(\mathbf{r}, \mathbf{r_2}, ..., \mathbf{r_N}). \tag{2.3}$$

It is important to remark that we are defining the single-particle electronic density in the many-body formalism, a quantity deriving from the many-body wave function, an extremely complicated object that is explicitly derivable only in very simple cases. In the following of the section, we often refer to $\rho(\mathbf{r})$ just with the term density. An equivalent expression for $\rho(\mathbf{r})$ in second quantization is:

$$\rho(\mathbf{r}) = \langle \Psi | \, \Psi^\dagger(\mathbf{r}) \Psi(\mathbf{r}) \, | \Psi \rangle \,, \tag{2.4}$$

being Ψ the field operator. Proof of the identity (2.3)= (2.4) is in Appendix B.1.

Three are the pillars of the entire DFT theory: the Hohemberg-Kohn (HK) theorems and the Kohn-Sham (KS) formulation [48, 49]:

Hohemberg-Kohn theorem 1. The first Hohemberg-Kohn theorem states that there is a one-to-one correspondence between the external potential, $v(\mathbf{r})$, and the ground state density, $\rho_0(\mathbf{r})$. Because the external potential uniquely defines the Hamiltonian, it is possible to conclude that the Hamiltonian is a functional of ρ_0. The theorem derives from two preliminary considerations. The first one is that, given a ground-state density $\rho_0(\mathbf{r})$, it is possible to calculate the corresponding ground state wave function $\Psi_0(\mathbf{r}, \mathbf{r_2}, ..., \mathbf{r_N})$. In other words, Eq. (2.3) can be inverted in the ground state and one can write $\Psi_0[\rho_0]$. The second consideration is that there is a one-to-one correspondence between $\Psi_0(\mathbf{r}, \mathbf{r_2}, ..., \mathbf{r_N})$ and $v(\mathbf{r})$. Very importantly, the theorem implies that the ground state energy of the system

can be expressed as a functional of $\rho_0(\mathbf{r})$. In fact, the energy is the expectation value of the Hamiltonian in the ground state, $\langle\Psi_0|\,H\,|\Psi_0\rangle$. As both H and Ψ_0 are functionals of $\rho_0(\mathbf{r})$:

$$E_{gs} = E_0[\rho_0].$$

Hohemberg-Kohn theorem 2. The second theorem states that the ground state energy is the minimum of the energy functional of $\rho(\mathbf{r})$, when $\rho(\mathbf{r})$ is varying in its definition field Υ. We can define the energy as a functional of $\rho(\mathbf{r})$ also for $\rho(\mathbf{r}) \neq \rho_0(\mathbf{r})$ as

$$E[\rho] = \min_{\Psi\to\rho}\langle\Psi|\hat{T} + \hat{V}_{ext} + \hat{W}|\Psi\rangle = \hat{T}[\rho] + \hat{W}[\rho] + \hat{V}_{ext}[\rho], \tag{2.5}$$

meaning $E[\rho]$ is given by $\Psi(\mathbf{r}, \mathbf{r_2}, ..., \mathbf{r_N})$ that reproduces $\rho(\mathbf{r})$ and minimize the energy. Then we can write

$$E_{gs} = E_0[\rho_0] = \min_{\rho\epsilon\Upsilon} E[\rho]. \tag{2.6}$$

Importantly, it is possible to obtain the ground-state energy by solving the functional derivative of $E[\rho]$ respect to ρ (for a review of functional derivatives we suggest appendix A of reference [50]). When written as a function of ρ, it is common to separate the non-interacting kinetic energy, $\hat{T}_0$ from the rest of the kinetic energy. Moreover it is use to separate the classic Coulomb contribution $\hat{J}$ from the rest of $\hat{W}$. In other words it is common to write:

$$E[\rho] = \hat{T}_0[\rho] + \hat{J}[\rho] + \hat{V}_{ext}[\rho] + \hat{E}_{xc}[\rho], \tag{2.7}$$

where:

$$\hat{E}_{xc}[\rho] = \hat{W}[\rho] - \hat{J}[\rho] + \hat{T}[\rho] - \hat{T}_0[\rho]. \tag{2.8}$$

This separation is important because $\hat{E}_{xc}[\rho]$ summarizes the parts of $E[\rho]$ that are difficult to express in the theory. In fact, while it is possible to explicitly express $\hat{J}$ as a function of ρ,

$$\hat{J}[\rho] = \frac{e^2}{2}\int d^3r \int d^3r' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \tag{2.9}$$

the same can not be said for the remaining part of $\hat{W}[\rho]$. Moreover in the next paragraph we will see how to write the non-interacting part of the kinetic energy, but an explicit expression for the total $\hat{T}[\rho]$ is laking. The functional $\hat{E}_{xc}[\rho]$ is called exchange-interaction (xc) term as it includes the exchange part of $\hat{W}[\rho]$, meaning the pure quantum-mechanical effect occurring anytime two identical particles interact, but also the so called correlation. The latter describes all the other many-body interaction effects, including $\hat{T}[\rho] - \hat{T}_0[\rho]$. Another reason to perform the separation in Eq. (2.7) is the fact that $\hat{J}$ is spin independent. $\hat{E}_{xc}$ is carrying the entire spin dependence, as clarified in Sec 2.2.

Kohn-Sham formulation. The last pillar of DFT allows us to obtain ρ_0 by solving a set of N auxiliary single particle equations instead of solving the many-body problem. The N Kohn-Sham (KS) equations (one for each electron i in the system) are

$$H_{KS}(\mathbf{r})\varphi_i(\mathbf{r}) = \left[-\frac{\hbar^2\nabla_i^2}{2m} + v_{eff}(\mathbf{r})\right]\varphi_i(\mathbf{r}) = \varepsilon_i\varphi_i(\mathbf{r}). \tag{2.10}$$

Once Eq. (2.10) with the correct potential $v_{eff}(\mathbf{r})$ is solved, it can be proved that

$$\rho_a(\mathbf{r}) = \sum_i^N \varphi_i^*(\mathbf{r})\varphi_i(\mathbf{r}) \equiv \rho_0(\mathbf{r}). \tag{2.11}$$

The ground-state density of the auxiliary non-interacting problem (from single-particle wave functions) is equal to the density of the many-body problem. This means that in the ground state we are able to calculate a very complicated quantity, Eq (2.3), by solving a much easier set of equations (2.10). Hence, thanks to the HK theorems, we can evaluate the total energy of the many-body system using $\rho_a(\mathbf{r})$. The correct single-particle potential v_{eff} is derived by calculating $\delta E[\rho]/\delta\rho$ (functional derivative of the energy) and it can be written as

$$v_{eff}(\mathbf{r}) = v(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}), \tag{2.12}$$

where $v(\mathbf{r})$ is the single particle external potential. The other two terms are ρ-dependent and writes:

$$v_H(\mathbf{r}) = v_H(\rho, \mathbf{r}) = \frac{\delta J[\rho]}{\delta\rho} = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r', \tag{2.13}$$

while

$$v_{xc}(\mathbf{r}) = v_{xc}(\rho, \mathbf{r}) = \frac{\delta\tilde{E}_{xc}[\rho]}{\delta\rho}, \tag{2.14}$$

with

$$\tilde{E}_{xc}[\rho] = \hat{T}[\rho] - \sum_i^N \int \varphi_i^*(\mathbf{r}) \left(-\frac{\hbar^2\nabla_i^2}{2m}\right) \varphi_i(\mathbf{r}) d^3r + \hat{W}[\rho] - \hat{J}[\rho] \tag{2.15}$$

This last energy term is the exchange-correlation energy functional expressed in the KS formulation of DFT. We stress again that it is the summary of what we are not able to express exactly in terms of ρ. The reasons why $\hat{W}$ and $\hat{T}$ are problematic in the theory are essentially the non-local nature of $\hat{T}$ and the impossibility to express entirely the two-particle operator $\hat{W}$ as a function of the single-particle density ρ. More details can be found in reference [50].

In conclusion, the theory so far is exact. It is proven that E is a functional of ρ in the ground state, although there is no explicit expression for $E[\rho]$. $\tilde{E}_{xc}[\rho]$ is the energy component not exactly describable and has to be approximated, leading to an approximated $v_{xc}(\mathbf{r})$ in the KS equations. The most common approximation is the Local Density Approximation (LDA) [49] where

$$\tilde{E}_{xc}[\rho] = \int \rho\varepsilon_{xc}^{hom}[\rho] d^3r \;\; => \;\; v_{xc}(\rho, \mathbf{r}) = \rho\frac{\partial\varepsilon_{xc}^{hom}[\rho]}{\partial\rho} + \varepsilon_{xc}^{hom}[\rho]. \tag{2.16}$$

Here $\varepsilon_{xc}^{hom}[\rho]$ is the exchange-correlation energy per particle of the uniform electron gas at the density ρ. More sophisticated approximations have been also widely used (*e.g.* the Generalized Gradient Approximation - GGA [51]), but the result presented in this thesis are based on the LDA and in particular on the implementation suggested by Perdew and Zunger [52].

It must be noticed that $v_{xc}(\rho, \mathbf{r})$ and $v_H(\rho, \mathbf{r})$ require the knowledge of the quantity ρ that is our unknown variable. The strategy is that to substitute ρ_a, as defined in Eq. (2.11), in Eq. (2.16) and (2.13) and find the solution of Eq. (2.10) self consistently.

It is worth mentioning that at this stage of the theory the eigenvalues of the KS equations, ε_i, are completely artificial objects (in the derivation of Eq (2.10) they assume the role of Lagrange multipliers necessary for imposing particle conservation in the functional minimization), but they do not have any strict physical meaning. However one can express the ground-state total energy as

$$E[\rho] = \sum_i^N \varepsilon_i - J[\rho] - \int \rho(\mathbf{r}) v_{xc}(\rho, \mathbf{r}) d^3r + \tilde{E}_{xc}[\rho]. \tag{2.17}$$

Moreover, even though conventional DFT is strictly limited to describe the electronic density of the ground-state, it is common to extend the KS equations to unoccupied orbitals and a definition of KS band gap is always given as the difference between the highest occupied KS eigenvalue and the lowest unoccupied one

$$E_g^{KS} = \varepsilon_{N+1} - \varepsilon_N \tag{2.18}$$

This definition is not obviously related to the real semiconductor gap (difference between the electron affinity and the ionization energy), in fact E_g^{KS} is on average 50% lower than the experimental band gap. However, much effort has been done in the past to correct DFT in order to obtain E_g^{KS} similar to the experimental one. This is the topic of the Section 2.7.

2.2 Kohn-Sham formalism for magnetic systems

The study of magnetic systems (systems where the potential is not a scalar, but spin dependent) requires to extend the definition of the density to a 2×2 tensor $\overline{\rho(\mathbf{r})}$ with components

$$\rho^{\alpha\beta}(\mathbf{r}) = \langle \Psi | \, \Psi_\beta^\dagger(\mathbf{r}) \Psi_\alpha(\mathbf{r}) \, | \Psi \rangle \,, \tag{2.19}$$

where α and β spin numbers, assuming the two possible values $\uparrow, \downarrow$. From the definition above derives the scalar electronic density,

$$\rho(\mathbf{r}) = \rho^{\uparrow\uparrow}(\mathbf{r}) + \rho^{\downarrow\downarrow}(\mathbf{r}), \tag{2.20}$$

and the magnetization density,

$$\mathbf{m}(\mathbf{r}) = Tr\left[\overline{\boldsymbol{\sigma}}\ \overline{\rho(\mathbf{r})}\right] = \begin{pmatrix} \rho^{\uparrow\downarrow}(\mathbf{r}) + \rho^{\downarrow\uparrow}(\mathbf{r}) \\ i(\rho^{\uparrow\downarrow}(\mathbf{r}) - \rho^{\downarrow\uparrow}(\mathbf{r})) \\ \rho^{\uparrow\uparrow}(\mathbf{r}) - \rho^{\downarrow\downarrow}(\mathbf{r}) \end{pmatrix}, \tag{2.21}$$

where $\overline{\boldsymbol{\sigma}}$ is the Pauli matrices vector. Therefore $\rho^{\alpha\beta}(\mathbf{r})$ can be rewritten as

$$\rho^{\alpha\beta}(\mathbf{r}) = \frac{1}{2}\rho(\mathbf{r})\delta^{\alpha\beta} + \frac{1}{2}\mathbf{m}(\mathbf{r}) \cdot \boldsymbol{\sigma}^{\alpha\beta}, \tag{2.22}$$

where δ is the Kronecker delta.

In this context, every ground-state expectation value is not only a functional of the scalar density, but also of the magnetization density. This brings some complications in proving the one-to-one relation between the external potential and the ground-state densities [53, 54, 55] The HK theorems need to be reformulated. Moreover, in the derivation of the KS equations, it is necessary to impose that both the density and the magnetization density coincide for the non-interacting and interacting system. We do not go into the theoretical details (that can be found in references [53, 56]), but prefer to take a more practical approach, presenting the relevant parts of the KS formalism for magnetic systems.

The density of the auxiliary system is a tensor as well,

$$\rho_a^{\alpha\beta}(\mathbf{r}) = \sum_i^N \varphi_i^{\beta *}(\mathbf{r})\varphi_i^{\alpha}(\mathbf{r}), \tag{2.23}$$

and Eqs (2.20) and (2.21) are still valid to define the magnetization density and scalar density. The KS equations now read:

$$\begin{bmatrix} -\frac{\hbar^2\nabla_i^2}{2m} + v_{eff}^{\uparrow\uparrow}(\mathbf{r}) & v_{eff}^{\uparrow\downarrow}(\mathbf{r}) \\ v_{eff}^{\downarrow\uparrow}(\mathbf{r}) & -\frac{\hbar^2\nabla_i^2}{2m} + v_{eff}^{\downarrow\downarrow}(\mathbf{r}) \end{bmatrix} \begin{bmatrix} \varphi_i^{\uparrow}(\mathbf{r}) \\ \varphi_i^{\downarrow}(\mathbf{r}) \end{bmatrix} = \varepsilon_i \begin{bmatrix} \varphi_i^{\uparrow}(\mathbf{r}) \\ \varphi_i^{\downarrow}(\mathbf{r}), \end{bmatrix} \tag{2.24}$$

where

$$v_{eff}^{\alpha\beta}(\mathbf{r}) = \delta^{\alpha\beta} \int \frac{\rho^{\uparrow\uparrow}(\mathbf{r}') + \rho^{\downarrow\downarrow}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r' + v_{xc}^{\alpha\beta}(\mathbf{r}) + v^{\alpha\beta}(\mathbf{r}). \tag{2.25}$$

The external potential can now include magnetic fields $\mathbf{B}(\mathbf{r})$:

$$v^{\alpha\beta}(\mathbf{r}) = v(\mathbf{r})\delta^{\alpha\beta} + \mu_B \mathbf{B}(\mathbf{r}) \cdot \boldsymbol{\sigma}^{\alpha\beta}, \tag{2.26}$$

with μ_B being Bohr magneton. However, it is not necessary to have an external field in order to have a magnetic ground state. In fact the xc energy defines in the potential a magnetic-field-like term, which is called exchange-correlation field: $\mathbf{B_{xc}}(\mathbf{r})$:

$$v_{xc}^{\alpha\beta}(\mathbf{r}) = \frac{\delta\tilde{E}_{xc}[\overline{\rho}]}{\delta\rho^{\alpha\beta}} = \frac{1}{2}v_{xc}(\mathbf{r})\delta^{\alpha\beta} + \frac{1}{2}\mu_B \mathbf{B_{xc}}(\mathbf{r}) \cdot \boldsymbol{\sigma}^{\alpha\beta} \tag{2.27}$$

The origin of magnetism was a longstanding problem for classical physics, but we now know that it originates from the exchange interaction, the part of the electron-electron coupling due to the Pauli exclusion principle. In DFT exact exchange is not accessible, but its effect is included in $\tilde{E}_{xc}$ that therefore becomes spin dependent. A more explicit expression for $v_{xc}(\mathbf{r})$ and $\mu_B\mathbf{B_{xc}}(\mathbf{r})$ in terms of $\delta\tilde{E}_{xc}[\overline{\rho}]/\delta\rho^{\alpha\beta}$ can be obtained by comparison with (2.20) and (2.21), however, so far in the theory, all these quantities are just definitions. As in standard DFT, $\tilde{E}_{xc}[\overline{\rho}]$ needs to be approximated, but here we have the additional

complication that $\tilde{E}_{xc}[\overline{\rho}]$ is a tensor dependent quantity. As such expressing it in a form derived from the homogeneous electron gas requires a more careful treatment. Two levels of approximation are currently in use and they are explained in the following two subsections.

2.2.1 Collinear spins

The most commonly used approximation is the so-called collinear spins case, where a single spin-quantization axis is assumed to be common to the entire system. With this assumption, the tensor $\overline{\rho}$ can be used in its diagonal form for every $\mathbf{r}$ and we call its diagonal elements $\rho^+(\mathbf{r})$ and $\rho^-(\mathbf{r})$. $\tilde{E}_{xc}$ depends on ρ^+ and ρ^- only or, equivalently, we can say that the exchange field $\mathbf{B_{xc}}(\mathbf{r})$ is pointing in one direction for each $\mathbf{r}$, so that it can be considered as a scalar. In this situation, an explicit expression for $\tilde{E}_{xc}$ in LDA can be derived. Electrons can be considered to form a spin polarized homogeneous electron gas with spin-up and spin-down densities and in a similar way to the non-magnetic case we can write:

$$\tilde{E}_{xc}[\rho^+,\rho^-] = \int \rho\varepsilon_{xc}^{hom}[\rho^+,\rho^-]d^3r, \tag{2.28}$$

where $\rho(\mathbf{r}) = \rho^+(\mathbf{r}) + \rho^-(\mathbf{r})$ is the scalar density. The magnetization density is also a scalar $m(\mathbf{r}) = \rho^+(\mathbf{r}) - \rho^-(\mathbf{r})$. Consequently,

$$\begin{aligned} v_{xc}^{\uparrow\uparrow}(\rho^+\rho^-,\mathbf{r}) &= v_{xc}^{+}(\rho^+\rho^-,\mathbf{r}) = \rho\frac{\partial\varepsilon_{xc}^{hom}[\rho^+,\rho^-]}{\partial\rho^+} + \varepsilon_{xc}^{hom}[\rho^+,\rho^-], \\ v_{xc}^{\downarrow\downarrow}(\rho^+\rho^-,\mathbf{r}) &= v_{xc}^{-}(\rho^+\rho^-,\mathbf{r}) = \rho\frac{\partial\varepsilon_{xc}^{hom}[\rho^+,\rho^-]}{\partial\rho^-} + \varepsilon_{xc}^{hom}[\rho^+,\rho^-]. \end{aligned} \tag{2.29}$$

Equivalent definitions are

$$\begin{aligned} v_{xc}(\mathbf{r}) &= \left(\frac{\delta\tilde{E}_{xc}[\rho^+,\rho^-]}{\delta\rho^+} + \frac{\delta\tilde{E}_{xc}[\rho^+,\rho^-]}{\delta\rho^-}\right) = \rho\left(\frac{\delta\varepsilon_{xc}^{hom}[\rho^+,\rho^-]}{\delta\rho^+} + \frac{\delta\varepsilon_{xc}^{hom}[\rho^+,\rho^-]}{\delta\rho^-}\right) + \varepsilon_{xc}^{hom}[\rho^+,\rho^-], \\ \mu_B\mathbf{B_{xc}}(\mathbf{r}) &= \mu_B B_{xc,z} = \left(\frac{\delta\tilde{E}_{xc}[\rho^+,\rho^-]}{\delta\rho^+} - \frac{\delta\tilde{E}_{xc}[\rho^+,\rho^-]}{\delta\rho^-}\right) = \rho\left(\frac{\delta\varepsilon_{xc}^{hom}[\rho^+,\rho^-]}{\delta\rho^+} - \frac{\delta\varepsilon_{xc}^{hom}[\rho^+,\rho^-]}{\delta\rho^-}\right). \end{aligned} \tag{2.30}$$

It can be noticed that in the collinear case, $v_{eff}^{\uparrow\downarrow}(\mathbf{r})$ and $v_{eff}^{\downarrow\uparrow}(\mathbf{r})$ are equal to zero and Eq. (2.24) consists of two equations algebraically decoupled, although one needs both ρ^+, ρ^- for the calculation of $v_{xc}^{+/-}$. In this context it is clear how to define spin-up and spin-down KS eigenvalues. This concept is the base for the understanding of the spin-dependent band structures and density of states in extended systems. From there the definition of other important magnetic quantities are given. For example the total magnetization of the system as the integral of $m(\mathbf{r})$ on the entire space. It is also common use to define atomic quantities like the atomic moment. However only global quantities have a clear theoretical foundation in DFT, the choice to divide the space in atomic contribution is arbitrary. In the collinear spins approximation there is no definition of spin direction. This limits the description of a magnetic system to ferromagnetic cases and some kind

of antiferromagnetism, but no spin spirals can be studied, although non-collinear ground states exist in magnetic systems.

2.2.2 Non-collinear spins

A non-collinear theory can be built on the idea that locally one can always diagonalize $\overline{\rho}$. Locally it is always possible to define a local quantization axis and hence to use locally the v_{xc} for the collinear LDA as from Eq. (2.29). In practice, for each $\mathbf{r}$, one can define a rotation matrix $\overline{U(\mathbf{r})}$ such that:

$$\overline{U(\mathbf{r})}\ \overline{\rho(\mathbf{r})}\ \overline{U^{\dagger}(\mathbf{r})} = \begin{pmatrix} \rho^{+}(\mathbf{r}) & 0 \\ 0 & \rho^{-}(\mathbf{r}) \end{pmatrix}. \tag{2.31}$$

One can then express $\delta\tilde{E}_{xc}[\overline{\rho}]/\delta\rho^{\alpha\beta}$ as a function of $\rho^{+}(\mathbf{r})$, $\rho^{-}(\mathbf{r})$ and of both the azimuth and elevation angles mapping the locale frame onto a fixed global one. We leave the derivation to reference [56], but present here the most compact way to define the xc potential:

$$v_{xc}^{\alpha\beta}(\mathbf{r}) = \frac{\delta\tilde{E}_{xc}[\overline{\rho}]}{\delta\rho^{\alpha\beta}} = \frac{1}{2}v_{xc}(\mathbf{r})\delta^{\alpha\beta} + \frac{1}{2}\mu_B B_{xc,z}(\mathbf{r})\cdot\sigma_z^{\alpha\beta}(\mathbf{r}) \tag{2.32}$$

where $v_{xc}(\mathbf{r})$ and $\mu_B B_{xc,z}(\mathbf{r})$ are exactly those of Eq. (2.30) and

$$\sigma_z(\mathbf{r}) = \begin{pmatrix} \cos\theta(\mathbf{r}) & \exp(-i\phi(\mathbf{r}))\sin\theta(\mathbf{r}) \\ \exp(i\phi(\mathbf{r}))\sin\theta(\mathbf{r}) & -\cos\theta(\mathbf{r}) \end{pmatrix}. \tag{2.33}$$

We point out once again that θ and ϕ are dependent on $\mathbf{r}$ because we are assuming a different rotation for each point in space. Practically, one usually assumes an atomic sphere approximation, meaning that each point within a sphere around an atom shares the same local direction of magnetization. This means that in real applications θ and ϕ are the same for each $\mathbf{r}$ within a sphere around each atom in the system. More details are in reference [56], where also the important task of choosing the correct U (and consequently θ and ϕ) for each sphere is discussed.

The theory brings naturally to the definition of atomic spin vector and allows the description of spin spiral states and complicated type of antiferromagnetism. In general cases, there is no coupling between the spin degree of freedom and the crystal lattice, so the global reference frame is arbitrary and only the relative orientation of the spin vectors matters. The situation is different, however, if the spin-orbit coupling is included in the theory. Section 2.3.2 explains the introduction of the spin-orbit coupling in DFT. In this section it is sufficient to mention that it brings a non collinear contribution to the external potential v. In fact, the spin-orbit interaction couples the spin with the lattice and allows, for example, calculations of the magnetocrystalline anisotropy.

2.3 Pseudopotential Theory

The core electrons have low energy and their wave functions, $\varphi_i^C(\mathbf{r})$, are highly localized at the nucleus. In order to reduce the computational demand of solving the KS equations, in DFT codes it is common to ignore the dynamics of the core electrons and to replace their effect by an effective potential, v_p, called the Pseudo-Potential (PP) [57, 58]. The idea is to construct a v_p able to solve an equation for the valence electrons (V) only

$$\left[-\frac{\hbar^2\nabla_i^2}{2m} + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}) + v_p(\mathbf{r})\right]\varphi_i^V(\mathbf{r}) = \varepsilon_i^V\varphi_i^V(\mathbf{r}), \tag{2.34}$$

with

$$\rho_a^V(\mathbf{r}) = \sum_i^{N_V}\varphi_i^{V*}(\mathbf{r})\varphi_i^V(\mathbf{r}). \tag{2.35}$$

At this point, $\rho_a^V(\mathbf{r})$ and not the full electron density goes inside $v_{xc}(\mathbf{r})$ and $v_H(\mathbf{r})$ to define the self consistent cycle. The potential v_p must reproduce the following properties:

- $\varepsilon_i^V = \varepsilon_i$ for each valence electron. ε_i are solution of the full all-electron KS equations.
- $\varphi_i^V(\mathbf{r}) = \varphi_i(\mathbf{r})$ for each valence state beyond a chosen cut off radius R_{ci}. $\varphi_i(\mathbf{r})$ eigenfunctions of the full all-electron KS equations.
- The integral from 0 to $R > R_{ci}$ of the all-electron and pseudo-atomic charge density has to be the same for each valence electron:

  $$\int_0^R drr^2|\varphi_i^V(\mathbf{r})|^2 = \int_0^R drr^2|\varphi_i(\mathbf{r})|^2. \tag{2.36}$$

  This means that the total charge in the core region add to the correct value, namely the norm is conserved.
- Other requirements to ensure transferability (the same PP should work in different chemical environments) are added as constraints in the construction of v_p. In particular it is common to impose that the logarithmic derivatives of the all-electron and PP wave functions coincide at R_{ci}:

  $$r\left[\frac{\partial}{\partial r}\ln(\varphi_i(\mathbf{r}))\right]_{R_{ci}} = r\left[\frac{\partial}{\partial r}\ln(\varphi_i^V(\mathbf{r}))\right]_{R_{ci}}. \tag{2.37}$$

A PP with the listed features above is commonly referred to as "Norm-Conserving PP".

Many different methods for the generation of PPs has been proposed [59, 60, 61]. Their common feature is that the construction of the PP starts from an all-electron calculation for the isolated atom. In this way the external potential is central.The atomic Hamiltonian commutes with one component of the angular momentum operator, L_z, and its magnitude squared, L^2. The wave function can then be written as

$$\varphi_i(\mathbf{r}) = Y_{lm}(\theta,\phi)R_{nl}(r), \tag{2.38}$$

with $Y_{lm}(\theta,\phi)$ being spherical harmonics. The wave function is uniquely defined by the three quantum numbers n, l and m. The wave-function radial component $R_{nl}(r)$ has to be found by solving the radial Schrödinger equation

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2}+\frac{\hbar^2 l(l+1)}{2mr^2}-\frac{Z}{r}+v_H^{IA}(r)+v_{xc}^{IA}(r)\right)rR_{nl}(r)=\varepsilon_{nl}rR_{nl}(r), \tag{2.39}$$

where $v_H^{IA}(r)$ and $v_{xc}^{IA}(r)$ are from Eq. (2.13) and Eq. (2.14) and are calculated for the isolated atom (IA). Here Z is the atomic number. Like any other KS equation, Eq. (2.39) has to be solved self consistently making use of Eq. (2.11) through Eq. (2.38). $R_{nl}(r)$ and ε_{nl} constitute the all-electron reference for the generation of the PP.

The next step is to find a v_l^P and $R_{nl}^V(r)$ that solve:

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2}+\frac{\hbar^2 l(l+1)}{2mr^2}+v_l^P(r)\right)rR_{nl}^V(r)=\varepsilon_{nl}^V rR_{nl}^V(r) \tag{2.40}$$

and brings to the requirements for $\varphi_i^V(\mathbf{r})$ stated at the beginning of this section. As mentioned before, there are many ways to reach this goal [59, 60, 61]. They are not detailed in this report, but it worth mentioning that the Troullier-Martin scheme [61] is the one employed in the generation of the PPs used in the project. The $v_l^P(r)$ obtained in Eq. (2.40) is not transferable because it derives from a procedure where $v_H(\mathbf{r})$ and $v_{xc}(\mathbf{r})$ of the isolated atom are employed, including the valence electrons contribution. However $v_H(\mathbf{r})$ and $v_{xc}(\mathbf{r})$ depend on the chemical environment. Hence it is necessary to subtract from $v_l^P(r)$ the contribution to $v_H^{IA}(r)+v_{xc}^{IA}(r)$ arising from the valence electrons (V, calculated from the pseudovalence charge density):

$$v_l^{ION}(r)=v_l^P(r)-v_H^{IA,V}(r)-v_{xc}^{IA,V}(r). \tag{2.41}$$

Even though there is no full consensus on the terminology, in the following of this thesis we refer to v_l^{ION} with the name "semilocal form of the PP". As a consequence of the derivation, each angular momentum component sees a different potential. The final ionic PP can be written in terms of an operator:

$$\hat{v}_p=\sum_l v_l^{ION}\hat{P}_l, \tag{2.42}$$

where $\hat{P}_l=|l,m\rangle\langle l,m|$ projects on the lth angular momentum component of the wave function.

An important final note is the fact that it is convenient to separate the PP in a local and non local part:

$$\hat{v}_p=v_{local}^P+\sum_l v_{non-local}^P\hat{P}_l=v_{local}^P+\sum_l v_{non-local}^{KB}\hat{P}_l. \tag{2.43}$$

The separation of a fully non-local part brings to substantial savings in computer time and the formalization of the theory is due to the work of Kleinman and Bylander (KB)

[62]. We leave aside the details, but it is important to say that there is no unique way to separate the local part from the non-local one. This is relevant for the understanding of the results contained in Section 3.1.

2.3.1 Non-linear core correction

Equation (2.41) is based on the assumption that there is a complete decoupling of the core charge in the determination of the exchange correlation potential seen by the valence electrons. It is implicit in the equation that

$$v_{xc}^{IA,V+C}(r) = [v_{xc}^{IA,V+C}(r) - v_{xc}^{IA,V}(r)] + v_{xc}^{IA,V}(r). \tag{2.44}$$

Again we remark that $v_{xc}^{IA,V+C}$ means that the total charge density is used to calculate the xc potential, while in $v_{xc}^{IA,V}$ only the pseudovalence charge density from Eq. (2.35) is used. The part in square brackets of Eq. (2.44) enters the PP, while the other part does not, since it is the one which is assumed to change with the environment. The aim is to have in the PP the contribution of the core electrons alone and this will be true if $v_{xc}^{V+C} = v_{xc}^{V} + v_{xc}^{C}$. However since v_{xc} is a non-linear function of the charge density, the latter equation is not necessary guaranteed, the core and the valence are not totally decoupled. This is particularly problematic in the case of magnetic systems as we might end up with a spin dependent PP. Recalling Eq. (2.41), in a spin polarized situation v_l^P would be spin dependent, but the spin dependence should cancel thanks to $v_{xc}^{IA,V}(r)$ in order to have a correct neutral v_l^{ION}. However if the cancellation is not complete, the PP will be spin dependent.

The solution to this problem proposed by Louie and coworkers is that to write [63]:

$$v_l^{ION} = v_l^P - v_H^{IA,V}(r) - v_{xc}^{IA,V+C}(r). \tag{2.45}$$

The entire xc contribution from the isolated-atom reference calculation is subtracted from v_l^{ION}, but in addition the charge of the core (ρ_{IA}^C) is stored. Then, when the calculation for a generic system is performed, the v_{xc} to use in equation 2.34 is:

$$v_{xc} = v_{xc}(\rho_a^V + \rho_{IA}^C) \tag{2.46}$$

where ρ_a^V is calculated self-consistently, while ρ_{IA}^C is always fixed. We are then improving the accuracy of the PP at the expense to retain information about the core charge. For computational reasons, it is not the entire core charge to be stored, but a partial one. This is because the core charge has significant effects only where the core and the valence densities are similar in magnitude, while it is unimportant close to the nucleus where most of the core charge resides. A partial core charge that is equal to the real one outside a radius r_0 and arbitrary inside is selected. Test shows that r_0 should be chosen so that the core density is between 1 to 2 times larger than the valence charge.

2.3.2 Relativistic Pseudopotentials

For heavy elements relativistic effects might be relevant, especially spin-orbit interaction. The PPs offer the chance to include relativistic effects in a rigorous way thanks to the theory of Bachelet and Schlüter [64]. The starting point is now the Dirac Hamiltonian for a central potential, that commutes with the total angular momentum operator $J = L + S$. A separation of the wave function in a radial and angular contribution is still possible by employing the definition of spherical spinors, eigenstates of J^2 and J_z. We do not report the entire theory. The crucial point is that it is possible to obtain two radial equations, one for the minor, $F(r)$, and one for the major, $G(r)$, component of the Dirac wave function of the system. The solution of these two radial equations is the all-electron reference for the generation of the PPs.

Since for valence electrons the major component is essentially decoupled from the minor one outside the core region, it is possible to show that $G(r)$ for the valence electrons outside the core region approximately obeys a Schrödinger-type equation instead of the full Dirac equation,

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2} + \frac{\hbar^2 k(k+1)}{2mr^2} + v_k^P(r)\right) rG_k^V(r) = \varepsilon_k^V rG_k^V(r). \tag{2.47}$$

This approach is correct up to, but not including terms of order α^2 (α is the fine-structure constant). It has the great advantage of allowing the use of a procedure similar to the non-relativistic case to generate the PP. The major difference with respect to the non-relativistic case is the fact that $v_k^P(r)$ (and consequently $v_k^{ION}(r)$) is now a function of the Dirac quantum number k,

$$k = -2(j-l)(j+1/2),$$

where l is the quantum number associated to L^2 and j the one associated to J^2. In other words, the PP is now labeled by j and l instead of l only as in the standard case. Since for each l , j assumes values $|l-s|$,$|l-s+1|$ ... $|l+s-1|$ $|l+s|$ and $s = 1/2$, for each l one have two PPs that can be labeled $l + 1/2$ and $l - 1/2$ except for $l = 0$ where one single $j = 1/2$ is allowed. Some other minor differences in the PP generation procedure are listed in reference [64]. Finally we discuss how to use these PPs (now expressed in the l, j dependent semilocal form) in a real calculation. In fact at the start of the derivation we made the separation of the wave function to a radial part and an angular part by employing the spherical spinors that are eigenvalues of J^2 and J_z. In other words, the derived form of the PPs acts on $|j, m_j\rangle$ states

$$\hat{v}_p = \sum_j v_{l,j}^{ION} |j, m_j\rangle \langle j, m_j| , \tag{2.48}$$

where the sum runs over all the possible js, meaning $j = 1/2$ deriving from $l = 0$, $j = 1/2$, $j = 3/2$ deriving from $l = 1$ and so on. However , the $|j, m_j\rangle$ states can be expressed in terms of of a tensor product of the regular angular momentum states $|l, m\rangle$ and states of

S. It can be proven that

$$\hat{v}_p = \hat{v}_{sc} + \hat{v}_{SO} = \sum_{l,m} \left[v_l^{sc} \overline{I} + v_l^{SO} \mathbf{L} \cdot \overline{\boldsymbol{S}} \right] |l,m\rangle \langle l,m| , \tag{2.49}$$

where for each $l \neq 0$

$$v_l^{sc} = \frac{1}{2l+1} \left[(l+1) v_{l,l+1/2}^{ION} + l v_{l,l-1/2}^{ION} \right] \tag{2.50}$$

and

$$v_l^{SO} = \frac{1}{2l+1} \left[v_{l,l+1/2}^{ION} - v_{l,l-1/2}^{ION} \right] , \tag{2.51}$$

$\overline{I}$ is the 2×2 identity matrix and $\overline{\boldsymbol{S}} = \hbar \overline{\boldsymbol{\sigma}}/2$.

The spin-orbit term emerges spontaneously from the theory as the non-scalar part of the PP when we pass from states of the total angular momentum J^2 to treat states of L^2 and S^2 separately. It can be noticed that a non-collinear treatment of the spin states is necessary anytime spin-orbit interaction is included.

Practical implementations in DFT codes might require additional modifications to the theory (for example on-site approximation in the SIESTA code, [65]), but the approach described here is common to any code employing PPs.

It is in use in the community also a third type of PPs, with intermediate characteristics between the full-relativistic PPs described in this subsection and the standard non-relativistic ones. These are called scalar-relativistic PPs. In practice they consists in the use of only one of the two j associated to each $l \neq 0$ of the full-relativistic treatment. This is equivalent to solve an effective Hamiltonian where some terms of the relativistic kinetic energy are considered, but not the spin-orbit interaction.

2.3.3 Practical aspects

We conclude this section with a list of choices one needs to make in the generation of a PP. These are of practical importance to understand Section 3.1.

- The exchange-correlation functional used for the generation of the PP needs to be coherent with the one used in the actual DFT calculation.

- The level of treatment of relativistic effects is not important for light elements. A full-relativistic treatment is necessary for heavy elements.

- The valence-core partition is most of the times obvious. However there are exceptions. For example K, Rb, Cs have a large polarizable core and PPs with just one electron in the valence might not work. Also, in some compounds containing Ga and In the contribution of d states to the bonding may be not negligible. We call semicore the shells that stays in the valence in the generation of the PP even if they are completely filled and have a n different with respect to that of the last occupied orbital. Of course the use of semicore states brings an additional computational cost to the DFT calculations because more electrons are considered to be part of the valence.

- The non-linear core correction is never harmful, so it can always be included. The core radius r_0 can be often chosen manually by the user and should be chosen so that the core density is between 1 to 2 times larger than the valence charge.
- The PP generation procedures available are several [59, 60, 61]. As already mentioned, in this project the Troullier-Martin scheme is employed [61].
- For each angular momentum channel, l, a radius R_c needs to be chosen. R_c marks the position where the PP and the all-electron wave functions match. Beyond R_c the PP should behave exactly as the real all electron potential. The choice of the cutoff radius is the discriminating factor between accuracy and efficiency. A really short cutoff radius ensures that the pseudo wave function is more similar with respect to the all electron one, but this will make the calculation slower. For bound states, a good compromise between accuracy and efficiency is usually given by choosing, for each l channel, an R_c in the vicinity of the outermost maximum of the corresponding all-electron wave function. This general statement needs however to be checked case by case.

Checking for the transferability of the PP is an essential part of the generation process. This can be done in the first place during the generation procedure, comparing the all-electron and PP logarithmic derivatives of the wave function as a function of the energy for a radius beyond R_c. However it is essential to perform at least one test for an extended system (for instance by comparing the equation of state for a crystal element). More details are presented in Section 3.1

Finally, it is necessary to mention that Norm-Conserving PPs are not the only one type of PPs in use in the electronic structure community. Different types are the Ultrasoft [66] and the PAW [67].

2.4 DFT for solids: the Bloch theorem

The KS Hamiltonian for an extended system spans over the entire space and describes a theoretically infinite number of ions and electrons. Since in DFT the external potential uniquely identifies the Hamiltonian and for a crystal it is periodic, the entire KS potential and the electron density hold as well the periodicity of the crystal. Thanks to the periodicity of the potential, it is possible to replace the difficult task of solving an infinite large Hamiltonian, with the task of solving a reasonably sized Hamiltonian, for an infinite number of what we call k-points. This is possible because of the Bloch theorem.

2.4.1 Bloch's theory

We will now summarize the main points introduced by the Bloch's theory. They apply not only to DFT, but they are the foundation of any electronic structure theory. However, we will keep the KS Hamiltonian as practical example. Even if the focus is on single-particle problems, the theory applies with no difference to many-body Hamiltonians.

The discussion requires first of all the definition of the reciprocal space. A crystal is a repetition of equivalent units and it is always possible to define the primitive lattice

vectors, *i.e.* three vectors, $\mathbf{a_i}$ such that any repetitive unit of the crystal is identified as

$$\mathbf{R} = n_1\mathbf{a_1} + n_2\mathbf{a_2} + n_3\mathbf{a_3}, \tag{2.52}$$

where n_i are integers. The associated three reciprocal primitive vectors are defined as

$$\mathbf{b}_1 = 2\pi\frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)} \quad \mathbf{b}_2 = 2\pi\frac{\mathbf{a}_3 \times \mathbf{a}_1}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)} \quad \mathbf{b}_3 = 2\pi\frac{\mathbf{a}_1 \times \mathbf{a}_2}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)} \tag{2.53}$$

and each point of the reciprocal lattice can be written as

$$\mathbf{G} = m_1\mathbf{b_1} + m_2\mathbf{b_2} + m_3\mathbf{b_3} \tag{2.54}$$

with m_i being integers. Very important statement (proof in chapter 2 of reference [68]) is that any periodic function can be expanded in Fourier series containing reciprocal vectors $\mathbf{G}$ only. This applies to any periodic potential (generically called $V(\mathbf{r})$ in the following) including, for instance, the $v_{eff}(\mathbf{r})$ of the KS theory,

$$V(\mathbf{r}) = \sum_{\mathbf{G}} V_{\mathbf{G}} e^{i\mathbf{G}\mathbf{r}}. \tag{2.55}$$

The single-particle Schrödinger equation (could be a KS equation) then writes:

$$\hat{H}\varphi(\mathbf{r}) = \left(-\frac{\hbar^2\nabla^2}{2m} + \sum_{\mathbf{G}} V_{\mathbf{G}} e^{i\mathbf{G}\mathbf{r}}\right)\varphi(\mathbf{r}) = \varepsilon\varphi(\mathbf{r}). \tag{2.56}$$

The next step consists to guess a form of the wave function $\varphi(\mathbf{r})$:

$$\varphi(\mathbf{r}) = \sum_{\mathbf{k}} c(\mathbf{k}) e^{i\mathbf{k}\mathbf{r}} \tag{2.57}$$

The choice is justified by the fact that $e^{i\mathbf{k}\mathbf{r}}$ is an eigenfunction of the translator operator $\hat{R}$ (the operator that transforms any function $f(\mathbf{r})$ to $f(\mathbf{r}+\mathbf{R})$) and it can be proven that $\hat{R}$ commutes with $\hat{H}$. Here $\mathbf{k}$ is called wave vector and its possible values are defined by the Born–von Karman boundary conditions applied to the wave function. Periodic boundary conditions are commonly assumed in the study of extended systems because they offer a convenient way to simulate infinite systems. In the case of a crystal, the Born–von Karman boundary conditions are:

$$\varphi(\mathbf{r} + N\mathbf{R}) = \varphi(\mathbf{r}), \tag{2.58}$$

where N (integer) goes eventually to infinity to describe the entire system. With our choice of the wave function we obtain

$$\varphi(\mathbf{r} + N\mathbf{R}) = \sum_{\mathbf{k}} c(\mathbf{k}) e^{i\mathbf{k}\mathbf{r}} e^{iN\mathbf{k}\mathbf{R}} = \sum_{\mathbf{k}} c(\mathbf{k}) e^{i\mathbf{k}\mathbf{r}} = \varphi(\mathbf{r}), \tag{2.59}$$

meaning that $e^{iN\mathbf{k}\mathbf{R}}$ needs to be equal to one, or in other words, $\mathbf{k} = 2\pi\mathbf{p}/N\mathbf{R}$, where the components of $\mathbf{p}$ assume integer values (negative, positive, zero). So $\mathbf{k}$ assumes the

discrete values allowed by the boundary conditions and it become a continuum for $N \to \infty$. Substituting Eq. (2.57) into Eq. (2.56), after some mathematics [68], a very interesting form of the Schrödinger equation can be reached:

$$\sum_{\mathbf{k}} \left(\frac{\hbar^2 |\mathbf{k}|^2}{2m} - \varepsilon \right) c(\mathbf{k}) e^{i\mathbf{kr}} + \sum_{\mathbf{k}} \sum_{\mathbf{G}} V_{\mathbf{G}} c(\mathbf{k} - \mathbf{G}) e^{i\mathbf{kr}} = 0 \tag{2.60}$$

Each Fourier component has to have the same coefficient so:

$$\left(\frac{\hbar^2 |\mathbf{k}|^2}{2m} - \varepsilon \right) c(\mathbf{k}) + \sum_{\mathbf{G}} V_{\mathbf{G}} c(\mathbf{k} - \mathbf{G}) = 0. \tag{2.61}$$

This equation proves one of the central points of the Bloch's theory. It is in fact clear that each $\mathbf{k}$ vector is related only to vectors that are distant $\mathbf{G}$ from it. This means that one has an independent problem for each $\mathbf{k}$ vector in the primitive cell of the reciprocal space (first Brillouin zone). The problem is now moved to the evaluation of the coefficients $c(\mathbf{k})$ and of course they will depend on the form of the potential $V(\mathbf{r})$. An independent wave function for each $\mathbf{k}$ can also be defined in this form [68]:

$$\varphi_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c(\mathbf{k} - \mathbf{G}) e^{-i\mathbf{Gr}} e^{i\mathbf{kr}} = u_{\mathbf{k}}(\mathbf{r}) e^{i\mathbf{kr}} \tag{2.62}$$

and it is easy to prove that $u_{\mathbf{k}}(\mathbf{r} + \mathbf{R}) = u_{\mathbf{k}}(\mathbf{r})$. This concludes the Bloch theory that we can summarize in this statement.

Bloch theorem. Given a periodic potential, it is possible to separate the Schrödinger equation in a set of independent equations, one for each $\mathbf{k}$ in the first Brillouin zone and defined by the periodic boundary conditions. The wave function of each independent $\mathbf{k}$-dependent problem assumes the form

$$\varphi_{\mathbf{k}} = u_{\mathbf{k}}(\mathbf{r}) e^{i\mathbf{kr}}, \tag{2.63}$$

where $u_{\mathbf{k}}(\mathbf{r})$ is a function with the same periodicity of the potential. An equivalent expression for the Bloch wave function of practical importance is

$$\varphi_{\mathbf{k}}(\mathbf{r} + \mathbf{R}) = e^{i\mathbf{kR}} \varphi_{\mathbf{k}}(\mathbf{r}). \tag{2.64}$$

Since one has an independent Schrödinger equation for each wave vector, the energy is dependent on $\mathbf{k}$ as well. $\varepsilon(\mathbf{k})$ is the dispersion relation of a the system and for real $\mathbf{k}$ it represents the so called "band structure" of a material.

The theory outlined here is taken from Kittel's book [68]. However a different proof of the Bloch theorem can be obtained just from the observation that $\hat{H}$ commutes with $\hat{R}$. This latter proof is for instance contained in Ashcroft-Mermin [69]. While the Kittel's discussion emphasizes the origin of the separation in independent problems, the Ashcroft-Mermin derivation points more attention to the fact that the Bloch theorem allows us to classify the eigenstates of the crystal Hamiltonian by $\mathbf{k}$, seen as the quantum number

associated to the commutation of the Hamiltonian and the translation operator. The Bloch theorem is of vital importance for the understanding of Section 2.6 and Chapter 5.

2.4.2 Plane-wave basis set

In any practical application of DFT, it is necessary to introduce a basis set on which to expand the wave functions. The main reason is to transform the KS differential equations into an algebraic problem involving the diagonalization of a matrix. By looking at the Bloch theory just described, it is understandable how a convenient basis set choice consists of plane waves, a complete basis for the Bloch's functions This is very general as it is independent from the type of crystal. The plane-wave expansion is exactly that introduced in Eq. (2.62) and the KS hamiltonin can as well be recasted in a form containing Fourier terms only. This leads to an algebraic equation for the coefficients $c(\mathbf{k}-\mathbf{G})$.

In theory, an infinite set of $\mathbf{G}$ should be considered for the exact solution. However, the plane-waves with small kinetic energy are the most important, so the basis can be limited to the elements with a kinetic energy lower than a cutoff E_{cut},

$$\frac{\hbar^2}{2m}|\mathbf{k}+\mathbf{G}|^2 < E_{\text{cut}}.$$

This leads to another advantage in using a plane-wave basis set: the accuracy can be systematically improved by increasing E_{cut}. However, in order to described localized functions around the nucleus, a huge number of plane-waves needs to be used and the diagonalization of big matrices is computationally expensive. In general the large number of of plane wave in use, makes almost impossible to store the Hamiltonian and the wave functions for system with more that few atoms. This justifies the use of local orbitals, described in the next section.

2.4.3 Local, non-orthogonal basis set for extended systems

The choice of localized orbitals is motivated by their computational efficiency, because a small number of them are usually sufficient to describe the KS wave functions. Hamiltonian and wave functions can be easily stored and used for post processing in a sort of tight-binding [70] spirit. Few types of local orbitals are in use, the most relevant are the Gaussian type and atomic orbitals. Atomic orbitals give the further advantage to provide a physical intuition of the atomic properties of the system. Local orbitals are clearly the choice for the study of isolated molecules, but many codes employ them for extended systems. The connection to the Bloch theory is not as straightforward as in plane-wave and it is the topic of the next few paragraphs. To be completely general, the basis set is allowed to be non-orthogonal.

The starting point is a single-particle Hamiltonian, again the theory is general, but we her focus in particular on the KS equation

$$H_{KS}(\mathbf{r})\varphi(\mathbf{r}) = \varepsilon\varphi(\mathbf{r}). \tag{2.65}$$

$H_{KS}(\mathbf{r})$ is periodic, but for the moment the periodicity is not explicitly considered. The $H_{KS}(\mathbf{r})$ and $\varphi(\mathbf{r})$ are defined over the entire space. If we call $\phi_\nu^{\mathbf{R}}(\mathbf{r})$ a generic element of the localized basis, the wave function expansion reads

$$\varphi(\mathbf{r}) = \sum_{\mathbf{R}}^{\infty} \sum_{\nu}^{orb} c_\nu^{\mathbf{R}} \phi_\nu^{\mathbf{R}}(\mathbf{r}), \tag{2.66}$$

where $\mathbf{R}$ is a generic vector of the crystal lattice. We are just separating the total sum in the sum over the basis elements inside the elementary cell (orb in total, labeled by ν) and the sum over the cells constituting the system. The functional form of $\phi_\nu^{\mathbf{R}}(\mathbf{r})$ is the same for each $\mathbf{R}$, however the center of the function is different. For this reason we keep the index $\mathbf{R}$ to distinguish between orbitals in different unit cells. We assume for simplicity one atom in unit cell, so that all the νs are centered at the same place for a given $\mathbf{R}$. In principle the sum runs over all the unit cells of the infinite crystal. The Schrödinger equation is

$$H_{KS}(\mathbf{r}) \sum_{\mathbf{R}}^{\infty} \sum_{\nu}^{orb} c_\nu^{\mathbf{R}} \phi_\nu^{\mathbf{R}}(\mathbf{r}) = \varepsilon \sum_{\mathbf{R}}^{\infty} \sum_{\nu}^{orb} c_\nu^{\mathbf{R}} \phi_\nu^{\mathbf{R}}(\mathbf{r}) \tag{2.67}$$

In bra-ket notation it writes

$$\langle \mathbf{r}|H_{KS}|\mathbf{r}\rangle \sum_{\mathbf{R}}^{\infty} \sum_{\nu}^{orb} c_\nu^{\mathbf{R}} \langle \mathbf{r}|\nu\mathbf{R}\rangle = \varepsilon \sum_{\mathbf{R}}^{\infty} \sum_{\nu}^{orb} c_\nu^{\mathbf{R}} \langle \mathbf{r}|\nu\mathbf{R}\rangle. \tag{2.68}$$

By multiplying both sides by $\phi_\mu^{\mathbf{0}*}(\mathbf{r}) = \langle \mu\mathbf{0}|\mathbf{r}\rangle$ (an element of the basis defined in the elementary cell) and integrating over the space we obtain:

$$\sum_{\mathbf{R}}^{\infty} \sum_{\nu}^{orb} c_\nu^{\mathbf{R}} \int \langle \mu\mathbf{0}|\mathbf{r}\rangle \langle \mathbf{r}|H_{KS}|\mathbf{r}\rangle \langle \mathbf{r}|\nu\mathbf{R}\rangle d^3r = \varepsilon \sum_{\mathbf{R}}^{\infty} \sum_{\nu}^{orb} \int c_\nu^{\mathbf{R}} \langle \mu\mathbf{0}|\mathbf{r}\rangle \langle \mathbf{r}|\nu\mathbf{R}\rangle d^3r. \tag{2.69}$$

The quantity $\int |\mathbf{r}\rangle\langle\mathbf{r}|d^3r$ defines a completeness relation for the real space representation, and, to be totally formal, the local operator $\langle \mathbf{r}|H_{KS}|\mathbf{r}\rangle$ should be written $\langle \mathbf{r}'|H_{KS}|\mathbf{r}\rangle\delta(\mathbf{r} - \mathbf{r}')$. Therefore on the left side one would have another integral in d^3r' and two completeness relations. In conclusion we can write:

$$\sum_{\mathbf{R}}^{\infty} \sum_{\nu}^{orb} c_\nu^{\mathbf{R}} \langle \mu\mathbf{0}|H_{KS}|\nu\mathbf{R}\rangle = \varepsilon \sum_{\mathbf{R}}^{\infty} \sum_{\nu}^{orb} c_\nu^{\mathbf{R}} \langle \mu\mathbf{0}|\nu\mathbf{R}\rangle \tag{2.70}$$

We call:

$$h_{\mu\mathbf{0},\nu\mathbf{R}} = \langle \mu\mathbf{0}|H_{KS}|\nu\mathbf{R}\rangle = \int \phi_\mu^{\mathbf{0}*}(\mathbf{r}) H_{KS}(\mathbf{r}) \phi_\nu^{\mathbf{R}}(\mathbf{r}) d^3r, \tag{2.71}$$

$$s_{\mu\mathbf{0},\nu\mathbf{R}} = \langle \mu\mathbf{0}|\nu\mathbf{R}\rangle = \int \phi_\mu^{\mathbf{0}*}(\mathbf{r}) \phi_\nu^{\mathbf{R}}(\mathbf{r}) d^3r. \tag{2.72}$$

Here $s_{\mu\mathbf{0},\nu\mathbf{R}}$ are the overlap integrals and are normalized, meaning $s_{\mu\mathbf{0},\mu\mathbf{0}} = 1$ for each $\mu = 1, .., orb$. The $h_{\mu\mathbf{0},\nu\mathbf{R}}$ are the Hamiltonian matrix elements. Since μ runs in the elementary cell, both these quantities are going to be zero for $\mathbf{R}$ outside a certain cutoff.

This is a consequence of the basis orbitals having strictly finite range. Such feature limits the sum in the Schrödinger equation to a finite number of cells (*cells*),

$$\sum_{\mathbf{R}}^{cells}\sum_{\nu}^{orb} c_{\nu}^{\mathbf{R}} h_{\mu\mathbf{0},\nu\mathbf{R}} = \varepsilon \sum_{\mathbf{R}}^{cells}\sum_{\nu}^{orb} c_{\nu}^{\mathbf{R}} s_{\mu\mathbf{0},\nu\mathbf{R}} \tag{2.73}$$

The Bloch theorem is introduced proving this identity:

$$c_{\nu}^{\mathbf{R}} = e^{i\mathbf{k}\mathbf{R}} c_{\nu}^{\mathbf{0}}. \tag{2.74}$$

The formal proof of this relation is in the Appendix B.2, it derives directly from Eq. (2.64). For each $\mathbf{k}$ allowed by the boundary conditions we have:

$$\sum_{\mathbf{R}}^{cells}\sum_{\nu}^{orb} c_{\nu}^{\mathbf{0}} e^{i\mathbf{k}\mathbf{R}} h_{\mu\mathbf{0},\nu\mathbf{R}} = \varepsilon \sum_{\mathbf{R}}^{cells}\sum_{\nu}^{orb} c_{\nu}^{\mathbf{0}} e^{i\mathbf{k}\mathbf{R}} s_{\mu\mathbf{0},\nu\mathbf{R}}. \tag{2.75}$$

Introducing the following definitions:

$$H_{\mu,\nu}(\mathbf{k}) = \sum_{\mathbf{R}}^{cells} e^{i\mathbf{k}\mathbf{R}} h_{\mu\mathbf{0},\nu\mathbf{R}}, \tag{2.76}$$

$$S_{\mu,\nu}(\mathbf{k}) = \sum_{\mathbf{R}}^{cells} e^{i\mathbf{k}\mathbf{R}} s_{\mu\mathbf{0},\nu\mathbf{R}}, \tag{2.77}$$

Eq. (2.75) becomes

$$\sum_{\nu}^{orb} H_{\mu,\nu}(\mathbf{k}) c_{\nu}^{\mathbf{0}} = \varepsilon \sum_{\nu}^{orb} S_{\mu,\nu}(\mathbf{k}) c_{\nu}^{\mathbf{0}}. \tag{2.78}$$

The above equation can be recasted in a matrix form defining $\bar{H}(\mathbf{k})$ as the $orb \times orb$ matrix whose elements are the $H_{\mu,\nu}(\mathbf{k})$, $\bar{S}(\mathbf{k})$ as the matrix of the overlaps $S_{\mu,\nu}(\mathbf{k})$ and $\mathbf{c}$ as the vector with elements $c_{\nu}^{\mathbf{0}}$. The matrix form of Eq. 2.78 is then:

$$\bar{H}(\mathbf{k})\mathbf{c} = \varepsilon \bar{S}(\mathbf{k})\mathbf{c}, \tag{2.79}$$

This last equation can be solved to find the $\varepsilon(\mathbf{k})$ and the $c_{\nu}^{\mathbf{0}}(\mathbf{k})$ for each $\mathbf{k}$. The wave function reads:

$$\varphi_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{R}}^{cells}\sum_{\nu}^{orb} c_{\nu}^{\mathbf{0}}(\mathbf{k}) e^{i\mathbf{k}\mathbf{R}} \phi_{\nu}^{\mathbf{R}}(\mathbf{r}) \tag{2.80}$$

This concludes the discussion on the use of a localized basis set for extended systems. As mentioned before, local orbitals basis sets are very efficient and intuitive, however they are not complete and they have the downside of not being single-parameter variational. More on the topic will be discussed in the next section.

2.5 SIESTA

With the term SIESTA (Spanish Initiative for Electronic Simulations with Thousands of Atoms [71]) we refer to a method and a code developed in the framework of DFT and with the following main characteristics: the expansion of the wave function is a linear combination of atomic orbitals and the Hamiltonian matrix elements are efficiently calculated so that the code scales linearly with the number of atoms in the system. PPs are employed and the extension to non-collinear systems with inclusion of spin-orbit coupling has been recently added [65].

The atomic basis orbitals used by SIESTA are products of a radial function and a spherical harmonic. The basic form of the radial part is calculated following the Sankey and Niklewski method [72, 73] and basically consists in the eigenfunctions $\phi_l(r)$ of the atomic PP, $V_l^{ION}(r)$, for an energy $\varepsilon_l + \delta\varepsilon_l$ chosen so that the first node occurs at the desired cutoff radius r_l^c. In this way r_l^c defines the point after which the wave function is considered to be zero. This is important to ensure the finite range of the elements of the basis. In addition, several orbitals with the same angular dependence but different radial part (called multiple-ζ) can be included in the basis. The second-ζ functions, $\phi_l^{2\zeta}(r)$, have the same tail as the first-ζ, $\phi_l^{1\zeta}(r)$, but behaves as a simple polynomial inside a 'split radius' r_l^s, namely it writes

$$\phi_l^{2\zeta}(r) = \begin{cases} r^l(a_l - b_l r^2) & r < r_l^s \\ \phi_l^{1\zeta}(r) & r > r_l^s \end{cases} . \tag{2.81}$$

The split radius is another parameter (together with the r_l^c) that needs to be chosen for the generation of the basis set. The standard choice of the number of l to include in the basis is that of considering the valence electron shells, but sometimes an extra angular degree of freedom needs to be added. Using pseudoatomic orbitals of higher angular momentum is unsatisfactory, because they might be too extended, or even unbound. In some cases the specific l-dependent PP channel may not be available. The solution is then to polarize the last occupied angular momentum $\bar{l}$. This is achieved by solving a pseudoatomic-like equation obtained in the first order perturbation theory with a small electric field as a perturbation. Details are in reference [71]. Importantly, selection rules imply that the resulting perturbed orbital will only have components with angular momentum $\bar{l} + 1$, so that an additional orbital with reasonable radial part and $\bar{l} + 1$ spherical harmonics as angular part is generated. Finally, note that the basis set described here is non-orthogonal, the theory in Section 2.4.3 is fully employed in the SIESTA code.

Once again we want to remark that the use of local orbitals is very efficient, but it is not obvious how to improve the accuracy of a calculation. First of all, the basis in not complete. Secondly, the parameters to chose are several and, even if a variational a approach can be used to compare the accuracy of the different basis, the presence of many parameters makes the entire process of optimization non trivial. In general, longer is r_l^c , better should be the calculation result at the expense of calculation efficiency. However, increasing r_l^c increases the overlap between orbitals and this might lead to some issues for particular applications (Section 5.2).

The second main feature of SIESTA, namely to scale linearly wuth the number of orbitals, it is not as relevant for the purpose of this thesis as the understanding of the basis set. We just mention that the idea is to calculate most of the contributions of $\langle \mu \mathbf{0} | H_{KS} | \nu \mathbf{R} \rangle$ on a radial grid in real space. This, combined with the key feature to have local orbitals as a basis, brings the code to scale linearly with the number of electrons. The details are in [71].

2.6 Group theory and band symmetries

Another relevant topic for the results of this PhD thesis is the understanding of the symmetry of the bands of a material, a concept that has been used for example in Section 1.2 to explain the symmetry filtering effect. An entire course on group theory applied to crystal structures should be reported in order to fully grasp the significance of the symmetries in electronic structure theory. Here we have to limit the discussion to a few targeted concepts, without being very formal. Books on the topics are from Heine [74] and Dresselhaus [75]. We follow the second one.

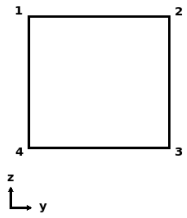
The first concept concerns the space group of a crystal. In Section 2.4 we have discussed how the Hamiltonian is invariant with respect to translations of the lattice vectors $\mathbf{R}$. However, there are also other operations that transform the crystal structure into itself, for example rotations and reflections with respect to certain axis. In the group theory terminology these are called point group symmetry operations (Chapter 3 of reference [75]). All the symmetry operators that commute with the Hamiltonian are indicated, following the notation of Dresselhaus, with the symbol $\{R_\alpha|\tau\}$ and provide quantum numbers for labeling the energy eigenvalues and eigenfunctions. R_α represents the point group operations, while τ indicates all the translations (not only the one of a lattice vector).

One could say that the collection of all the $\{R_\alpha|\tau\}$ defines the space group of a crystal, however the operations are an infinite number. It would be then convenient to factor out the pure translational properties of the space group and select a finite number of elements that fully describe the space group. One can define a subgroup of the space group containing pure translation operations ($\{\varepsilon|\tau\}$). However individual elements of the translations subgroup do not commute, in general, with other element of the space group. It is not possible then to say that the point group only is sufficient to describe a space group. Instead, one can write the space group operations as:

$$\{R_\alpha|\tau\} = \{\varepsilon|R_n\}\{R_\alpha|\tau_\alpha\} \tag{2.82}$$

where $\{\varepsilon|R_n\}$ is a translation with respect to a generic vector of the crystal lattice, and the vector τ_α (associated with each of the point group operators R_α) is either zero or describes a translation that is not a primitive translation. Section 9.1 of Dresselhaus' book [75] discusses how some selected elements of $\{R_\alpha|\tau_\alpha\}$ (all the pure point group operations plus some others) and the three primitive lattice vectors are sufficient to fully describe a space group.

The possible crystallographic lattices that can form three dimensional structure are



representation	basis functions	E	C_4^2	$2C_4$	$2iC_4^2$	$2iC_2'$
Δ_1	$1, x, 2x^2 - y^2 - z^2$	1	1	1	1	1
Δ_2	$y^2 - z^2$	1	1	-1	1	-1
Δ_2'	yz	1	1	-1	-1	1
Δ_1'	$yz(y^2 - z^2)$	1	1	1	-1	-1
Δ_5	$y, z; xy, xz$	2	-2	0	0	0

Figure 2.1: Character table of the point group C_{4v} and schematic representation of a square. Picture adapted from Ref. [75], Chapter 10.

14 in total (the Bravais lattices) and only 230 space groups can be generated by placing different atomic structures in the Bravais lattices. The number is limited by the fact that the requirements imposed by the translational symmetry limit the possible rotation angles and axis [76].

The next concept to clarify is the definition of the symmetry of a particular wave vector $\mathbf{k}$. In fact, only the Γ point, $\mathbf{k} = (0, 0, 0)$, respects the full set of symmetries of the crystal. For other $\mathbf{k}$ points, only a subgroup of the space group symmetry operations brings the $\mathbf{k}$ in itself or in an equivalent $\mathbf{k} + \mathbf{G}$. The set of space group operations which transform $\mathbf{k}$ into itself, or into an equivalent $\mathbf{k}$ is called "group of the wave vector". For a generic $\mathbf{k}$ the identity only takes the wave vector into itself and in this case the wave functions describing electron states only see the translational symmetry of the space group. However, there are higher symmetry points whose group of the wave vector contains point group operations. These are interesting for our work and in particular it is important to understand how to interpret the so called "character table" of the group of the wave vector.

In order to formally define what a "character table" is, one requires too many concepts of group theory (Chapter 1-3 Dresselhaus [75]). Here we just present an operative explanation of why it is useful for us, following a specific example.

Taking a simple cubic lattice (space group 221), a generic $\mathbf{k}$ point along the line Γ−X is often called a Δ point. $\mathbf{k}$ vectors along that line respect the symmetries of a square, rather than those of a cube. These are the symmetry operations of point group C_{4v}. In other words, the "group of the wave vector" for Δ points is the point group C_{4v}, whose character table is in Figure 2.1, together with a schematic representation of a square where the four angles have been labeled with numbers. Leaving aside the "representation" and "basis functions" columns for a moment, the title line lists the 8 symmetry operations of the square:

- E is the identity.
- C_4 is a $\pi/2$ rotation around the four fold axis (along the x direction in Figure 2.1). There are two of such rotations, one clockwise (moving the angles 1234 to 4123) and one anti-clockwise (moving 1234 to 2341). The fact that there are two is an information that is included in the title line of Figure 2.1. In fact there is a pre-factor 2 before C_4.
- C_4^2 is a π rotation around the same axis, leading 1234 to 3412. There is only one of such

rotation.

- iC_2^4 is a π rotation with respect to a two-fold axis passing through the center of two opposite sides. There are two equivalent of such axis, the first one parallel to the y axis (brings 1234 to 4321) and the second parallel to the z axis (brings 1234 to 2143).

- Finally iC_2' is a π rotation respect to a two fold axis connecting two opposite angles, so in the direction of the diagonals of the square. Again we have two, the first brings 1234 to 1432 and the second brings 1234 to 3214.

We explained the 8 symmetry operations in a way that could be as much intuitive as possible. Of course there are equivalent ways to define an operation, using for example reflections and so on. This is why the labeling of each symmetry operation is probably not so intuitive, for a better understanding we point again to chapter 3 of Dresselhaus' book [75].

We now move to explain the relevance for our work of the so call "irreducible representations" of the point group, the first column in the Figure 2.1. For a formal definition Chapter 2 of [75] comes to help. We are interest in the interpretation of the numbers on the table associated to each irreducible representation. By looking at the first line, one can observe that all the numbers are one. This means that the Δ_1 representation brings any spacial function $f(\mathbf{r}) = f(x, y, z)$ in itself when any of the symmetry operations listed above is applied to the function. The Δ_2 representation instead, brings $f(x, y, z)$ in itself except for the operations C_4 and iC_2', that transform f into $-f$. Similar arguments apply to $\Delta_{2'}$ and $\Delta_{1'}$. The last row is a bit different as the number 2 appears as first entry. The first column refers to the identity E, so certainly any function maps on itself when the identity is applied. The 2 in the table means that this irreducible representation has dimension two instead of dimension one like all the others. This means that at least two functions are necessary to describe the representation and the transformation is defined by a 2×2 matrix. The number in the line in Figure 2.1 is thus the trace of the transformation matrix. To better clarify, we now list for each symmetry operation the corresponding representative matrix in the Δ_5 irreducible representation:

E	C_4^2	C_4 1^{st}	C_4 2^{nd}	iC_4^2 1^{st}	iC_4^2 2^{nd}	iC_2' 1^{st}	iC_2' 2^{nd}
$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
2	-2	0	0	0	0	0	0

The choice of the matrices above is arbitrary. In fact a similarity or equivalence transformation would produce an equally good set of matrices. The trace of the matrix (reported in the last line) though is not arbitrary, it is the same for any similar matrix. The trace is called character and corresponds to the last line in Figure 2.1.

Final mention goes to the previously ignored "basis functions". They are an example of spacial function $f(\mathbf{r}) = f(x, y, z)$ that transform accordingly to the respective representation. For instance $f(\mathbf{r}) = yz$ goes in itself for any applied operation except for C_4 and iC_2^4. In fact y goes in z if we apply a 90° rotation of the square in anti-clockwise direction,

but z goes in $-y$. If the rotation is clockwise y goes in $-z$ while z goes in y. Either ways yz goes in $-yz$. A similar discussion can be done for iC_2^4.

The importance of symmetries and group theory for the electronic structure goes way beyond the concepts mentioned here. For instance the degeneracy and splitting of bands could be predicted from a group theory investigation of the crystal. Nevertheless, the discussion outlined here is sufficient for the purpose of understanding how to label the symmetry of the bands of a material. We summarize in a more systematic way the various steps. Firstly, the "group of the wave vector" g_k for the high symmetry direction of interest needs to be found. Section 10.4 of Dresselhaus [75] discuss how to do it, however it is enough to say that luckily the g_k for every $\mathbf{k}$ of every space group is listed in tables or websites [77, 78, 79]. Next step is to figure out the "character table" of g_k. Once that is available, one needs to understand according to which irreducible representation of g_k the wave function of a particular band transforms. This is probably a concept that was not stressed enough in the previous paragraphs: the symmetry operations of g_k (or of any space group) bring the lattice in itself, however the wave functions are not invariant respect to the former operations, they in fact follow one particular representation of g_k. In other words, the wave functions for a certain $\mathbf{k}$ are basis functions for the representations of g_k. Finding out which particular representation, allows us to label the symmetry of the corresponding band.

The concepts outlined here are used in Section 4.2, where we explain the final practical actions one needs to take in order to label the symmetry of a band.

2.7 Band gap in DFT and Atomic Self Interaction Correction

In Section 2.1 we have underlined how the density is the only quantity deriving from the KS equations to hold a strict physical meaning. The KS eigenvalues for the occupied orbitals bear only a semiquantitative resemblance with the true energy spectrum, while the empty bands are not to be trusted at all. The real band gap could be obtained from exact DFT by looking at the total energy. In fact the correct ionization energy, I, and electron affinity, A could be extracted from

$$I = E_0(N-1) - E_0(N), \qquad A = E_0(N) - E_0(N+1), \tag{2.83}$$

where $E_0(M)$ is the ground-state energy of the system with M electrons. The band gap is $E_g = I - A$. Therefore, in theory, with three DFT total energy calculations one could obtain the real many-body gap. Unfortunately, there is no exact DFT and this justifies the efforts to try to estimate the real gap by other means. This section deals with the understanding of the problematic calculation of band gaps in DFT and possible solution to the issue, in particular the correction used for this PhD project, namely the Atomic Self-Interaction Correction (ASIC), is presented in details.

2.7.1 The problem of the gap in approximated DFT

For the description of the band gap in DFT, of great importance is the finding of Sham and Schlüter [80]. They were able to rigorously show that

$$E_g = E_g^{KS} + E_g^{xc}, \tag{2.84}$$

where E_g^{KS} is the difference between the first unoccupied KS eigenvalue and the highest occupied one (see Eq. (2.18)) and E_g^{xc} can be expressed in the following way:

$$E_g^{xc} = \frac{\delta \tilde{E}_{xc}}{\delta \rho_>} - \frac{\delta \tilde{E}_{xc}}{\delta \rho_<}, \tag{2.85}$$

where

$$\frac{\delta \tilde{E}_{xc}}{\delta \rho_>} = \lim_{M \to N+1} \frac{\delta \tilde{E}_{xc}(M)}{\delta \rho} \quad \text{and} \quad \frac{\delta \tilde{E}_{xc}}{\delta \rho_<} = \lim_{M \to N-1} \frac{\delta \tilde{E}_{xc}(M)}{\delta \rho} \tag{2.86}$$

and δ is the functional derivative. Eq. (2.84) is extremely insightful for two reasons. Firstly, it makes explicit why any approximation currently in use gives a wrong gap (even with (2.83)). It is in fact a well-known property that the derivative of the true xc functional has to be discontinuous, since it is dictated by the difference in the chemical potential of the N and $N \pm 1$ particle system. For a fractional occupancy the true xc energy is piecewise linear. In fact, it can be defined only in terms of a mixture of systems with integer number of electrons (see chapter 2 in reference [81]) and, as a consequence, it is not differentiable at the points with integer occupancy. LDA and GGA approximations describe instead the xc energy as a smooth function of the fractional occupancy and therefore they wrongly account E_g^{xc} to be zero. Secondly, it is clear from Eq. (2.84) that E_g^{KS} always underestimate the real gap. This justify the approach to start from the KS gap and try to "open it up" by means of some corrections to the DFT theory.

2.7.2 Strongly correlated systems and DFT+U

An incorrect gap is just one of the failures of DFT. It is in fact one of many symptoms of the most important problem of DFT: the approximated xc part. Other failures include, just to mention some: wrong magnetic ground-state predictions and bad description of electron localization over defects for transition metal oxides [82]; problems in capturing the properties of high-temperature superconductors in doped phase [83]; the failure in reproducing the metal to insulator transition under structural change in VO_2 [84] and the famous problem in describing the dissociation of the H_2 molecule. All the situations just mentioned are associated to the category of the so-called strongly correlated systems, that are known to be characterized by an important electron-electron correlation contribution. It is then the lack of correct correlation in LDA/GGA that brings DFT to fail in such cases. The latter observation is the motivating origin of one of the most popular corrections to DFT, called DFT+U [85, 86].

The DFT+U, also called Hubbard correction, aims at embedding in DFT a local correlation effect (Hubbard term) for some localized orbitals. In formula, the DFT+U energy

functional writes:

$$E_{DFT+U}[\rho] = E_{DFT}[\rho] + E_{hub}[\{\rho_i\}] - E_{dc}[\{\rho_i\}], \tag{2.87}$$

where, in its simpler formulation takes the form:

$$E_{hub}[\{\rho_i\}] = \frac{U}{2}\sum_{i\neq j}\rho_i\rho_j, \qquad E_{dc}[\{\rho_i\}] = \frac{U}{2}(N-1)N. \tag{2.88}$$

The sum runs over all the electrons belonging to the same shell (different shells gives different contributions). $E_{dc}[\{\rho_i\}]$ is called the double counting term necessary to remove the energy contribution of the orbitals in $E_{hub}[\{\rho_i\}]$ that is already included in $E_{DFT}[\rho]$. $N = \sum_i \rho_i$, so the double counting term is the mean field value of the Hubbard term. This correction is usually applied to partially filled d and f shells and U is a parameter that needs to be chosen for each shell. The effective implementation of DFT+U is much more complicate and there are many different flavors (a partial review is in reference [87]). DFT+U is not however used in any part of this PhD project.

2.7.3 The self-interaction problem and the ASIC

The lack of correct correlation is not, however, the only problem of DFT. This becomes clear by looking, for example, at the dissociation curve of the H_2^+ molecule (trend of the energy as a function of the two hydrogen atoms distance for the molecule without one electron). H_2^+ has just one electron, so no correlation effect is present. In fact, the system is correctly described by non-interacting theories. However, when calculated within the DFT framework, the wrong trend is obtained (see Figure 2.2). This is due to self-interaction problem.

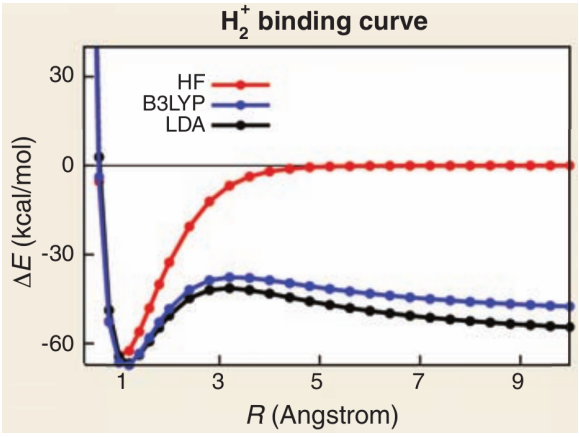


Figure 2.2: Binding energy of the H_2^+ molecule as a function of the H-H distance (R). HF corresponds to a non-interacting theory. LDA is the DFT calculation in local density approximation. B3LYP is a more sophisticated choice of the xc approximation. Picture adapted from reference [88].

We call self interaction the spurious interaction of an electron in a given KS orbital with the Hartree and xc potential generated by itself. Thinking of the hydrogen atom it is

clear that the only relevant interaction is the one with the external potential (nuclei), since no other electron is present. In DFT this translates into $v_H(\mathbf{r}) + v_{xc}(\mathbf{r}) \equiv 0$. However, because $\tilde{E}_{xc}[\rho]$ is only approximated, such cancellation does not take place. In fact, the ionization potential of the hydrogen atom in LDA-DFT is not -13.6 eV. The same is true for a many electron problem. The condition $J[\rho_i] + \tilde{E}_{xc}[\rho_i] = 0$ for the orbital density $\rho_i = \varphi_i^*(\mathbf{r})\varphi_i(\mathbf{r})$ is not satisfied. The obvious solution to this problem consists in removing the self-interaction part by defining a functional

$$E^{SIC}[\rho] = E[\rho] - \sum_{i}^{occupied} (J[\rho_i] + \tilde{E}_{xc}[\rho_i]) \tag{2.89}$$

However, this would lead to orbital-dependent KS equations. The dependence on the total ρ is lost and $v_{eff}(\mathbf{r})$ would contain explicit dependence on φ_i.

The introduction of two sets of orbitals, the usual KS, φ_i, and localized orbitals ϕ_i minimizing $\tilde{E}_{xc}^{SIC}[\rho] = \tilde{E}_{xc}[\rho] - \sum_i^{occupied}(J[\rho_i] + \tilde{E}_{xc}[\rho_i])$ is an appropriate solution [89, 90, 91] to the KS equations defined through Eq. (2.89). Now the localized orbitals will be used to evaluate the correction through $\rho_i = \phi_i^*(\mathbf{r})\phi_i(\mathbf{r})$. The correspondent KS equation then writes:

$$\left[-\frac{\nabla^2}{2} + v_{eff}(\mathbf{r}) + \sum_m v_m^{SIC}(\mathbf{r})\hat{P}_m^{\phi}\right]\varphi_i(\mathbf{r}) = \varepsilon_i\varphi_i(\mathbf{r}), \tag{2.90}$$

with

$$v_m^{SIC}(\mathbf{r}) = v_H(\rho_m, \mathbf{r}) + v_{xc}(\rho_m, \mathbf{r}). \tag{2.91}$$

Here ρ_m is associated to localized orbitals, and

$$P_m^{\phi}\varphi_i(\mathbf{r}) = \phi_m(\mathbf{r})\langle\phi_m|\varphi_i\rangle. \tag{2.92}$$

The idea is that the self-interaction-corrected potential for canonical orbitals is a weighted average of the self-interaction correction for localized orbitals and their weighting factor is the non-local projector P_m^{ϕ}.

When applied to extended systems, such correction method demands considerable additional computational costs, mostly because a larger cell is needed to describe the localized states. In order to make the correction more scalable some approximations can be introduced:

- Incorporating part of the correction into the definition of the pseudopotential. The atomic self interaction part of the core electrons is subtracted already in the free atom calculation that is necessary to generate the pseudopotential. This is consistent only after assuming that the projection of valence states the on core ones is negligible.

- The localized states are assumed to be atomic-like orbitals, Φ:

$$\sum_m v_m^{SIC}\hat{P}_m^{\phi} \rightarrow \lambda\sum_m \tilde{v}_m^{SIC}\hat{P}_m^{\Phi},$$

  where $\tilde{v}_m^{SIC}$ is defined in Eq. (2.91) but replacing ϕ with atomic-like orbitals Φ in the

expression for ρ_m and the projector P_m^Φ also refers to atomic-like orbitals. An empirical scaling factor λ is also introduced.

- The non-local projector is replaced with its expectation value, but considering the orbital occupation f_n:

$$P_m^\Phi \to \langle P_m^\Phi \rangle = p_m = \sum_n f_n \langle \varphi_n | P_m^\Phi | \varphi_n \rangle.$$

This scheme goes under the name of Atomic Self Interaction Correction (ASIC) [92, 93] and the associated KS equations are

$$\left[-\frac{\nabla^2}{2} + v_{eff}(\mathbf{r}) + \lambda \sum_m \tilde{v}_m^{SIC}(\mathbf{r}) p_m \right] \varphi_i(\mathbf{r}) = \varepsilon_i \varphi_i(\mathbf{r}). \tag{2.93}$$

Some comments are necessary. The first one is about the scaling factor λ. This quantity can be interpreted as a measure of the deviation of the ASIC potential from the exact SI potential of Eq. (2.91). Ultimately λ reflects the deviation of the ASIC projectors $|\Phi\rangle\langle\Phi|$ from the localized orbital $|\phi\rangle\langle\phi|$ defining the correct SI ground state. Therefore, one expects λ to be close to 1 for systems with an atomic-like charge density and to vanish for metals , whose valence charge is almost uniform. The second comment regards the total energy of the system. Because localized orbitals are assumed to be atomic-like and not obtained by minimizing $\tilde{E}_{xc}^{SIC}[\rho]$, the theory is not variational. There is no functional $E[\rho]$ from which equation (2.93) is derived by a variational principle. The expression adopted for the calculation of the total energy is (2.89) with $\rho_i = p_i |\Phi_i|^2$, but the physical interpretation of this energy is not trivial.

Although the total energy in this situation is an ill-defined quantity, the ASIC gives us a powerful tool to correct the gap by varying the parameter λ, proving to be very useful when the research requires a good description of the band gap more then a clear interpretation of the total energy. In the ASIC theory λ is an empirical factor, but it can hopefully be connected to other physical quantities. The high-throughput analysis of band gap calculations with ASIC is Section 3.2 has exactly the purpose to prove by means of data the correlation of λ with some easily accessible quantities and ultimately to give a tool to automatically chose the λ parameter for each calculation. A variational form of ASIC has been implemented [94], but is not necessary to the scope of this thesis as we do not have any other aspiration apart from using ASIC as a toll to open up the band gap.

We conclude the section saying that in reference [93] there is an interesting discussion about the formal connection between ASIC and DFT+U. This two methods start from different observation, but the final effect on the KS eigenvalues is pretty similar. We also need to remark that DFT, practiced in the ASIC or DFT+U approximations, is not a rigorous many-body theory anymore, but a mean-field theory (a very sophisticated mean field one could say). So in general one should be very careful in approaching them.

2.8 High-throughput approach

The technological development in many fields of science opened the possibility to measure/calculate in a "cheap" way quantities that were difficult to access in the past. For example early methods of DNA sequencing in 1970s could process only few thousands of pair of basis (kb) per day, while nowadays we reached the speed of 10^{12} kb/day [95]. When high volumes of data can be produced without much effort, it is convenient to calculate/measure properties of an entire category of choice (screening) and to analyze the data afterwards: this is the core of the high-throughput approach to science.

High-throughput differs from standard approaches because the focus moves from the study of a particular system (often with techniques not widely transferable), to the standardization and automatization of the calculation/measurement process (workflow) in order to target many different systems and to produce a database of quantities that can be consistently compared. Most of the science is then done *a-posteriori*, during the data analysis, where statistical tools can be used to find trends and correlation between quantities and ultimately guide the search for systems with ad-hoc characteristics. High-throughput screening is a solid reality in genomics and drug discovery, where the use of robotics and control softwares allow the researchers to conduct millions of chemical, genetics or pharmacological tests. In the field of material science, high-throughput screening is not yet popular for what concerns experimental studies. However, in recent years, the world of simulations is starting to explore this approach to science, the reason being the huge development in computational power of supercomputers.

2.8.1 High-throughput DFT Simulations: AiiDA

In 1964/65, when the papers of Hohenberg, Kohn and Sham pioneered the field of DFT simulations, the most powerful supercomputer was the Control Data Corporation (CDC) 6600. It was a relatively compact machine (about the size of four filing cabinets) and it had only one CPU able to perform about 3 million Floating Point Operations Per Second (3 megaFLOPS) [96]. Nowadays average laptops easily reach pick performances of hundreds of gigaFLOPS and the year 2018 marked the birth of the first exascale supercomputer. This exponential increase in computer power, brought DFT to be recently seen as an application for High-Throughput (HT) computing rather than for High Performance Computing (HPC) as it was in the past. The difference is significant. HPC requires the use of massive amount of computing for a short period of time, while the HT paradigm is characterized by the use of many small resources for a longer period of time. Moreover, the management of the data become a relevant part of the process.

It is now possible to use DFT (in combination with advanced thermodynamics and other electronic-structure methods) to create large databases of existing and hypothetical material properties and to combine them with intelligent data mining in the search of a material with tailored properties targeting the technology of interest. This is the foundation of the so called high-throughput materials design [97].

The creation of the database requires first of all a systematic way of submitting calcu-

lations and the interesting properties have to be stored in an intelligent way. The AiiDA (Automated interactive infrastructure and DAtabase) package [98] provides a framework for this purpose. The four pillars of the project are:

- Automation. The aim is to abstract away the tasks of preparing, submitting and retrieving large number of calculations. By means of python classes AiiDA is able to standardize structures, k-points, bands, input and output parameters. Once a computer and a code are set in AiiDA, a daemon running in the background of the local machine will manage the submission of the calculations.

- Data. AiiDA is able to parse the results of calculation and store them in a database. All the inputs and relevant information of the calculation are also saved ensuring to keep the origin of each calculation. A querying system is also provided to facilitate the analysis of a large amount of calculations.

- Environment. AiiDA provides the users with a way of coding all the workflow of one project. This implies to keep provenance of each step of the work and to ensure the reproducibility of an entire scientific process, not only of a single calculation. Complex sequences of steps involving more then one code can be coded in order to create turnkey solutions.

- Sharing. As registered user of AiiDA, one is able to share part of his local database with other users or to make it public. Also workflows can be shared. A web interface will be available to explore projects and calculations.

In this PhD work we performed high-throughput DFT calculations with the code SIESTA. The interface between AiiDA and SIESTA has been developed in collaboration with researchers of the Institut de Ciència de Materials de Barcelona (ICMAB) and it is publicly available at `https://aiidateam.github.io/aiida-registry/plugins/siesta.html`.

Of extreme importance in the high-throughput approach is the creation of data that are consistent among each other. With this we mean that a database needs to contain quantities calculated from the same level of theory. It would not have much sense, for instance, to compare DFT data coming from two different approximation of the xc functional. Moreover, in order to be compared, the data needs to have a certain standard of precision. This is the reason why it is important, before starting an high-throughput DFT study, to asses the possible sources of imprecision in the calculations. This is the topic of Chapter 3.

Chapter 3

High-throughput DFT

The very core of the high-throughput approach to science is the automatization of a the calculation/measurement process. In the case of Density Functional Theory (DFT) calculations, this translates into choosing automatically the DFT input parameters. We can divide the DFT inputs in 3 categories. The first one includes all the inputs that do not require testing; they are common to any DFT code, although often expressed in different formats in different codes. Examples of these are the crystal structure and the high symmetry paths along which to calculate the electronic/phononic dispersion relations. AiiDA takes care of these inputs without any effort from the user. Once the structure is passed to AiiDA from a file (in any format), dedicated modules deal with the structure, the high symmetry points in the Brillouin zone and the k-point path, passing them to the DFT code in the correct format. The second category gathers the parameters that determine the convergence, for instance the k-point mesh for the calculation of the integrals in reciprocal space, the energy cutoff, E_{CUT}, for plane-wave codes and others. The task of testing these parameters until reaching a satisfactory convergence is easily encodable in a workflow. However, it is often clear how to define ultra-converged values that are safe to be used in any situation. This choice is, of course, made at the expense of the calculation efficiency, but, especially in the case of simulations performed on small systems and with localized basis set codes, the computational time is not an issue. The last category includes inputs whose choice is definitely less systematic, for example the pseudopotentials (PPs), the basis set in localized basis codes or the U parameter in DFT+U. These parameters are always chosen starting from an educated guess, but often modified manually until one hits the desired target result. This is because there is not a clear strategy that guarantees an improvement of the results.

This chapter wants to critically reflect on the role of this latter category of DFT inputs in high-throughput material science and to present some ideas on how to make systematic their selection. The work done on the generation of optimal SIESTA PPs and the choice of the correct λ parameter for Atomic Self Interaction Correction (ASIC) simulations is presented in Section 3.1 and Section 3.2 respectively.

A detailed analysis of the choice of the SIESTA basis set would also belong to this chapter, however not enough work has been done on the topic to justify the drafting of

a dedicated section. Technical work has been done with the purpose to create a stand alone basis manager to include in AiiDA that would facilitate systematic test of basis sets. However, this is not ready yet for production purposes. Once the computational tool is ready, two possibilities are foreseeable: to test the basis set versus a plane wave reference or to utilize an improved version of the "simplex" approach, whose details are in reference [99]. In any case, the choice of the ideal basis set is very much system dependent (the optimal basis of O in MgO or in O_2 is different). Any of the two approaches above should be performed before SIESTA calculations on any new system, with evident computational costs. The basis set choice automatization is the missing piece needed to complete the setting up of a rigorous procedure to fully automatize the SIESTA DFT calculations, with guaranteed standard of precision. For the results of this work, a DZP basis set (two ζs for each angular momentum in the valence plus a polarized orbital, see Section 2.5) has always been used, with no testing on the reliability of the choice.

3.1 Optimized pseudopotentials

In this section we summarize the work done in order to generate a set of reliable norm-conserving PPs to use for high-throughput SIESTA simulations. The level of approximation for the exchange-correlation (xc) potential is the Local Density Approximation in the implementation of Perdew and Zunger (LDA-PZ) [52]. From the PP theory highlighted in Section 2.3 we have learned that the generation of a PP starts from a DFT calculation for the isolated atom and its transferability is ensured by various means. This justifies the expectation for a PP to be relatively system independent, meaning it will behave well no matter what chemical environment it is placed into. Limitations to the latter statement, like the situations where semicore states are need, have been discussed in Section 2.3. It is then reasonable to work on the creation of a set of reliable PPs to be used for all possible systems. The electronic structure community is working actively in this direction. Moreover, in accordance with the high-throughput spirit, it is also important to determine a measure of the accuracy of each PP. To this end, in recent years, the community adopted a standard to test PPs and, more in general, to compare DFT calculations across different platforms: the Delta test described in reference [100, 101]. More details on this method will follow in subsection 3.1.2. The basic idea is to compare the equations of state of standard crystal elements calculated with the PP to be tested and a reference all-electron calculation. The reliability of the Delta test is debatable; one could, for example, claim that the correct description of the occupied band dispersions is an important factor to include in the process of judging the accuracy of a PP, while the Delta test considers only quantities deriving from the total energy. This criticism is surely legitimate, but there is no doubt that the Delta test is now fully accepted in the electronic structure community as the standard for DFT calculations comparison.

Recent efforts to evaluate the accuracy of PP libraries based on the Delta test are the PPPS project [102] and the PseudoDojo project [103]. Both these efforts are focused on testing PPs in the so called Unified Pseudopotential Format (UPF), a format created

with the purpose to establish a flexible file standard where to store different type of PPs: norm-conserving, ultrasoft, PAW and others. The format is "unified" in the sense that it is a common platform for different kind of PPs, but it is not universal because up to now only a few DFT codes are able to import UPF PPs. For example SIESTA uses a different standard, the 'Froyen' (PSF) format.

To the best of our knowledge there is no available library of tested LDA-PZ PPs in PSF format and this is the starting point of our work. A straightforward application of the Delta test using the SIESTA code is, in our opinion, not fully reliable because we consider important to decouple the effect of the PP from that of the basis set choice. In fact, while one expects the PP to be relatively system independent, the optimal basis set could dependent on the chemical environment. In order to tackle this problem we do not see any other possibility than to test first PPs using a plane-wave code (namely Quantum Espresso - QE [104]) that imports UPF PPs and then convert the optimized pseudo in the SIESTA-readable PSF format. Of course the performance of any plane-wave code depends as well on other parameters, but, in contrast to the SIESTA basis, they are all of the second kind defined in the introduction. The use of ultra-converged parameters enable us to isolate the effects of the PP.

As we are not aware of any LDA-PZ UPF library as well, we decided here to bring the analysis a step further and create a workflow to automatically generate optimized PPs in both UPF and PSF format. The fundamental tool for our investigation is the PP generator Atomic Pseudopotentials Engine (APE [105]), which allows us to generate from the same input a UPF and a PSF file. The simple idea is to test many UPF PPs comparing the QE calculation with an all-electron code. The aim is to find for each element in the periodic table a choice of the PP generation parameters which gives an accurate PP. Once the optimum UPF has been chosen, the corresponding PSF PP is assumed to be optimum as well. In theory, any discrepancy between a SIESTA calculation with the optimized PP and the corresponding all-electron calculation should be associated to the basis set. We anticipate, however, that the procedure described above is not perfect because of a technical difference between PSF and UPF formats. PSF contains the PPs in a semi-local form (see Eq. (2.41)), while into UPFs there is already the division between the local part and the non-local projectors. SIESTA constructs such separation inside the code and there is no guarantee that the choice is same as that used for the UPF. Hence, we have two PP coming from the same semi-local form, but effectively used with two different strategies concerning the separation into local Kleinman and Bylander (KB) projectors. Unfortunately there is no way at the moment to measure the effect of the change in KB projectors within the same code. However an *a-posteriori* comparison of the local part makes us confident that the effect should be small. Further efforts in order to unify the PP treatment among different codes are *in fieri* in the electronic structure community, although the decisive step still has to be taken [106].

Before going into the details of the exact procedure for the PP generation, some technical aspects must be mentioned. As mention above, the LDA-PZ is the *xc* approximation of choice. This is motivated by the fact that our next step in the workflow is the creation

of PPs including the ASIC correction which is implemented for LDA-PZ PPs only. The crystal structures of the reference crystal elements used throughout the entire work are available in the supplementary material of reference [100]. They are the same used in reference [101] to prove the reproducibility of DFT results coming from different codes.

The steps in our workflow are outlined in the following subsections:

- Construct a set of all-electron references computed by using the code FHI-AIMS [107] (subsection 3.1.1)
- Generate the UPF optimized PPs by varying the cutoff radii (subsection 3.1.2)
- Perform SIESTA calculation tests with the PSF optimized PPs (subsection 3.1.3)
- Generate the same PP with a modified version of the PP generator ATOM [108], able to include the ASIC (subsection 3.1.4)

During the presentation of the outcomes of the study, explicit results on C (graphite) only are reported. When possible, general consideration are drawn at the end of each section and specific problematic cases are discussed.

3.1.1 All-electron calculations: FHI-AIMS

The code of choice for the generation of the all-electron references is FHI-AIMS, an all-electron code using numerically tabulated atom-centered orbitals as basis set. Despite being a localized basis set code, FHI-AIMS presents a substantial difference with respect to SIESTA: it uses a standard set of basis for each element of the periodic table that has been chosen through a complex iterative construction procedure. In other words, no intervention by the user is necessary for the basis selection. The claim is that the choice is accurate up to a certain energy threshold because a new element is added iteratively to the basis in a way that ensures a variational improvement of a target total energy [107]. Moreover the choice is transferable since the optimization target is the total energy of a set of elemental dimers having various bond distances. The functions "pool" from which to select new basis elements consists of valence and excited-state radial functions of single atoms with different confining potentials, radial functions of cations and hydrogen-like radial functions. Each of the described radial functions (u_{il}) are associated to an angular momentum channel l. When included in the basis set, all the $2l+1$ functions $u_{il}Y_{lm}$ become part of the basis (Y_{lm} are the spherical harmonics) [107].

If one trusts the method above, the localized basis becomes no longer an input of the third kind. The choice is then only between (using the FHI-AIMS terminology) "tiers" of accuracy. Of course, more accurate tiers contain larger basis, with the consequent increase in the computational demands. Thanks to the features described, FHI-AIMS was part (in contrast to SIESTA) of the set of codes used in the study presented in reference [101], where the reproducibility of DFT results among different codes was proved. In that work ultra-converged calculations are performed and the input parameters are clearly listed in the supplementary material (page 23 of the supporting information of reference [101]). We use the same set of parameters for our calculations:

	KP Mesh	Num KP	V_0(Å^3/atm)	B_0(GPa)	B_1	B_0(eV/Å^3)	gap(eV)
H	28x28x20	7840	17.9925	9.4042	2.66582	0.058696	Dir 9.10
He	40x40x22	17600	10.2231	4.7208	5.80302	0.029465	Ind 17.37
Li	38x38x38	27436	18.9509	15.1335	3.49195	0.094456	Metal
Be	52x52x28	37856	7.54916	131.8402	3.32551	0.822882	Metal
B	26x26x24	8112	7.00116	251.8718	3.47013	1.572060	Ind 1.41
C	48x48x12	13824	11.3427	220.5460	3.51255	1.376540	Metal
N	16x16x16	2048	28.0819	56.4927	3.6754	0.352600	Ind 7.61
O	26x24x24	7488	16.1803	68.6392	4.11436	0.428412	Metal
F	16x28x14	3136	17.5268	40.7442	3.97679	0.254305	Dir 3.32
Ne	22x22x22	5324	14.3840	8.9515	7.11756	0.055871	Dir 12.00
Na	32x32x32	16384	33.3669	9.2102	3.74247	0.057485	Metal
Mg	36x36x20	12960	21.6645	39.7139	3.92472	0.247875	Metal
Al	24x24x24	6912	15.7994	84.0507	4.35479	0.524603	Metal
Si	32x32x32	16384	19.7096	96.9336	4.24656	0.605012	Ind 0.47
P	30x8x22	2640	20.3054	79.1429	4.46103	0.493971	Ind 0.02
S	38x38x38	27436	16.1744	100.9841	3.93925	0.630293	Metal
Cl	12x24x12	1728	35.6435	22.7477	4.46749	0.141980	Dir 2.31
Ar	16x16x16	2048	30.1682	7.1706	7.09083	0.044755	Ind 8.76
K	20x20x20	4000	64.4324	4.5203	3.77266	0.028213	Metal
Ca	18x18x18	2916	37.8754	17.8861	4.70942	0.111636	Metal

Table 3.1: Equation of state parameters obtained from non-relativistic AIMS calculations. The second and third columns report information about the k-point mesh used in the calculation. V_0 is the equilibrium volume per atom, B_0 the bulk modulus per atom (reported in both units of GPa and eV/Å^3) and B_1 its derivative respect to the pressure. The last column reports information about the KS band gap calculated at the equilibrium volume.

- A gaussian smearing with a fictitious broadening corresponding to 0.01 eV for the description of the occupation at the Fermi Energy
- A basis set tier2
- The k-point mesh for the calculation of integrals in reciprocal space is different for each element. They are listed in Table 3.2.

The study reported in reference [101] refers to Perdew-Burke-Ernzerhof (PBE [109], a type of GGA *xc* functional) scalar relativistic calculations, but we assume that for LDA calculations the same set of parameters are enough for convergence. The choice is motivated by the fact that in the original FHI-AIMS paper [107] the possibility of different basis for different *xc* potentials was discussed and the conclusion is that the optimized basis does not depend on the *xc* functional. Moreover the converged k-points mesh might depend on whether the system is metallic or not and we do not see any crystal element changing from insulator to metal when calculated with PBE or LDA.

In Table 3.1 the result for non-relativistic calculations for elements up to Ca are presented. The equilibrium volume per atom, V_0, and the bulk modulus per atom, B_0, listed in the table are obtained from a fit of the energy versus volume (per atom) curve calculated for 7 equidistant volumes. The function used for the fit is the Birch–Murnaghan equation of state [110]. Table 3.2 reports the same results for elements up to Sr and also including In, Sn, Sb, Te, I, but calculated at a scalar-relativistic level. Other relativistic effects (spin-orbit interaction, requiring a full-relativistic treatment) are expected to be relevant for heavier elements and they are not treated in this work.

The difference between scalar relativistic and non-relativistic calculations is quantified in Figure 3.1, where the relative percentage error for the bulk modulus per atom and the

	KpointMesh	NKpoint	V_0(Å^3/atm)	B_0(GPa)	B_1	B_0(eV/Å^3)	gap(eV)
H	28 x 28 x 20	7840	17.991	9.4051	2.66564	0.0587022	Dir 9.10
He	40 x 40 x 22	17600	10.2222	4.7208	5.80355	0.0294648	Ind 17.37
Li	38 x 38 x 38	27436	18.9438	15.1420	3.4783	0.0945091	Metal
Be	52 x 52 x 28	37856	7.54736	131.8759	3.31016	0.823105	Metal
B	26 x 26 x 24	8112	6.99888	251.8894	3.47165	1.57217	Ind 1.41
C	48 x 48 x 12	13824	11.3366	220.5765	3.51322	1.37673	Metal
N	16 x 16 x 16	2048	28.0685	56.4604	3.67639	0.352398	Ind 7.61
O	26 x 24 x 24	7488	16.1742	68.6743	4.09238	0.428631	Metal
F	16 x 28 x 14	3136	17.5334	40.7533	3.96138	0.254362	Dir 3.31
Ne	22 x 22 x 22	5324	14.3708	8.9751	7.11036	0.0560182	Dir 11.96
Na	32 x 32 x 32	16384	33.2666	9.2508	3.74189	0.0577388	Metal
Mg	36 x 36 x 20	12960	21.6337	39.6912	4.04086	0.247733	Metal
Al	24 x 24 x 24	6912	15.8049	83.0993	4.44734	0.518665	Metal
Si	32 x 32 x 32	16384	19.7011	96.3139	4.29096	0.601144	Ind 0.45
P	30 x 8 x 22	2640	20.3324	78.3679	4.46079	0.489134	Ind 0.02
S	38 x 38 x 38	27436	16.1598	101.5141	4.01195	0.633601	Metal
Cl	12 x 24 x 12	1728	35.6652	22.6948	4.48824	0.141649	Dir 2.27
Ar	16 x 16 x 16	2048	30.0872	7.2191	7.08069	0.0450584	Ind 8.68
K	20 x 20 x 20	4000	63.9726	4.5601	3.76431	0.0284622	Metal
Ca	18 x 18 x 18	2916	37.9319	17.7643	3.64172	0.110876	Metal
Sc	34 x 34 x 20	11560	22.5265	62.0289	3.30935	0.387154	Metal
Ti	40 x 40 x 22	17600	16.1176	126.101	3.51946	0.787066	Metal
V	34 x 34 x 34	19652	12.5387	210.7375	3.62276	1.31532	Metal
Cr	36 x 36 x 36	23328	10.8543	300.4001	4.19975	1.87495	Metal
Mn	28 x 28 x 28	10976	10.0892	333.9641	4.60174	2.08444	Metal
Fe	36 x 36 x 36	23328	10.3764	252.0320	4.92541	1.57306	Metal
Co	46 x 46 x 24	25392	10.0044	274.9944	4.58965	1.71638	Metal
Ni	28 x 28 x 28	10976	10.0352	254.1116	4.89247	1.58604	Metal
Cu	28 x 28 x 28	10976	10.9025	188.4881	4.9375	1.17645	Metal
Zn	44 x 44 x 20	19360	13.6485	103.4817	5.36299	0.645882	Metal
Ga	22 x 12 x 22	2904	18.386	66.9038	5.21028	0.417581	Metal
Ge	30 x 30 x 30	13500	22.229	71.8192	4.76465	0.44826	Dir 0.08
As	30 x 30 x 10	4500	21.1477	80.7792	4.23619	0.504184	Metal
Se	26 x 26 x 20	6760	27.6842	57.0609	4.44972	0.356146	Ind 1.01
Br	12 x 24 x 12	1728	35.797	29.6201	4.86353	0.184874	Dir 1.30
Kr	16 x 16 x 16	2048	37.8194	6.9040	6.99958	0.0430917	Dir 6.88
Rb	18 x 18 x 18	2916	77.4247	3.5749	3.52628	0.0223129	Metal
Sr	16 x 16 x 16	2048	48.3809	14.5859	3.44484	0.0910378	Metal
In	30 x 30 x 20	9000	24.6233	49.9413	4.75486	0.311709	Metal
Sn	26 x 26 x 26	8788	33.9092	45.1266	4.92782	0.281658	Ind 0.04
Sb	26 x 26 x 8	2704	29.6583	60.3689	4.46212	0.376793	Ind 0.02
Te	26 x 26 x 16	5408	32.8261	54.3873	4.56894	0.339459	Dir 0.23
I	12 x 22 x 10	1320	45.5322	25.2516	5.0187	0.157608	Ind 0.90

Table 3.2: Equation of state parameters obtained from scalar-relativistic AIMS calculations. The explanation of the various columns is in the caption of Table 3.1.

equilibrium volume per atom is plotted. Relativistic effects are definitely not relevant for elements up to Mg as the error of both V_0 and B_0 is below 0.5 %. For the 3p series the error on the bulk modulus is more significant, however, still very small. Moreover the error on the volume is basically zero. The decision is then to treat elements up to Ar in a non-relativistic manner while all the other elements are calculated at the scalar-relativistic level.

In looking at our results a few comments needs to be made. Firstly, the LDA equilibrium volumes are very different with respect to PBE (smaller). This is a known shortfall of LDA, which tends to over-bind. Secondly, the equation of state is computed by just varying the volume, no atom relaxation or strain minimization are considered. This may result in errors in the bulk modulus, but it was the same choice adopted in reference [101]. Finally, all the calculations are performed without spin polarization except for Fe, Co and Ni for which the ferromagnetic ground state is considered. In reference [101] the elements

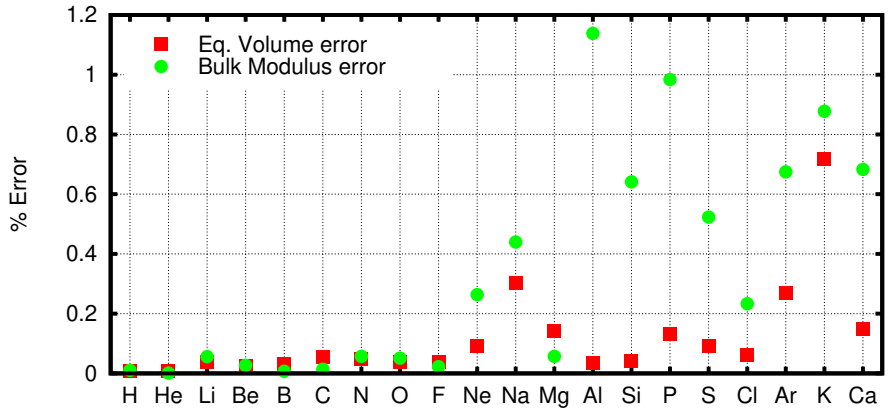


Figure 3.1: Percentage difference between scalar-relativistic and non-relativistic calculations (perform with FHI-AIMS) for elements up to Ca. This is defined as 100 $\delta X/\overline{X}$, where δX is the difference between quantity X computed with scalar-relativistic and non-relativistic and $\overline{X}$ is their average.

O, Cr and Mn are calculated in the antiferromagnetic configuration, here the non-magnetic ground state is explored instead.

The results in the Table 3.1 and Table 3.2 are the all-electron reference used for the following steps of the work.

3.1.2 Identification of optimal UPF PPs by means of QE calculations

The second step of our study consists in generating the UPF PPs using APE in using them in QE in order to calculate the equation of state. The PP radii are varied until a PP with sufficient accuracy is obtained.

Firstly, let us discuss some technical aspects. Non-linear core corrections are always included; the core radius cutoff is not an input parameter for the generation of the PPs because it is chosen automatically by APE as the radius where the core charge is 1.5 times the valence one. The inputs are then only the PPs cutoff radii, in theory one for each angular momentum channel. Experience shows that only the cutoff radii of bound states are relevant (even if empty). Unbound states can be recognized by looking at the wave function. They present a free-electron form that is practically zero in the vicinity of $r = 0$ (see Figure 3.9). For each element up to Ca, the cutoff radii of f and d channels are fixed to a value took from literature, while the cutoff radii associated to p and s orbitals are varied around the outermost maximum (peak) of the all-electron wave functions. The choice is due to the fact that a cutoff around the outermost peak of the wave-functions leads to PPs showing a good compromise between accuracy and efficiency. For $3d$ transition metals also the cutoff associated to d states is varied. In the latter situation though, the outermost peak of the wave function is situated very close to $r = 0$, creating PPs that are unnecessarily inefficient in term of computational performances. The strategy then is to vary the cutoff radius around two times the outermost peak.

The QE calculation parameters are also extracted from the supplementary material (page 139) of the paper in reference [101], the most important being the energy cutoff

E_{CUT} equal to 100 Ry and the k-point meshes that turn out to be the same as the one used for the FHI-AIMS calculations (see Table 3.2). By using the same argument as the once employed with FHI-AIMS, we consider the ultra converged PBE parameters good enough for the LDA as well.

The ultimate quantity that defines the accuracy of the PP is the already mentioned Delta value:

$$\Delta = \sqrt{\frac{\int_{0.94V_0}^{1.06V_0} (E_{AE}(V) - E_{PP}(V))^2 dV}{0.12V_0}}. \tag{3.1}$$

The quantity condenses in one single number the difference in the equation of states calculated with the PP with respect to the all-electron one. Here, V_0 is the equilibrium volume of the all-electron calculation, E_{PP} is again from the Birch–Murnaghan fit of the energy per atom calculated with th PP for at 7 equidistant volumes, and E_{AE} is the all-electron reference. The Δ is naturally larger for elements with large bulk modulus because the equation of state over a given volume interval reaches higher energies. To take this factor into consideration, rescaled quantities have been proposed [111, 112]. However, in order to maintain a quantitative comparison with the study in reference [101], we decided here to use the original definition contained in Eq. (3.1) and report the bulk modulus together with the calculated Δ. In fact, in reference [101] a discussion about the quantification of the error associated to a certain value of Δ is outlined. By taking averages over the entire set of 71 crystal elements considered in the study, the authors were able to estimate that a $\Delta = 1$ meV corresponds to, on average, a volume per atom deviation of 0.14 Å^3 and a bulk modulus deviation of 1.6 GPa. Moreover high-quality experimental data on Cu, Ag and Au shows errors corresponding to $\Delta = 1$ meV. This allows us to conclude that a value for Δ of 1 meV is to be considered very small, with the additional remark that Δ is expected to be small for materials having a small bulk modulus.

As a final note, we remark here that, in contrast to other works, we are not concerned about the efficiency of the PPs because we are aiming to use the PPs for high-throughput simulations of small elementary cells. Our main target is to maintain a reliable and certified accuracy across the entire dataset.

The entire procedure for the selection of the optimized PP for each element of the periodic table is implement in an AiiDA workflow. From a first APE calculations one obtains the outermost peaks of the all-electron wave function for the angular momentum channels of interest. Subsequently the workflow creates a 4×4(×4) mesh of test cutoff radii that are used to generate the PPs, set the equations of state and calculate the Δs. A minimum accuracy of $\Delta = 1$ meV is set as a threshold. If no combination of radii in the mesh satisfies the criteria, new radii will be generated and new calculations performed until a satisfactory result is achieved.

The typical 3D plots showing the dependence of V_0 and B_0 on the various cutoff radii (s and p in the presented case of C) used in the generation of the PP are reported in Figure 3.2 and Figure 3.3 (red surface). We observe that in the case of C, the volume decreases when the radius of the s and the p channels increases. The bulk modulus, instead, has an opposite behavior comparing s and p radii, it decreases while p becomes larger and

increases when s increases. Of course our speculation is limited to cutoff radii in the close vicinity of the outermost maximum of the wave functions (equal to 1.20 Bohr for both s and p channels in case of C) as we do not have information for other radii. The trend just described is observed for the first elements of the second period of the periodic table (namely from Li to N), but it can not generalized further. No other clear trend is observed concerning the dependence of volume and bulk modulus on the cutoff radii. In Figure 3.2 and Figure 3.3 also the surfaces representing the all-electron reference value are reported (blue surface). The equilibrium volume of the reference is matched in the range of cutoff radii considered here, the same can not be said for the bulk modulus. However, the difference between the reference B_0 and the one obtained with any of the PPs considered here is very small, less than 0.8 %.

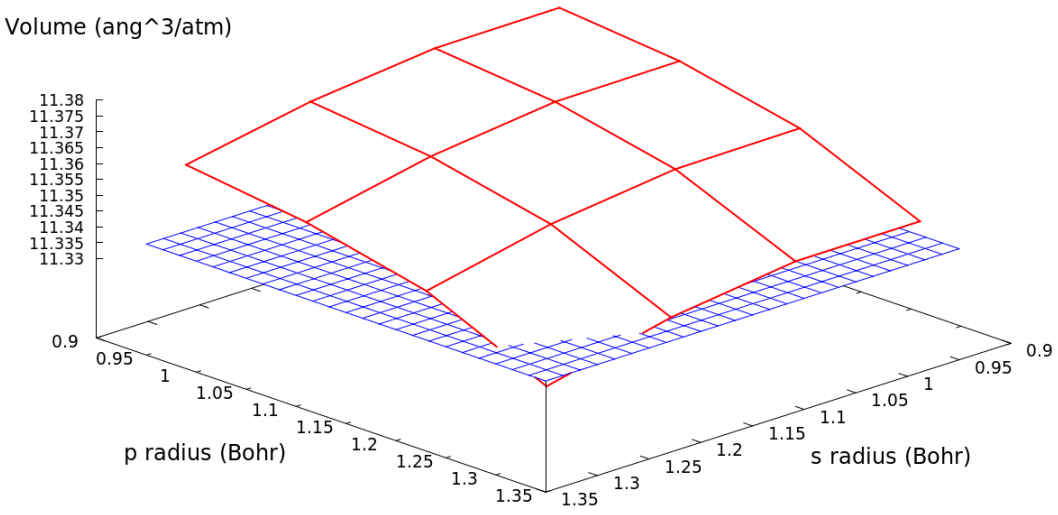


Figure 3.2: Equilibrium volume of C (graphite) as a function of the PP p and s radius (red curve). The AIMS reference of 11.3427 Å^3/atm is indicated by the blue surface.

Figure 3.4 reports the dependence of the Δ parameter on the cutoff radius of the s and p channels used for the generation of the PPs. Like the previous two pictures, it refers to the element C (graphite). The blue surface now sets the zero. We notice that the Δ follows more or less the trend of the equilibrium volume in Figure 3.2. This means that the change in volume is the most relevant error in the case in analysis. The only difference is the presence of a minimum, at which the surface is not differentiable. This is due to the fact that the equilibrium volume reaches the reference value and then it keeps decreasing, assuming values that are more distant respect to the reference. Consequently Δ shows a minimum.

At the end, for C, the PP which gives best agreement with all-electron calculation has been generated with the following parameters: cutoff of the s, p, d, f orbitals are respectively 1.20, 1.31, 1.56, 1.56 Bohr. Giving a Δ of 0.13 meV. A closer look at the results is presented in Table 3.3.

A *a-posteriori* comparison between the band structure calculated with FHI-AIMS and QE with the optimized PP is in Figure 3.5). The comparison is more than satisfactory to

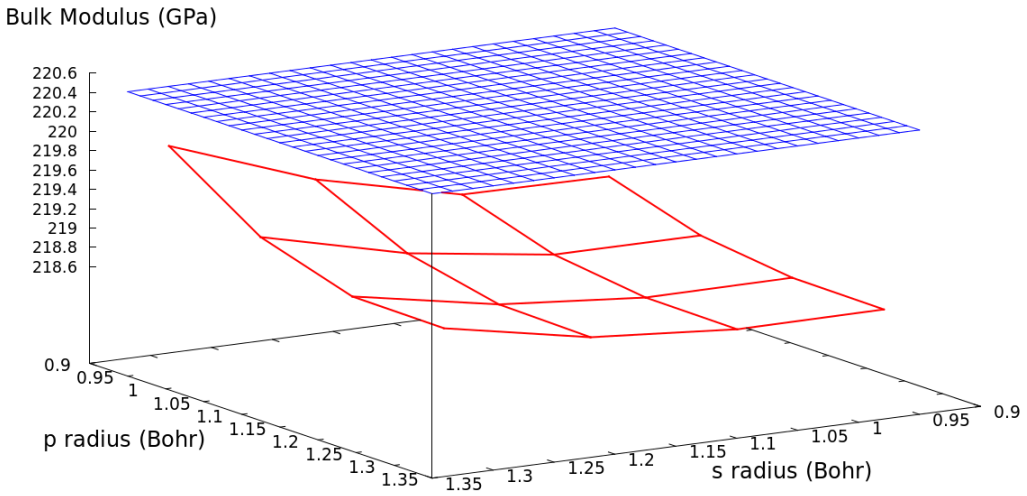


Figure 3.3: Bulk modulus of C (graphite) as a function of the PP p and s radius (red curve). The AIMS reference of 220.5460 GPa is indicated by the blue surface.

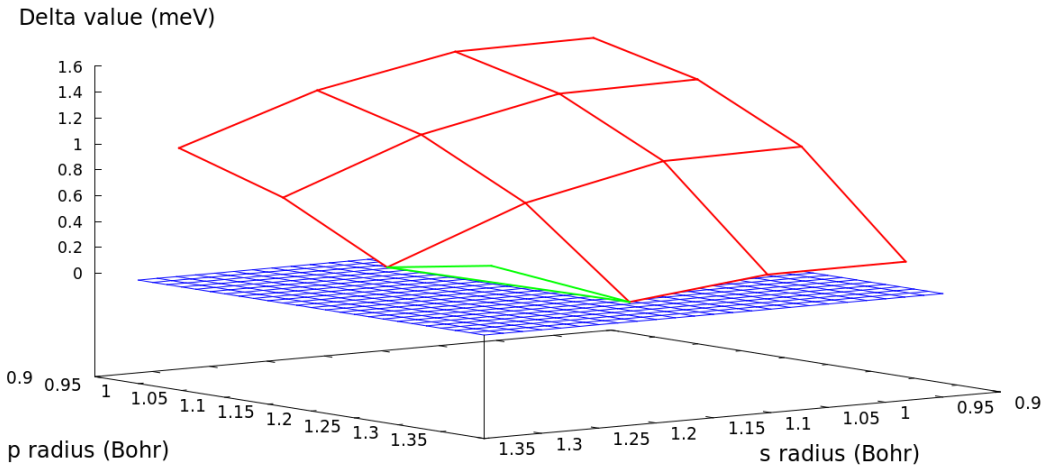


Figure 3.4: Delta parameter (Δ) of C (graphite) obtained from the comparison between QE and AIMS results as a function of the QE PP radius. The blue surface is the ideal situation of Δ=0.

	V_0(Å^3/atm)	B_0(GPa)	B_1	B_0(eV/Å^3)
AIMS Reference	11.3427	220.5460	3.5125	1.3765
Optimized UPF	11.3447	218.7432	3.5224	1.3651

Table 3.3: Equation of state results comparison between the all-electron AIMS calculation and the QE calculation obtained using the PP generated at the end of the optimization procedure. The explanation of the various columns is in the caption of Table 3.1.

capture the main electronic features of graphite. This reassures on the fact that it is not necessary to include a comparison of the bands in the optimization procedure. The good description of the equation of states translates in a good description of the bands as well, at least for the present case of graphite.

Table 3.4 presents the cutoff radii for the optimal PPs identified by the automatic pro-

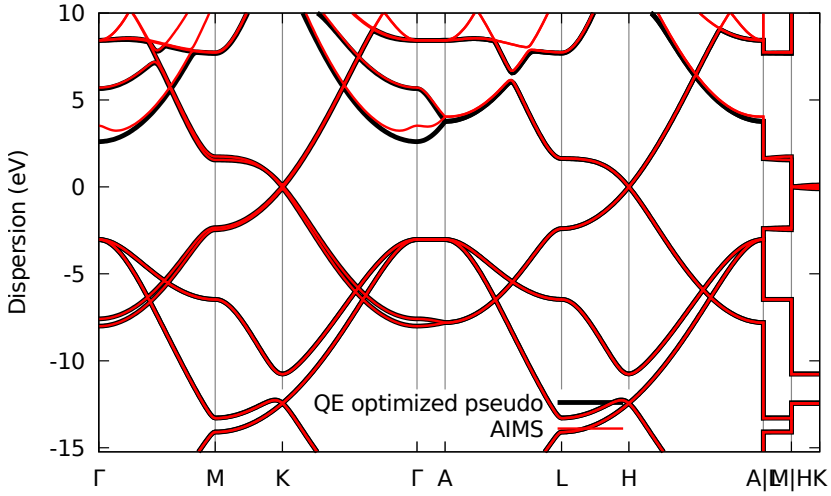


Figure 3.5: Band structure of C (graphite). Comparison between the bands computed with AIMS all-electron code and QE with the PP obtained though the optimization procedure described in this section. Fermi energy at 0 eV.

cedure for elements from Li to Ge. Furthermore the corresponding Δ values are reported. It is necessary to underline how the PPs obtained are not the absolute best in terms of Δ, since the algorithm stops when a Δ smaller than 1 meV has been achieved. This explains also the fact that there is no correlation with the value of Δ obtained and the magnitude of the bulk modulus (also reported in Table 3.4). A further refinement is possible if even more accurate PPs are needed. Two peculiar cases need to be discussed, the noble gases Ar and Ne. For them a PP have been produced, but the Delta test was not performed because they present a discontinuity in the equation of state calculated with QE. In QE the number of plane wave depends on the volume of the cell under investigation and this might lead to discontinuities on the equation of state for not converged calculations. At convergence, however, this should not happen. For this reason we questioned the choice of the plane-wave cutoff of 100 Ry for these situations. We increase the cutoff to 150 Ry, but we obtained the same discontinuity. The two noble gases are pathological cases, they are in fact kept together by Van der Waals forces that are not described in DFT. Their minimum in the DFT equation of state is then so shallow that not even a cutoff of 150 Ry is able to produce accurate enough calculations to capture it.

Except for these two cases, we where successful in generating a set of optimized PPs for the element of the periodic table from Li to Ge. The PPs, obtained in UPF format, have a certified accuracy of $\Delta < 1$.

3.1.3 Optimized pseudopotatntial in PSF format

Once the optimal UPF PP is selected, the corresponding PSF, generated automatically by APE, is considered accurate as well. In order to verify this assumption, the SIESTA equation of state employing DZP as a basis set is calculated. The ultra-converged parameters for the SIESTA calculation are: mesh cutoff 900 Ry, k-points as in AIMS and QE (see Table 3.1).

	s	p	d	f	Δ (meV)	B_0(GPa)
Li	2.69	2.26	2.59	2.59	0.47	15.1335
Be	1.85	1.85	2.43	2.43	0.09	131.8402
B	1.37	1.37	1.76	1.76	0.26	251.8718
C	1.20	1.31	1.56	1.56	0.13	220.5460
N	1.31	1.26	1.48	1.48	0.37	56.4927
O	1.23	1.34	1.01	1.01	0.95	68.6392
F	1.45	1.26	1.10	1.10	0.39	40.7442
Ne	-	-	-	-	-	8.9515
Na	2.56	3.41	3.13	3.13	0.46	9.2102
Mg	2.54	2.99	2.59	2.59	0.33	39.7139
Al	1.86	2.65	2.22	2.22	0.92	84.0507
Si	1.62	2.57	2.11	2.11	0.79	96.9336
P	1.56	2.18	2.02	2.02	1.00	79.1429
S	1.42	1.77	1.94	1.94	0.94	100.9841
Cl	1.28	1.43	1.88	1.88	0.24	22.7477
Ar	-	-	-	-	-	7.1706
K	2.64	0.81	3.33	3.33	0.05	4.5601
Ca	2.66	0.94	2.66	2.66	0.60	17.7643
Sc	2.44	3.23	1.31	2.48	0.28	62.0289
Ti	2.32	2.68	1.18	2.48	0.89	126.101
V	2.48	2.55	1.41	2.48	0.77	210.7375
Cr	2.03	2.51	1.92	2.48	0.65	300.4001
Mn	1.61	2.20	1.79	2.2	0.18	333.9641
Fe	2.00	2.47	1.70	2.00	0.24	252.0320
Co	2.38	2.44	2.23	2.23	0.54	274.9944
Ni	2.70	2.38	2.15	2.15	0.38	254.1116
Cu	1.80	2.10	1.92	1.9	0.57	188.4881
Zn	2.01	1.68	1.64	1.9	0.90	103.4817
Ga	2.18	2.35	2.18	2.56	0.89	66.9038
Ge	2.13	2.55	2.38	2.16	0.76	71.8192

Table 3.4: Cutoff radii (in Bohr) for the optimized PPs. The Δ value is obtained from the comparison between the all-electron calculation and the QE calculation with the optimized PPs. The bulk modulus (in GPa) is from the AIMS reference.

	V_0(Å^3/atm)	B_0(GPa)	B_1	B_0(eV/Å^3)	Δ (meV)
AIMS Reference	11.3427	220.5460	3.51255	1.376540	-
DZP	11.6173	227.5084	3.49473	1.42000	13.47
DZDP	11.4190	227.2668	3.64248	1.41848775	3.83

Table 3.5: Volume per atom V_0, bulk modulus per atom B_0 (in GPa and eV/Å^3) and the bulk modulus derivative B_1 for the crystal element C (graphite). We present the comparison between all-electron AIMS and SIESTA calculations employing the optimized PP.

The result (reported in the second line in Table 3.5) is disappointing as it leads to a value of Δ of 13.47 meV, not even close to the level of accuracy displayed by the corresponding PP in UPF format. However, by increasing the richness of the basis (DZDP - third line in the table), a significant improvement is achieved. This finding supports our hypothesis that the error on the basis set is relevant and it is important to try to separate it from the PP accuracy.

We here tested also the effects of a minimal variation of the cutoff of the basis, in particular we increase and decrease the cutoff radius of the s orbital of C by 5% with respect to the automatic cutoff of DZDP. The p orbitals and their polarized ζ remain the same. The results presented in Table 3.6 show a change in the equation of state parameters that is not large, but still significant. We need to recall, however, that among all the crystal elements considered in this study, C is one of the few with very large value of bulk modulus. For such crystal elements a bigger variability of Δ is expected, as we

	V_0(Å^3/atm)	B_0(GPa)	B_1	B_0(eV/Å^3)	Δ (meV)
AIMS Reference	11.3427	220.5460	3.51255	1.376540	-
DZDP stand	11.41897	227.2668	3.6424769	1.41848775	3.83
DZDP long	11.44216	224.2840	3.56149	1.3998707	4.87
DZDP short	11.417312	227.3124	3.6427386	1.41877	3.75

Table 3.6: Table reporting the volume V_0, bulk modulus B_0 (in GPa and eV/Å^3) and the bulk modulus derivative B_1 for the crystal element C (graphite). The SIESTA calculations are performed employing the optimized PP and three different choices of the basis cutoff for the s orbitals.

have already mentioned in the previous subsection.

An additional source of difference, as explained in the introduction, might be the different treatment of the local part of the PP in SIESTA compared to the UPF format. SIESTA makes the separation inside the code, so only after running the calculation one can access the form of the local potential (for example from the file C.ion.xml under the name ReducedVLocal). The local part in the UPF is instead directly accessible from the PP file. A comparison between the two is in Figure 3.6. The difference is not large, but still relevant. As mentioned in the introduction, there is no way at the moment to analyze the influence on the calculation accuracy of the different local part of the PP separately from the basis set effect. Similar magnitude of the difference are observed for other crystal elements.

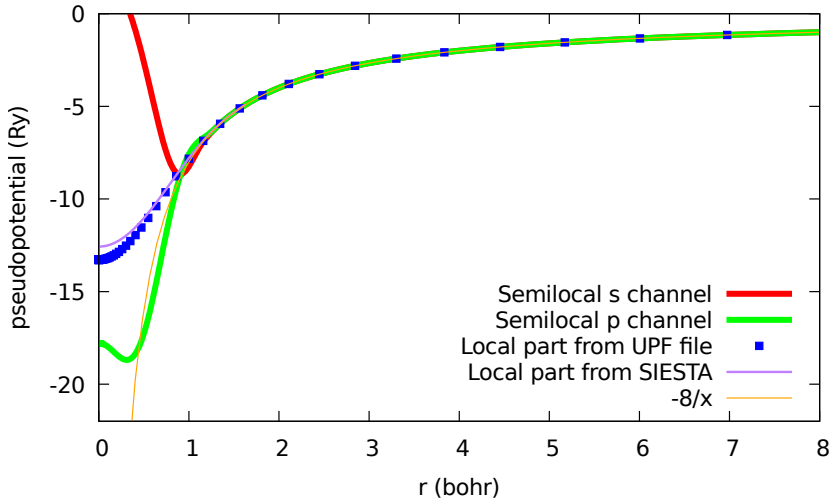


Figure 3.6: Optimized PP for C. The local part of the PP is reported in blue dots for the UPF, while the SIESTA choice is in purple. The semilocal part for s and p channel is also reported. The dependence -8/x is the expected Coulomb dependence in presence of 4 atoms in valence.

The check on the performances of the PSF PP has not been done for each element. We demonstrated that the accuracy depends also on the basis, but it is not the scope of our work to optimize the SIESTA basis for each crystal element under investigation. The information about a good basis for the crystal elements would not be transferable to other systems.

3.1.4 Pseudopotentials with self-interaction correction

Our final step consists in producing the same PP using the code ATOM, in which a modification has been added in order to create, alongside the standard PP, a self-interaction corrected PP (as explained in Section 2.7). It is possible to compare the charge, PP and wave-functions for the ATOM and APE outputs. A perfect match for every quantity (Figures 3.7, 3.8, 3.9) ensures us that the fact that we are producing the same PP.

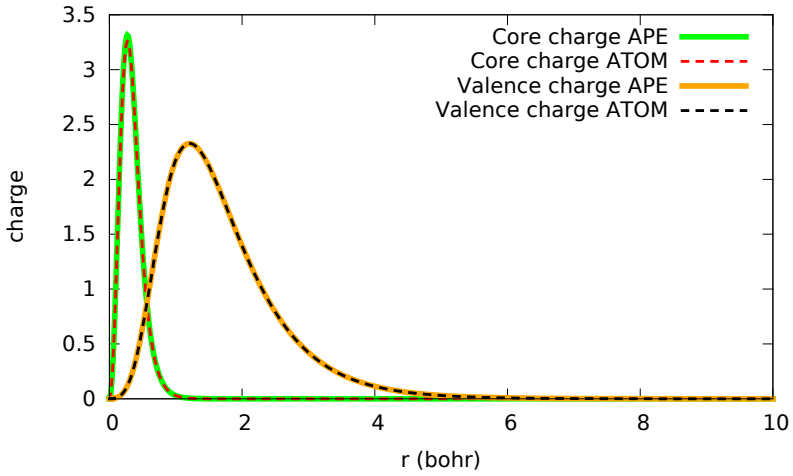


Figure 3.7: PP valence and core electronic charges multiplied by $4\pi r^2$ generated by ATOM and APE.

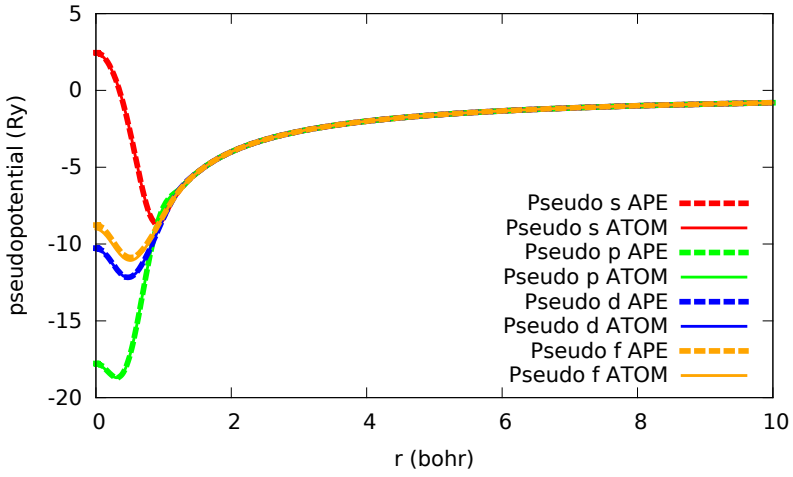


Figure 3.8: PPs comparison between ATOM and APE. Different colors refer to different angular momentum channels.

Also the equation of state is computed and results for graphite are in Table 3.7. The slight difference between the two Δ values is due to the different radial grid used by the two code for the generation of the PP. SIESTA adapts the cutoff of the basis to match one point present in the radial grid. As the grid is different between the two codes, the actual cutoff for the s and p basis elements are slightly different. Namely for APE they are (in Bohr) 4.088 and 4.750; for ATOM 4.048 and 4.800.

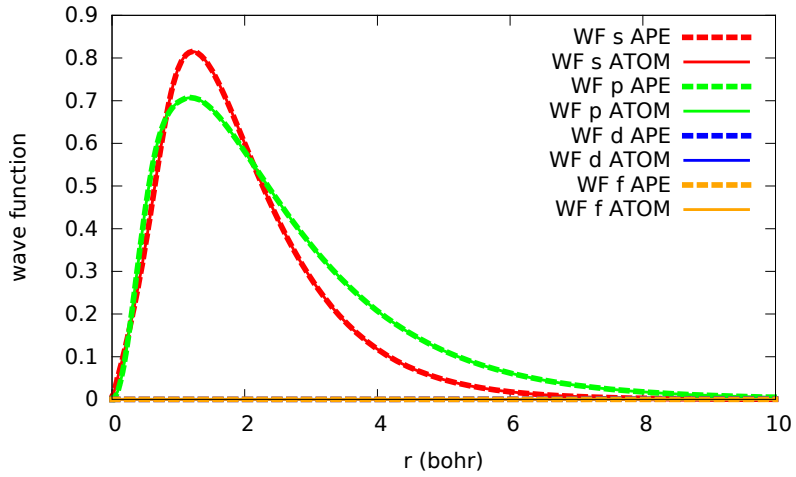


Figure 3.9: PP wave-functions including a factor of r. The f and d orbitals are unbound. Their wave function is zero close to the origin.

	V_0(Å^3/atm)	B_0(GPa)	B_1	B_0(eV/Å^3)	Δ (meV)
AIMS Reference	11.3427	220.5460	3.51255	1.376540	-
APE DZDP	11.4190	227.2668	3.64247	1.41849	3.83
ATOM DZDP	11.4227	226.8607	3.59085	1.415953	3.99

Table 3.7: Results from the equation of state of C (graphite) obtained with the PP generated by APE and by ATOM employing the same generation parameters. The cutoff radii are the one of the optimized PP.

An extra quantity we easily obtain during the PP generation with ATOM is the logarithmic derivative of the wave function as function of the energy for a radius longer than the radius chosen for the PP generation. The all-electron and PP results are compared for both the s and p orbitals and the perfect match is an indication of the good transferability of the PP.

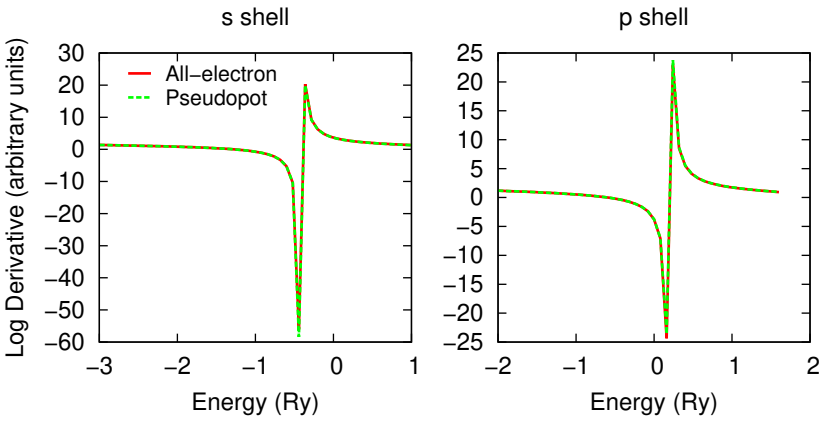


Figure 3.10: Logarithmic derivative of the wave function as function of the energy for a radius longer than the radius chosen for the PP generation.

We then succeed in creating a set of reliable PPs for elements between Li and Ge in PSF format. These PP can be used for high-throughput SIESTA calculations. The exact

accuracy can not be precisely discussed because of the problem of having a different local part respect to the same PP in UPS form. However, a PP deriving from the same semi-local form have been tested with QE giving an accuracy corresponding to a Delta value of less than 1 meV. We plan in the future to make publicly available the data and the PPs produced in this section, with the help of a web interface.

3.2 ASIC

The problem of calculating the band gap in DFT was discussed in Section 2.7, where the ASIC and DFT+U scheme were also presented. Using these gap correction methods in an high-throughput study is limited by the fact that both of them require the introduction of an empirical parameter (λ for ASIC and U for DFT+U) and there is no systematic way to select it.

In recent years we also saw the publication of many studies investigating the possibility of predicting the band gap using machine learning techniques [113, 114, 115, 116, 117]. For the sake of this work, it is sufficient to think about machine learning as a sophisticated fitting procedure: a set of variables (collectively called feature vector - $\mathbf{X}$) are associated to a target quantity y (either scalar or vectorial) and the machine learning model tries to find the relation $y{=}f(\mathbf{X})$. In some of the aforementioned studies, DFT calculated properties are used as a feature vector. For instance in reference [113, 114] the KS gap is used in combination with other chemical properties of the material.

In this section we have collected a series of experimental data of binary compound band gaps and we performed high-throughput DFT simulations with ASIC corrections in order to answer two main questions:

- In case one wants to predict the band gap starting from DFT data, is it better to consider standard DFT calculations or the ASIC-corrected gap is a better starting point? (subsection 3.2.2)

- Is it possible to obtain through machine learning a model that can predict the parameter λ to use in a DFT-ASIC simulation in order to calculate the correct experimental gap? In other words, we are looking for a model able to suggest *a-priori* the correct λ to use. (subsection 3.2.3)

A successful answer to the second question finds its applicability in all the situations when the energy gap is not the final quantity of interest, but its accurate description is necessary. An example are the simulations of tunneling transport that can be performed using Green's function formalism together with DFT [34]. Our work also somehow fits in a series of recent studies devoted to formalize a self-consistent U theory [118, 119, 120, 121, 122], where the U parameter of DFT+U is chosen automatically without user intervention.

Before presenting the results, a detailed description of the criterion we have used to collect the experimental data and of the workflow built in order to run several ASIC calculations is explained in the next subsection.

3.2.1 Experimental data collection and High-throughput workflow

We first collect experimental data for 257 binary compounds from various references [123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134]. In order to enter the list, a space group must accompany the experimental band gap. The density in g/cm^3 is also taken when present. Most of the times it is not so easy to understand what type of measurement was employed to obtain the gap (transport, optical, etc.). As we are interested in the semiconductor gaps, measurements deriving from electronic transport studies (temperature dependence of conductivity, hall coefficient and others) are privileged. However, for most of the compounds only optical measurements are available. We decide to use them as well, being careful to distinguish the excitonic properties from the actual gap when possible.

The next step consists in associating a crystal structure to every compound composition for which a space group and an experimental band gap have been collected. To this end, we have used the Inorganic Crystal Structure Database (ICSD) [135], a project aiming to gather structural information of all the synthesized inorganic compound published in literature since 1913. For each compound and space group, there are several corresponding structures. When a density was associated to the experimental gap, we looked for the entry of ICSD with a similar density. In the remaining cases we have selected a random one, expecting anyway all the entries to have comparable densities as they are coming from experimental data.

The association with a structure was possible only for 190 compounds. The complete list of such compounds is in Appendix C, where the ICSD number identifier is also reported.

A workflow is then generated to conduct an automatic inspection of some of the compounds in the list mentioned above. Only compounds containing elements with atomic number smaller than 38, plus In, Sn, Sb, Te and I are considered in the study. For heavier elements we expect spin-orbit effects to be important and it is impossible for the moment to perform calculations with both spin-orbit and ASIC. For each compound, the workflow prepares, sends and retrieves the outputs of a standard DFT calculation ($\lambda = 0$) and a fully corrected calculation ($\lambda = 1$). We have now to remind the significance of the λ parameter in the ASIC theory. It is a factor introduced to take into account the difference between the atomic like orbitals (used to calculate the self interaction correction) and the minimized SIC orbitals of the system. A value of lambda equal to one is expected to work well in the case of very ionic compounds, while it is smaller for systems described by more delocalized orbitals. A value $\lambda = 0.5$ was predicted in reference [93] to give good performances for a range of covalent compounds.

Let us conclude this section with a few technical notes. As the structure is taken from an experimental database, no relaxations of the atomic positions or cell optimization is performed. The exchange correlation potential is LDA-PZ, the SIESTA ultra converged parameters are a mesh cutoff 900 Ry and the k-point mesh is constructed by imposing 30 k-points in the direction with the shortest lattice vector and rescaled in the other two directions according to the ratio of the lattice vectors. The PPs are not the one obtained in Section 3.1 because this work was done in parallel to the extraction of the optimized PPs. We used instead the PPs available in the ICMAB database [136], that, as explicitly

written on the web site, are not tested. They are obtained translating into the PSF format a set of PPs in use for the DT code Abinit [137].

At the end was possible to obtain results for about 90 compounds only. These are presented in the next two subsections.

3.2.2 Gaps analysis

The first question we address concerns whether it is ASIC ($\lambda = 1$) or LDA that provides the best starting electronic structure to be use in the construction of a machine learning model to compute the band gap. Figure 3.11 reports the comparison between the experimental gap and the Kohn-Sham (KS) gap for both standard LDA (left panel) and $\lambda = 1$ ASIC-corrected band structures. Perfect agreement is obtained for points situated on the blue line.

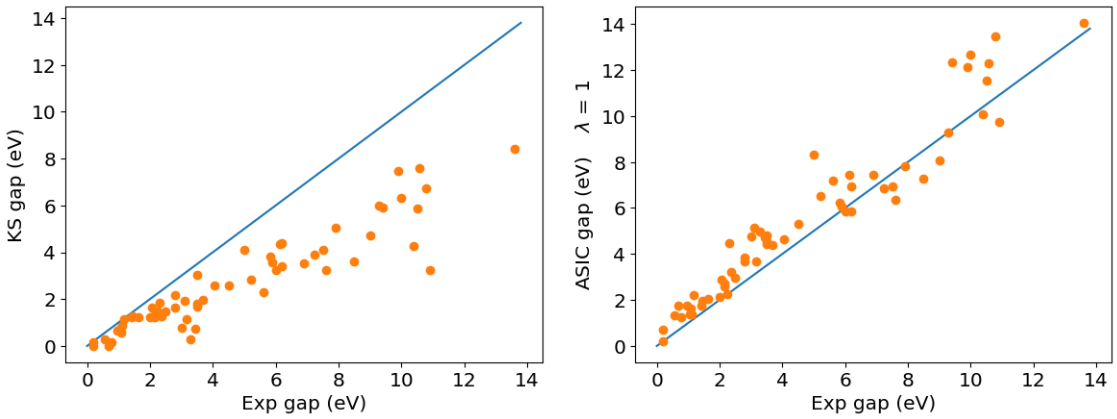


Figure 3.11: Comparison between the experimental gap and the Kohn-Sham (KS) gap for both standard LDA calculations (left panel) and $\lambda = 1$ ASIC-corrected situations (right panel). The blue line marks the perfect agreement between the calculated gap and the experimental one.

Focusing on the right panel, we notice that most of the points are above the blue line. This shows how the choice of $\lambda = 1$ over-corrects the gap of most of the compounds, as expected from the theory. The points falling below the line $y{=}x$ are, instead, the materials for which not even $\lambda = 1$ is enough to open the gap. Luckily, these are not too far from the line, meaning that this feature is probably due to inaccuracy of the calculations rather then to a systematic problem of the methodology. Moreover they are all ionic compounds, for which we know a value of $\lambda = 1$ should be optimal.

Going back to the original question, just by looking at Figure 3.11 it seems that the ASIC-corrected gap is closer to the real one with respect to the non-corrected case, even with this choice of $\lambda = 1$ that is, at the end, a choice of convenience looking at the totality of the compounds in consideration (λ has not been optimized here). In order to have a more quantitative result, we decide to apply a machine learning method to our data. The choice fell on the Random Forest method, known to perform better than other models when a few data are available for the training [138]. We used it in conjunction with 3-fold cross-validation, meaning that the set of our data is divided into three subsets and three different models are considered, each one of them takes two subsets as training set and one as a test

set. From the software side, we employed the function ExtraTreesRegressor of the scikit-learn python package [139] with n_estimators=100 and max_depth=4. Random forest consists of a large number of individual decision trees that try to predict y by splitting the data available into groups. The splitting decisions derive from the comparison of feature vectors $\mathbf{X}$. The individual decision trees are uncorrelated and each one of them makes a prediction. The individual predictions are then averaged together to get the final prediction. The quantity n_estimators is the number of decision trees that are considered and max_depth is the number of splits allowed for each decision tree. It is out of the scope of this project to detail all the technicalities of the Random Forest method, but the details can be found in reference [138].

Inspired by the idea of including some information about the ionicity of the bonds present in each compound, we decided to use a feature vector composed by two quantities only: the difference in Pauling electronegativity of the two elements composing the material and the DFT gap. Since we want to compare the ASIC calculation and the non corrected case ($\lambda = 0$), the DFT gap corresponds to the $\lambda = 1$ ASIC gap in one case and to the standard KS gap in the second.

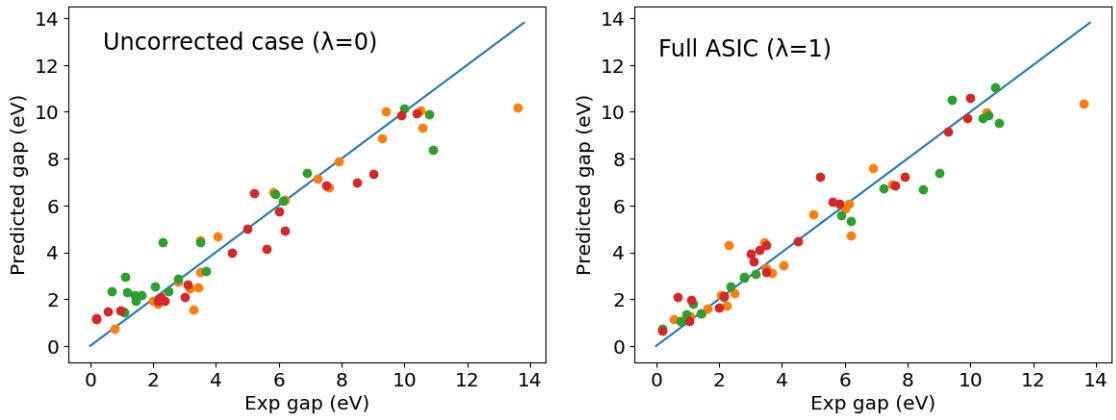


Figure 3.12: Machine learning prediction of the band gap using the following quantities as a feature vector: on the left panel the Pauling electronegativity difference and the uncorrected KS gap, on the right panel the ASIC ($\lambda = 1$) gap and again the Pauling electronegativity difference. The blue line correspond to the perfect prediction. Each colors of the dots correspond to one subset of the the 3-fold cross-validation procedure.

A pictorial representation of the fitting result is in Figure 3.12. For both panels, the x axis describes the experimental gap of the test set and on the y axis the machine learning predicted gap. The three cases corresponding to the 3-fold cross-validation procedure are all reported. They appear as points having three different colors in Figure 3.12. The prediction is surprisingly good in both cases. The inclusion of just two chemical features in the feature vector is enough to predict the experiment gap with good accuracy. A measure of accuracy is given by calculating the coefficient of determination, denoted R^2 [138]. An R^2 larger than 0.90 is obtained in both cases. The very good prediction ability of the model is due, in part, to the correlation between the experimental gap and the KS gap (see Figure 3.11), however it also means that the difference in Pauling electronegativity is quite a good descriptor for the data in our possession. This is shown in Figure 3.13, where

the difference in Pauling electronegativity is reported as a function of the experimental gap. An almost linear dependence between the two quantities is observed. Studies that try to connect the gap to the elecronegativity of the constituents of a material are present in literature [140].

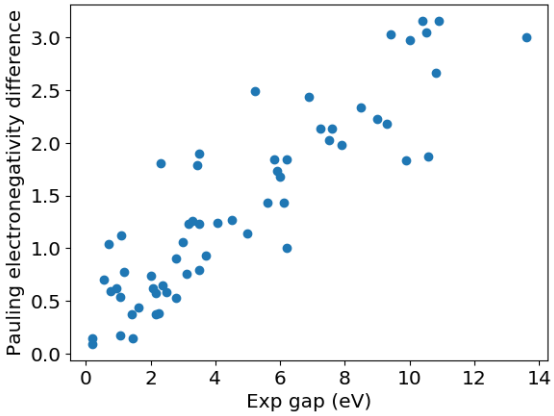


Figure 3.13: Pauling electronegativity difference versus experimental band gap.

Comparing now the LDA uncorrected situation and the ASIC $\lambda = 1$ case, it is found that the R^2 is around 0.90 for LDA and around 0.93 for ASIC. We conclude that the ASIC corrected case works slightly better than the uncorrected case. This is an important result, especially recalling that a correction $\lambda = 1$ is applied to all the materials. As already mentioned, this is a choice of convenience and it is expected to lead to the correct gap only for more ionic compounds. On the other hand, also the LDA uncorrected gap combined with the difference in Pauling electronegativity is a good descriptor to predict the experimental gap.

3.2.3 Selection of λ

The final part of this chapter discusses the idea of using the available DFT data to construct a model that is able to predict the correct λ to use in a ASIC calculation in order to obtain the correct band gap.

The first observation is that the band gap increases linearly with the λ parameter. This can be seen in Figure 3.14, where an example for GaP is reported. This observation allows us to obtain the λ giving the experimental gap, λ_{exp}, by simple linear interpolation, namely by using the equation:

$$\lambda_{exp} = \frac{E_{ex} - E_{KS}(\lambda = 0)}{E_{KS}(\lambda = 1) - E_{KS}(\lambda = 0)}, \tag{3.2}$$

where E_{ex} is the experimental gap and E_{KS} is the KS gap, calculated with the value of λ reported in the bracket. Therefore, λ_{exp} is the targeting quantity we now want to predict with a machine learning model.

We use again the Random Forest algorithm with various combinations of the following quantities in the feature vector:

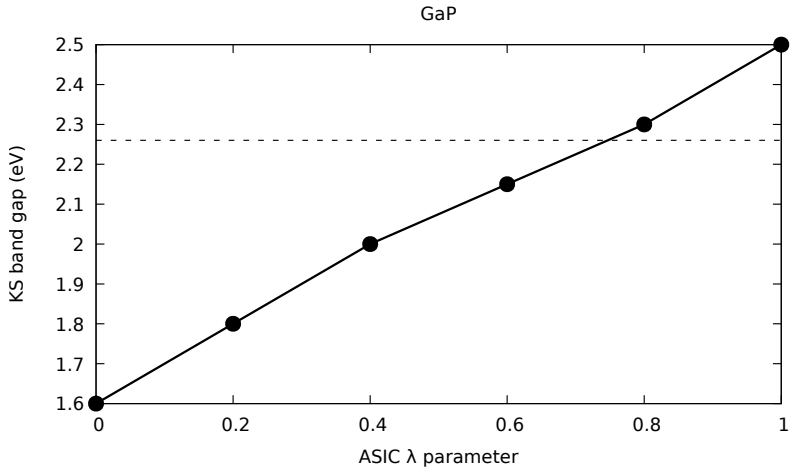


Figure 3.14: Dependence of the KS gap on the ASIC λ parameter. The dashed line represents the value of the experimental gap.

- Pauling electronegativity difference.
- Density of the material.
- Space group.
- Difference of the atomic charges extracted from the Mülliken population analysis [141] of the LDA calculation.
- KS gap for the uncorrected LDA calculation

The parameters for the Random Forest are the same used in the previous section, namely n_estimators=100 and max_depth=4. The results of R^2 are summarized in Table 3.8. None of the combinations we have tried could lead to a prediction of the λ_{exp}. The rationale for the choice of the feature vector is explained in the following. We first used the combination of Pauling electronegativity and LDA gap that so well performed in predicting the experimental gap (first line in Table 3.8). Secondly, we replace the Pauling electronegativity with a quantity containing similar chemical information, but deriving from DFT, namely the difference of the atomic charges calculated thanks to the Mülliken population analysis (second line in Table 3.8). Finally, some other information related to the crystal structure are added (third line in Table 3.8). For all the described combinations, R^2 is always smaller than 0.30. The same analysis is repeated using different values of the parameter max_depth, namely max_depth=2,6,8,10. No improvement is observed.

We discuss now the possible reasons of this failure by looking at Figure 3.15 that reports the λ_{exp} versus the difference in Pauling electronegativity, a quantity that so well correlates with the experimental band gap.

First we note that there are two compounds with very low Pauling electronegativity for which λ_{exp} almost one is necessary to reproduce the experimental gap. One of them is GaP. This compound is interesting because the result obtained with the automatic workflow can be compared with a previous calculation used to produce Figure 3.14. The main difference

Pau	Dens	SG	LDAgap	Mül	R^2
✓	–	–	✓	–	0.26
–	–	–	✓	✓	0.30
✓	✓	✓	✓	✓	0.25
✓	✓	✓	–	–	0.35

Table 3.8: R^2 values obtained from the Random Forest prediction of λ_{exp}. Each line correspond to a choice of the feature vector. "Pau" stands for Pauling electronegativity difference, "Dens" is the density of the material in g/cm^3, SG is the space group, "LDAgap" is the LDA KS gap, "Mül" stands for difference of the atomic charges extracted from the Mülliken population analysis.

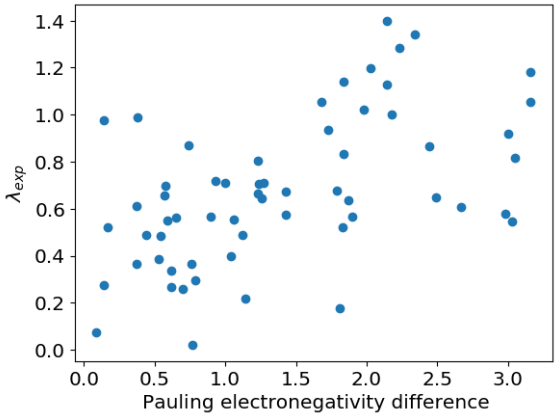


Figure 3.15: The quantity λ_{exp} as defined in Eq. 3.2 versus the Pauling electronegativity difference.

between the two calculations is the use of a different PP. It turns out that the result of the simulations obtained in the workflow leads to a gap of 1.475 eV for $\lambda = 0$ and a gap of 2.268 eV for $\lambda = 1$, fairly different with respect to the calculation presented in Figure 3.14. The difference for both corrected and uncorrected gap is about 10%. If we look at the quantity λ_{exp} though, in one case $\lambda_{exp} = 0.73$ and in the other $\lambda_{exp} = 0.99$, a 25% difference. This shows how sensible is λ_{exp} to the correct description of gap and, in turn, to the accuracy of the calculations.

A second point to make emerges by looking at the very bottom of Figure 3.15 where we have compounds for which a very little correction ($\lambda_{exp} < 0.1$) is enough to obtain the experimental gap. Among them there is the binary GeTe. This compound is known to be predicted a small gap semiconductor by DFT calculations. However, in any experimental sample the presence of Ge vacancies plays a fundamental role, turning the compound into a degenerate semiconductor [142]. In this case then, it is the experimental data not to be trusted, as it probably derives from the study of a sample with great amount of vacancies.

The two situations just presented are two examples of problematic cases that are present in our study. Our difficulty in finding a model that predicts λ_{exp} is then probably due to the inconsistency of the data in our possession. At this point, it is also not surprising that most of the works in the literature presenting the energy gap prediction through machine learning, uses GW calculated gaps as a target quantity rather then a set of experimental data. GW is a computationally very demanding way to obtain the semiconductor gap of a material and has been proven to be very accurate when compared to experimental

data. GW calculations allow to factor out the variability due to the experimental error, making the machine learning more reliable. In our study, the bad propagation of errors acting on λ_{exp} and explained for the case of GaP, adds even more complexity to the prediction. Finally, it is important to remind, at this point of the discussion, that the self-interaction error is just one aspect of the wrong prediction of the band gap employing DFT. For instance, in very covalent compounds, correlation errors are expected to be more relevant, making difficult to reach a satisfying correction with ASIC. Elemental crystals were excluded from the study, but other compounds with strong covalent bonding (such as silicon carbide) are in the list of the analyzed materials. This underline difficulty of ASIC includes a further level of complexity that the machine learning model needs to predict.

We conclude that the attempt to predict λ_{exp} is, so far, unsuccessful, however a couple of ideas might improve the situation. Firstly, we plan to perform again the ASIC calculations with more reliable choices of the input parameters, namely a richest basis set and more reliable PPs. Secondly, we will extract more information from the DFT calculations, for instance the so call "off-diagonal terms" of the Mülliken population analysis, representing the overlaps between the charges on two atoms. Also structural information might be included, for example the coordination of each element in the compound and the bond distances. Finally, we are looking to expand the set of experimental data in order to include more binary compounds.

Chapter 4

Ferromagnet selection: bands symmetries

This chapter explains the first high-throughput study carried out in this thesis in order to gather some insightful information for the improvement of the Spin Transfer-Torque Magnetic Random Access Memory (STT-MRAM) technology. The role of the ferromagnet in the memory was extensively discussed in Chapter 1 and here we want, first of all, to summarize some essential requirements that a bulk material needs in order to be considered a candidate for the technology:

- The most important feature is that the material stays ferromagnetic at the operational temperature of the memory, estimated to reach up to 600 K. This means that a Curie temperature, T_c, higher than 600 K is required. It is interesting to notice that high T_c means also strong exchange coupling between atoms and in Section 1.4 we have explained how this helps to rotate the spin moments uniformly, favoring the overall magnetization switching. High T_c is then never harmful. Although in actual devices the ferromagnet is in a thin film form, the bulk T_c is a good descriptor for the ferromagnetic robustness of the material as it was demonstrated that even down to 4-5 monolayers the T_c of thin films is similar to the bulk one [143, 144].

- The ferromagnet should be magnetically soft enough to allow magnetization reversal in presence of a relatively low current, but hard enough to be thermally stable. Section 1.4 went in great details on the competition between low switching current and stability in small dimensions. The most important quantity to consider here is the anisotropy and it was underlined how interfacial effects play a relevant role. Therefore, in this respect, very little one can say to help the selection of good materials just by looking at bulk properties.

- In Section 1.6 we have also discussed the relevance of the Gilbert damping parameter, that should be as small as possible. A systematic analysis of this quantity is, however, not accessible at the moment. The interplay of the different contributions critically complicates the estimation of the bulk value of the Gilbert damping for a given material. Even theories limiting the discussion to the so called "intrinsic" damping, (generic

dissipation induced by the presence of spin-orbit interaction) lead to computationally expensive strategies to evaluate the Gilbert damping [40], certantly not feasible for an high-throughput study.

- A final feature is the importance of the ferromagnet bands and their symmetry. Their study is necessary to predict the presence of the symmetry filtering effect, a vital element to achieve high Tunnel-Magneto-Resistance (TMR) ratios (see Section 1.2). Moreover, the ferromagnet bands are useful to infer the presence of spatial oscillations of the torque, as explained in Section 1.3

With the present list in mind, we have adopted a strategy to create a database of useful properties for the STT-MRAMs. The first point in the list (high T_c) is used to pre-select a pool of interesting ferromagnets. One could say that this is the quickest and easiest way to discard materials that for sure are not compatible with the technology. The Curie temperature is, in fact, a widely studied quantity in the magnetic materials community. The selection is based on experimental data and is explained in Section 4.1.

On the preselected materials, an accurate Density Functional Theory (DFT) calculation is performed in order to obtain the bands dispersion along high symmetry direction of the Brillouin zone. Further analysis allows us to understand the symmetry of all the bands crossing the Fermi energy for spin-up and spin-down electrons, a valuable information for the design of STT-MRAMs as expressed in the last point of our list. Our procedure is detailed in Section 4.2, while the discussion of the results is in Section 4.3.

4.1 Preselection based on the Curie temperature

The pre-selection of ferromagnetic materials on which to perform band structure calculations is based on the simple idea that only materials showing Curie temperature higher than 600 K should be consider for the study. The T_c is a very important quantity for magnetic applications and it is available experimentally for a wide range of crystals. In particular we used the data presented in the work of Nelson and Sanvito [145] that were able to collect the T_c values for about 2500 ferromagnrtic materials.

In contrast to what we have done in Section 3.2, in reference [145] only the experimental T_c and the composition of the material are collected, without any information about the crystal structure, the space group or the density. This makes the association of the T_c with an experimental structure, on which to perform further analysis, more problematic. The final choice was to select, for each compound, the ICSD structure [135] showing the smallest enthalpy of formation per atom calculated at the DFT level. The latter information is available in the AFLOW database [146].

The AFLOW project contains DFT+U calculations of all the entry of the ICSD database. Among other quantities, the enthalpy of formation per atom is calculated simply as the difference in total energy of the material and the crystal elements composing it [147]. Taking for instance a ternary compound $A_iB_jC_l$ with $i+j+l=1$ and naming the energy per atom ε, the enthalpy of formation is $\varepsilon(A_iB_jC_l) - i\varepsilon(A) - j\varepsilon(B) - l\varepsilon(C)$, where the single

ICSD number	T_c(K)	SG
Al1Co2Mn1_ICSD_57618	693	225
Al1Cu2Mn1_ICSD_57695	603	225
Al1Fe1_ICSD_165165	640	221
Al1Mn1_ICSD_608471	650	123
Al1Mn2V1_ICSD_181805	730	225
B1Co3_ICSD_612882	747	62
B1Fe2_ICSD_391330	1013	140
B1Fe3_ICSD_260757	897	62
B2Fe5P1_ICSD_614140	628	140
B2Fe5Si1_ICSD_614183	784	140
C1Co2Mn2_ICSD_44353	808	123
Co1Fe1_ICSD_187982	1253	221
Co1Fe2O4_ICSD_166200	790	227
Co1_ICSD_53806	1387	194
Co2Fe1Si1_ICSD_52958	980	225
Co2Ga1Mn1_ICSD_185084	694	225
Co2Ge1Mn1_ICSD_185088	905	225
Co2Mn1Si1_ICSD_624141	985	225
Co2Mn1Sn1_ICSD_102532	829	225
Cr1Fe3_ICSD_188258	1083	225

ICSD number	T_c(K)	SG
Fe1_ICSD_76747	1041	229
Fe1Ni1_ICSD_103555	790	123
Fe1Ni3_ICSD_40334	941	221
Fe2Mg1O4_ICSD_24493	653	227
Fe2Ni1O4_ICSD_52387	859	227
Fe2O3_ICSD_201097	961	167
Fe3Ga1_ICSD_631740	760	194
Fe3Ge1_ICSD_632004	647	194
Fe3P1_ICSD_43365	716	82
Fe3Si1_ICSD_157941	810	225
Fe3Sn1_ICSD_108474	743	194
Fe3Sn2_ICSD_71	612	166
Fe4N1_ICSD_53502	761	221
Ge2Mn5_ICSD_40462	710	72
Mn1Ni1_ICSD_104917	662	221
Mn1Ni1Si1_ICSD_165241	622	62
Mn1Ni3_ICSD_104920	750	221
Mn4N1_ICSD_76055	745	221
Ni1_ICSD_76668	628	194
Ni2Sc1_ICSD_105335	630	227

Table 4.1: The 40 compound with T_c larger than 600 K considered in this study. The first column reports, together with the material composition, the ICSD reference number. The second column is the experimental T_c collected in reference [145] and the last column is the space group (SG) of the associated structure.

element energies refer to crystal element calculations. Of course the same level of theory is used for the compound and the crystal element calculations.

In conclusion, for a given T_c, anytime multiple structures are associated to it (different space groups or just density), the corresponding AFLOW calculations are checked to select the structure with the lower enthalpy of formation. The method, of course, does not guarantee to pair the experimental T_c to the correct structure on which the experiment was performed, however it offers a systematic way to select the most stable structure that likely was the object of the experiment. We plan in the future to go through the database once again and recover the missing information directly from the original reference experimental work from which the data in reference [145] are collected.

The 2500 experimental data set contains also alloys to which no crystal structure could be associated. At the end of the analysis, 792 materials were matched with an ICSD entry. Among them, only 82 showed a T_c larger than 600 K, proving the fact that the ferromagnet choice is already pretty limited. We also decided to exclude, for the moment, compounds containing very rare metals (Au, Pt, Os, Ir) and f elements as they would require the inclusion of spin-orbit effect for a correct description of the electronic properties. Our plan is to extend the study to these compounds in the future.

The procedure described leaves us with 40 compounds in total, listed in Table 4.1, where also the space group of the associated structure is presented in the last column. Cubic space groups are the most present, with predominance of the rock-salt structure (space group number 225). We will comment further on this point in Section 4.3.

4.2 Bands and symmetries

This section aims to discuss in details the method used to calculate the bands structure of the 40 ferromagnets listed in Table 4.1 and especially the procedure to understand the symmetry of the bands close to the Fermi energy.

The DFT calculations are performed with the code SIESTA, using the Pseudo-Potential (PP) obtained in Section 3.1. Only for the element Sn an optimized PP was not available and therefore we used the one available in reference [136]. The SIESTA basis set DZP is employed. An ultra-converged value of 900 Ry is used for the mesh cutoff parameter and also for the k-point mesh we applied very stringent conditions: the standard mesh is 35×35×35, rescaled in one or two direction in case of non isotropic elementary cell. Spin-polarized calculations are performed in the collinear-spin approximation leading to spin dependent band structures.

The DFT calculation in itself does not present particular concerns, the part of the workflow that requires more careful treatment is the understanding of the bands symmetries. In Section 2.6, the concepts leading to the significance of the symmetry of a band has been discussed. Here we want to explain how to practically apply those concepts to assign a symmetry label to every band passing in the vicinity of the Fermi energy. Following the theory of Section 2.6, one needs, first of all, to identify the group of the wave vector for the direction in reciprocal space under investigation (the possible transport direction). Secondly, all the symmetry operations of the group of the wave vector should be applied to the wave function of the band in order to understand according to which representation the band is invariant. The incorporation of this procedure in a automatic workflow implies many difficulties:

- Given the space group of the crystal structure, the group of the wave vector for every high-symmetry wave vector **k** is tabulated in books or can be obtained from the Bilbao Crystallographic Database under the section "The k-vector types and Brillouin zones of Space Groups" [79]. The website offers a nice interface where one can type the crystal space group and the **k** value (referred to the primitive or conventional cell) in order to obtain all the needed information. In particular, the group of the wave vector is reported under the name "Little co-group". In order to include this step in an automatic workflow, one requires either a complicated interface to call the Bilbao Crystallographic Database web-page or to reimplement from scratch the group of the wave vector finder (the scripts of the Bilbao Crystallographic Database project are not publicly available).

- The group of the wave vector for a high-symmetry **k** is a point group (acually the statement is true only for "symmomorfic" space groups, but we do not go into more complicated situations). There are 32 of such point groups in total and implementing all the transformations corresponding to all the symmetry operations is a challenge in itself. Moreover, determining the correct representation essentially requires to encode the character table for each point group.

- Assuming to have a way to select the correct representation according to which the wave

function transforms, there is the final problem to express the finding in an understandable way. One needs to relate back the information obtained from the character table of the point group to the terminology of the original group of the wave vector.

For the many reasons described above, the idea to have a completely automatic procedure must for the moment abandoned. A workflow can, however, be created to gather in a database all the information one needs. The symmetries will be then assigned using, for example, the tools of the Bilbao Crystallographic Database described before.

Our final choice is to systematically produce data on two important quantities (apart from the spin-dependent band dispersion). Firstly, for every high-symmetry direction, we list the values of **k** where the bands cross the Fermi energy (there might be more than one). Moreover we check if two bands are degenerate at each one of these k-points. This is done for both the spin-up and spin-down bands. Secondly, we calculate, for each listed **k**, the projection of the wave functions in the plane perpendicular to the direction under investigation. Only the wave functions of bands in the vicinity of the Fermi energy are considered.

The procedure is clarified here with an example, the first entry in Table 4.1, namely the compound Al1Co2Mn1_ICSD_57618. Figure 4.1 shows the spin-dependent band dispersion of the material, calculated with SIESTA on the experimental structure taken from the ICSD database. Note that the high-symmetry path is automatically chosen by AiiDA thanks to the SeeK-path package [148].

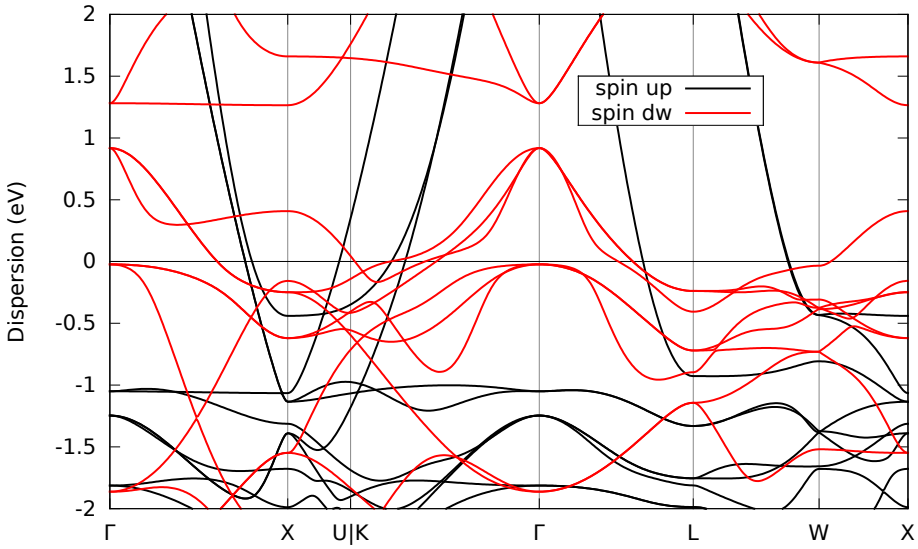


Figure 4.1: Spin-dependent bands dispersion of Al1Co2Mn1, structure from ICSD entry with identifier ICSD_57618. Spin-up bands are in black, spin down in red.

Let us now focus on the direction Γ−X. As already said, we are interested in listing the k-points at which a band crosses the Fermi energy. From Figure 4.1, we see that there are three of such bands, two spin-up bands and one spin-down. An algorithm implemented in AiiDA as a post process tool of the band structure data is able to store in a database information of the kind reported in Table 4.2. In our example, the algorithm identifies two

[a] direction Γ−X, spin up, **k** index 1

cartesian	[0.0, 0.82434, 0.0]
reciprocal lattice vectors	[0.37850, 0.0, 0.37850]
energy multiplicity	2

[b] direction Γ−X, spin up, **k** index 2

cartesian	[0.0, 0.86505, 0.0]
reciprocal lattice vectors	[0.39719, 0.0, 0.39719]
energy multiplicity	1

[c] direction Γ−X, spin down, **k** index 1

cartesian	[0.0, 0.66151, 0.0]
reciprocal lattice vectors	[0.30374, 0.0, 0.30374]
energy multiplicity	2

Table 4.2: Information about the band crossing the Fermi energy in direction Γ−X for the compound Al1Co2Mn1. Each subtable corresponds to one k-point at which a band crosses the Fermi energy. The value of each k-point is reported in cartesian (1/Bohr) and reciprocal lattice vectors units. The multiplicity of the energy at each k-point is also reported.

bands crossing the Fermi energy for spin-up electrons (subtables [a] and [b] of Table 4.2). Information about the corresponding k-points are listed in both cartesian coordinates and fractional coordinates respect to the reciprocal lattice vectors. The value appearing in the lines "energy multiplicity" is the number of energies that are present for each k-point. For one **k** (subtable [a] in Table 4.2) the multiplicity is 2. This means that the band crossing the Fermi energy at that k-point is double degenerate. In the following, it will be clear how this is already a valuable information. The second spin-up **k** shows instead one single energy (subtable [b]). The last subtable in Table 4.2 refers to the spin-down bands. The down bands cross the Fermi energy only at one **k** point between Γ and X. Two degenerate energies are present. By looking at Figure 4.1 it can be noticed that two more down bands are very close to the Fermi energy in the vicinity of Γ. However, for the moment, the algorithm focuses only on bands that effectively cross the Fermi level, so that these two bands are not considered. This should probably be modified in the future.

The second quantity that is automatically produced by our workflow is a projection of wave functions of interest over the plane perpendicular to the direction under investigation. In our example the direction Γ-X is under scrutiny. We then want to project on a plane perpendicular to Γ−X the wave functions of the five energies identified in Table 4.2. They are grouped by k-vector in Figure 4.2, Figure 4.3 and Figure 4.4, meaning that the figures with two panels report the wave function of the two degenerate energies belonging to the same **k**. The projections are obtained thanks to a modified version of the code "denchar", belonging to the SIESTA distribution. All the inputs are basically listed in Table 4.2. Starting from the converged density matrix of the SIESTA calculation, it is possible to run again SIESTA (non self-consistently) to obtain the wave function of selected bands at selected k-points. The latter are then passed to "denchar", together with the direction perpendicular to the projection plane (Γ−X direction in cartesian coordinates in our example) and the data plotted in Figure 4.2, Figure 4.3 and Figure 4.4 are produced.

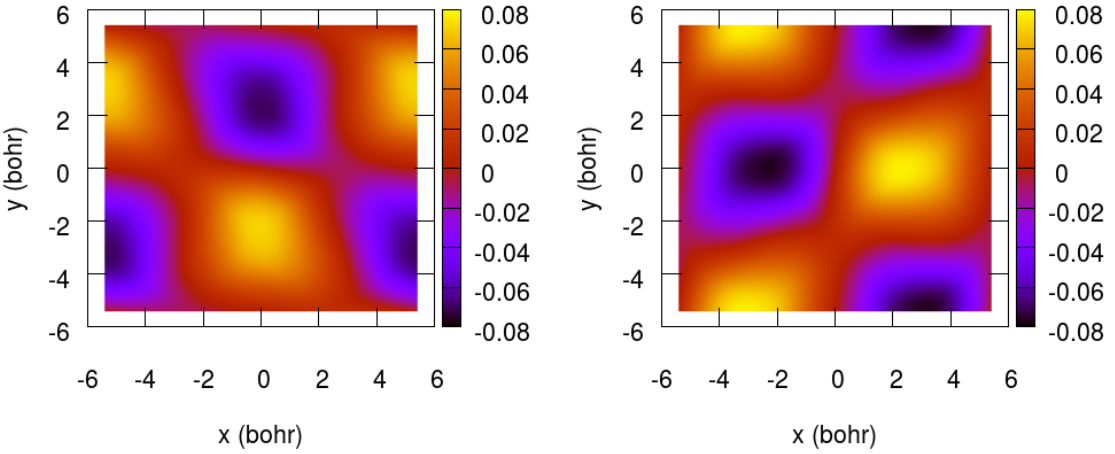


Figure 4.2: Projection on the plane perpendicular to the direction Γ-X of the real part of the wave function for the spin up band crossing the Fermi energy at $\mathbf{k}$ = (0.0, 0.82434, 0.0) 1/Bohr. Two energies are degenerate at this $\mathbf{k}$.

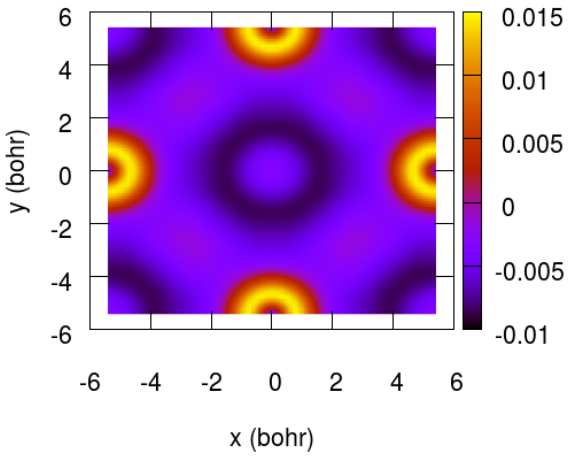


Figure 4.3: Projection on the plane perpendicular to the direction Γ-X of the real part of the wave function for the spin up band crossing the Fermi energy at $\mathbf{k}$ = (0.0, 0.86505, 0.0) 1/Bohr.

Our automatic procedure stops here. It is performed for every direction and every material in the list in Table 4.1. General considerations on the results will follow in Section 4.3. However, there is still to explain why we have decided to gather the described information about the projection of the wave function. The reason is the fact that, in easy cases, one can identify at sight the symmetry of the band by just looking at its projected wave function. For instance, for any direction, wave functions presenting spherical symmetry (like the one in Figure 4.3) always transform according to the irreducible representation that brings the function in itself for all the symmetry operations (the one with all entries equal to one in the character table, see Section 2.6). Moreover the presence of

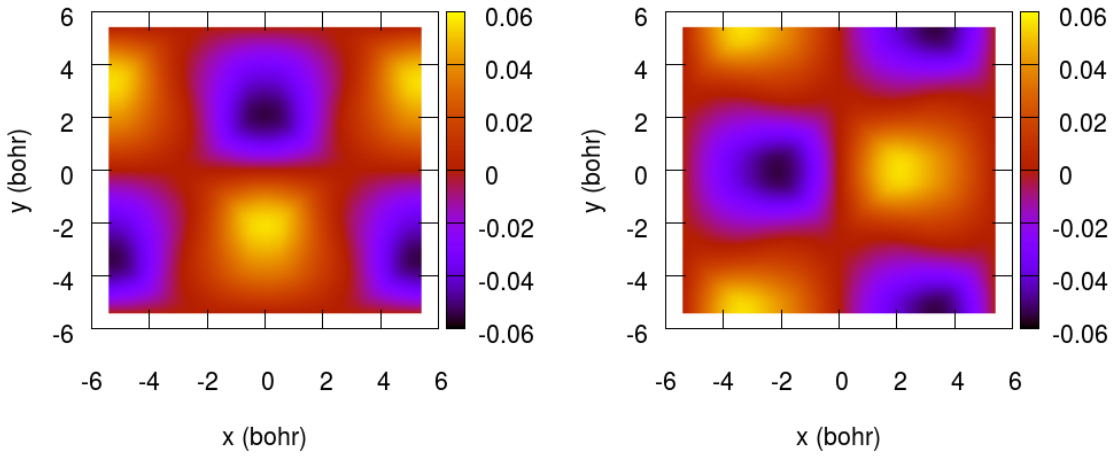


Figure 4.4: Projection on the plane perpendicular to the direction Γ-X of the real part of the wave function for the spin down band crossing the Fermi energy at $\mathbf{k}$ = (0.0, 0.66151, 0.0) 1/Bohr. Two energies are degenerate at that $\mathbf{k}$.

degenerate energies is already an indication on the symmetry as well. We did not prove it extensively, but in presence of degenerate energies, the corresponding wave functions must transform according to a representation with dimension larger than one. In the example of Al1Co2Mn1 along Γ−X, the group of the wave vector is the point group C_{4v}, whose character table is reported in Figure 2.1 of Section 2.6. The Δ_5 representation is the only one of dimension two and therefore we can conclude straight away that for the first (up) and last (down) $\mathbf{k}$ listed in Table 4.2 the symmetry of the band is Δ_5. Not all the cases are as simple as the one just described. However the information contained in Table 4.2, together with the space group of the structure, can be used in conjunction with the Bilbao Crystallographic Database in order to work out the correct symmetry labeling for the bands crossing the Fermi energy.

In conclusion to this section, a final note is necessary. We always refereed to the projection of the wave function along a symmetry direction. However, in general, the wave function is a complex quantity. The Figures 4.2, 4.3, and 4.4 report, actually, only the real part of the wave function. This is sufficient for any situation when no degeneracy of the energy is present. In that case, in fact, real and imaginary part of the wave function transform in the exact same way. The situation is slightly more complicated when two or more energies are degenerate. In fact, any linear combination of the two wave functions is then an admissible eigenvector. We are not guaranteed that the real and imaginary part of one single wave function transform in the same way, but a combination of the two wave functions with this characteristic must exist. Luckily, in the particular case of Figure 4.2 and Figure 4.4, the real and imaginary part transform in like manner even if two energies are degenerate, but this must not be considered the general case.

4.3 Results discussion

In this final section, we attempt a discussion about some interesting conclusion that can be derived looking at data collected for the 40 compounds in Table 4.1.

We first notice that many of the compounds are cubic (space group ≥ 195). In addition to Body Centered Cubic (BCC) Fe, the space group 221 is heavily present. The primacy goes however to the rock salt structure (space group 225) that embeds many ternaries compound. They are Heusler alloys [149]. Finally we see some compounds with space group 227. For all these systems the group of the wave vector along the direction Γ-X (equivalent to Γ-H of BCC Fe) is the point group C_{4v}. A direct comparison with BCC Fe can then be discussed.

To conclude something general regarding the band symmetries, however, is not easy. We take the approach of looking for same peculiar characteristics of interest for the STT-MRAM technology. In particular we look at situations where bands are present at the Fermi energy in the Γ-X/H direction only for one spin channel. This feature is showed by the following four compounds: Al1Fe1_ICSD_165165, Fe4N1_ICSD_53502, Mn1Ni1_ICSD_104917, and Mn4N1_ICSD_76055. Only Mn1Ni1 turns out to be a full half metal, with however a very small gap for the minority spins (of about 0.25 eV). For the others the absence of one spin stands only for the Γ−X direction. The symmetry analysis does not show any particular trend. Al1Fe1 has Δ_5 bands at the Fermi energy, Mn1Ni1 a Δ_1 band, Mn4N1 a $\Delta_{2'}$ band, while Fe4N1 shows three bands at the Fermi energy, the double degenerate Δ_5 and a Δ_2.

The described case where only one spin has states at the Fermi energy (in one particular direction) is of interest because it is supposedly the situation where the highest values of TMR ratio can be achieved. In fact, the conduction of one channel will always be suppressed. The symmetry matching between the states of the barrier and the states of the ferromagnet determines now only the transmission of the majority channel. In the case of a barrier selecting Δ_1 states for example, Mn1Ni1 will show a much higher current compared to the other three compounds.

The same exercise of looking for situations where bands are present at the Fermi energy only for one spin has been performed also for the L direction of the 225 space group lattices. The group of the wave vector along Γ−L corresponds to the point group C_{3v} (symmetries of a triangle). Two compounds allow only one spin at the Fermi energy in this direction. They are Cr1Fe3_ICSD_188258 and Co2Fe1Si1_ICSD_52958. While the first has two bands (Λ_3 and Λ_2), the second has only one (Λ_2). Similar discussions can be explored in all the high symmetry directions of cubic crystals.

Apart from the cubic lattices, another recurrent space group in Table 4.1 is the Hexagonal 194. The most interesting direction for this crystal is the so called A direction, perpendicular to base of the hexagonal prism that form the conventional cell of the lattice. In this situation, however, we do not register any situation like the one described for cubic lattices where one single spin has bands at the Fermi energy.

The discussions outlined so far are just examples of the many speculations one can carry out looking at the data collected in this work. At the moment, we are looking for

effective ways to make these information available to the public as we think it would be of great benefit for any research where the symmetry of the bands at the Fermi energy is important.

The small discussion presented here, however, was sufficient to identify Mn1Ni1 as a promising candidate to be matched with a barrier selecting the Δ_1 symmetry in order to have a high TMR ratios. This situation is interesting because insulators selecting the Δ_1 symmetry are quite abundant in nature. All the insulators with cubic structure and presenting a predominant s character of the valence or conduction band tend to select the Δ_1 symmetry for transport along the X/H direction of the crystal. The fact that Mn1Ni1 is an half-metal adds even more interest to the situation. Half-metals are promising materials for many applications, including high TMR ratio devices, however half-metallicity often is not very robust property. The presence of defects modifies the electronic properties of half-metals, making them less performing in real devices compared to the theoretical prediction. The fact that Mn1Ni1 presents Δ_1 symmetries of the band at the Fermi energy, should make more robust the half-metallic property of an hypotentical junction containing this material.

Chapter 5

Insulator selection: the Complex Band Structure

The role of the insulator in a Spin Transfer-Torque Magnetic Random Access Memory (STT-MRAM) junction is essentially that to create the conditions for an efficient symmetry filtering. This may lead to a fully spin-polarized current in the case of parallel configuration of the ferromagnets and to a much lower current in the anti-parallel configuration. Moreover the overall resistance of the junction depends on the insulator's characteristics. The required features for an ideal insulator in the STT junction are:

- Large enough band gap, to have tunneling transport in the presence of a bias.
- Clear symmetry filtering. The selection of one single symmetry is desirable.
- Low damping coefficient for the most transmitted symmetry to have a small overall resistance.

All these information are intimately connected to the tunneling properties of the junction and their quantification can be achieved by looking at the Complex Band Structure (CBS) of the insulator calculated at the Density Functional Theory (DFT) level. This is the topic of the entire chapter.

The importance of the CBS for the STT-MRAM technology, introduced already in Section 1.2, is recalled in Section 5.1 and 5.4, In the rest of the chapter the focus moves to the study of the CBS itself. After a rigorous definition of what we mean with the term CBS, a method to calculate it is presented: the Transfer Matrix Method (TMM). This is not the only way to obtain the CBS of a material, however it is the one that seems more suited for an high-throughput treatment and to investigate the symmetries of complex bands. The method is implemented as a post processing tool of the SIESTA code, but the first results showed the presence of what we call "flat bands". These bands are introduced by the presence of a non-orthogonal basis set. This statement is formally proved leading to a critic review of all the analytical properties of the CBS. A practical solution to deal with the flat bands is suggested. Finally, calculations of the CBS on few materials are presented in Section 5.4, following a discussion about their symmetries.

5.1 Complex band structure definition

The central quantity in the study of the electronic structure of a crystal is the dispersion relation, $E(\mathbf{k})$, obtained as a solution of the time-independent Schrödinger equation for the system under investigation (for instance the set of KS equations (2.10)). When E and $\mathbf{k}$ are real quantities, the relation between them is also called band structure of the material. For every energy E, the allowed wave-vectors $\mathbf{k}$ (if any) describe propagating states in the crystal, whose mathematical expression for the wave function, according to the Bloch theorem (Section 2.4), is:

$$\varphi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}) \tag{5.1}$$

where $u_{\mathbf{k}}(\mathbf{r})$ is a periodic function with the same periodicity of the crystal.

The wave vector $\mathbf{k}$ needs to be real in order to ensure finite wave-functions over the entire space, but this is not a requirement of the Schrödinger equation itself. It is then possible to generalize the dispersion relation, $E(\mathbf{k})$, considering complex wave-vectors $\mathbf{k}$. Wave functions of the type described by Eq. (5.1) with a complex $\mathbf{k}$ vector represent evanescent states that grow or decay from one unit cell to the next. Even though these latter states are forbidden by translational symmetry, they become important when such symmetry is broken via, for example, a surface or an interface.

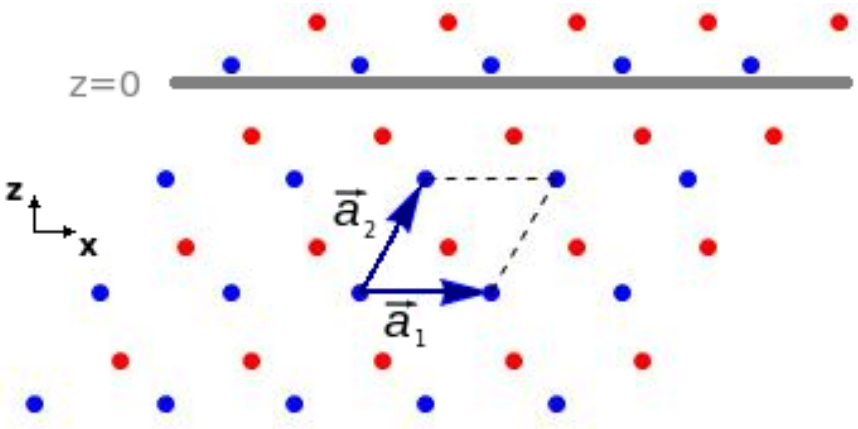


Figure 5.1: Pictorial representation of a 2D crystal. the line z=0 defines a line at which the translational symmetry is broken. $\vec{a}_1$ and $\vec{a}_2$ are the lattice vectors of the 2D crystal.

Take, for instance, Figure 5.1 as an example, and imagine that the plane $z = 0$ breaks the translational symmetry along z. Then one can consider the two subspaces $z > 0$ and $z < 0$ independently. A wave functions of the form:

$$\varphi_{\mathbf{k}}(\mathbf{r}) = e^{i(k_x x + k_y y)} u_{\mathbf{k}}(\mathbf{r}) e^{i\mathfrak{Re}(k_z)z} e^{-\mathfrak{Im}(k_z)z} \tag{5.2}$$

with $\mathfrak{Im}(k_z) > 0$ is perfectly acceptable for $z > 0$ as it does not diverge in that subspace.

The study of evanescent functions finds its applicability in many different fields, most notably in surface science and electron transport problems. The relevance of the decaying

functions in the study of surface states goes back to the theory of Heine, who postulated how the total wave function in presence of an interface can be obtained by matching the wave function on the right and on the left of the discontinuity when the evanescent function are included in the treatment [150]. More relevant to our study is the role played by evanescent states in tunneling transport. Take for instance the simple model of the one dimensional potential barrier of length a. It is possible to demonstrate how the transmission probability is proportional to $\exp(-2\mathfrak{Im}(k_z)a)$ (see proof in Appendix A). This means that an electron passing through an insulating barrier (like in the case of STT-MRAMs) has its wave function exponentially suppressed with a decay coefficient proportional to the imaginary part of k_z (k_z is the component of the wave vector along the direction of transport).

If one allows $\mathbf{k}$ to be complex, $E(\mathbf{k})$ will become a multi-valued $\mathbb{C} \mapsto \mathbb{C}$ function that can be studied with the tools of complex analysis, giving an entire new prospective to the understanding of the electronic structure of a crystal. The analytical properties of $E(\mathbf{k})$ were fist derived by Kohn [151] for strictly 1D situations and then extended by Heine [150] and Prodan [152] to include 2D/3D geometries when $\mathbf{k}$ is allowed to be complex in only one direction. This is the most interesting situation, as it includes effectively one dimensional materials, but also 2D/3D materials with a surface, where the components of the vector $\mathbf{k}$ parallel to the surface are conserved and can be treated as parameters.

Even though the study of the complex quantity $E(\mathbf{k})$ is of interest in itself, only energies with $\mathfrak{Im}(E) = 0$ hold a physical meaning. The collection of real energies from complex $\mathbf{k}$ is what is generally called the Complex Band Structure (CBS) of a material. Having information on the CBS of a material is equivalent to know, for each energy, the allowed complex $\mathbf{k}$ vector, meaning the decay coefficient of the corresponding evanescent wave-function. Here it resides the importance in studying this quantity. Heine showed, in a pioneering paper [150], that the CBS can be seen as a collection of "line of real energies". These are lines formed by energies with $\mathfrak{Im}(E) = 0$ that can be followed continuously from $E = -\infty$ to $E = +\infty$ and that have a well defined behavior in the complex plane. Properties of the "line of real energies" and of the complex dispersion relation $E(\mathbf{k})$ are the focus of Section 5.3.

Concerning actual calculation of CBS for real materials, the literature is plenty of examples and techniques, but only recently we have seen some efforts to give a unified prospective to the study of the CBS thanks to the review of Reuter [153]. Such review is the starting point for the understading of the TMM presented in the next section, a method for the calculation of the CBS that we have implemented as a post processing tool of the SIESTA code.

5.2 Complex band structure calculations: the Transfer Matrix Method

The first step in the calculation of the CBS of a material consists in identifying a layered structure of the system and in writing a layered Hamiltonian accordingly. An example of

what we intend for layered structure is presented in Figure 5.2. The subspace $z > 0$ is divided in repetitive units (layer 1, layer 2 and so on) in the direction z. In our example, each layer has the same hight of the elementary cell, but this is not a requirement of the theory; larger layers can be considered. In the previous section we mentioned how the CBS assumes meaning when the translational symmetry is broken, otherwise no complex $\mathbf{k}$ would be allowed. The layered structure is a mathematical stratagem to break the translational symmetry in one direction. The choice of the layered structure, in fact, dictate the direction along which the translational symmetry is broken, hence the direction where $\mathbf{k}$ is allowed to be complex. For instance, if we want to investigate the CBS in the [111] direction of a crystal (meaning the evanescent functions in that direction), the layers will be perpendicular to the vector [111] and each of them should form a repetitive unit in that direction.

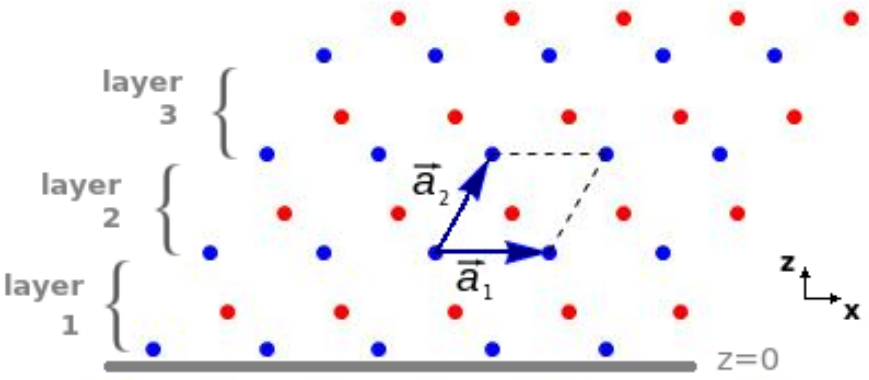


Figure 5.2: Layered structure of a 2D crystal. z=0 defines the line at which the translational symmetry is broken.

With this formalism it is not possible to study situations where the translational symmetry is broken in two directions. However, as explained in the previous section, the case when $\mathbf{k}$ is complex in one direction only is the most relevant and we will limit the discussion to that. In the rest of the chapter, we refer with k to the component of the wave-vector in the direction where the translational symmetry is broken (direction perpendicular to the layers, z in the example), while the other two components (for which the wave vector is strictly real) will be implicitly treated as parameters, $\mathbf{k}_{\parallel}$.

We now assume to introduce an expansion of the wave function over a localized basis set and to be able to define the Hamiltonian matrix elements. The layered Hamiltonian has the form:

$$H = \begin{bmatrix} H_{11} & H_{12} & H_{13} & H_{14} & ... \\ H_{21} & H_{22} & H_{23} & H_{24} & ... \\ H_{31} & H_{32} & H_{33} & H_{34} & ... \\ H_{41} & H_{42} & H_{43} & H_{44} & ... \\ ... & ... & ... & ... & ... \end{bmatrix}. \tag{5.3}$$

The entries of H, H_{ij}, are matrices composed by the Hamiltonian matrix elements. In particular, H_{ij} is the part of the Hamiltonian that couples layer i and layer j, $i = 1$ is the surface layer. Neglecting any form of disorder at the surface and each layer being a

repetitive unit, the interaction between two layers depends only on their relative distance so that all the elements on the diagonal are the same and equal to $H_{11} == H_D$. Using the same argument we can identify with H_{S1} all the first nearest layers interactions, with H_{S2} the interaction between second nearest layers and so on. In general H_{SN} is the interaction between layer 1 and layer 1+N. This would eventually be zero after a certain N because of the finite range of the interaction. The total Hamiltonian then writes:

$$H = \begin{bmatrix} H_D & H_{S1}^\dagger & H_{S2}^\dagger & H_{S3}^\dagger & ... \\ H_{S1} & H_D & H_{S1}^\dagger & H_{S2}^\dagger & ... \\ H_{S2} & H_{S1} & H_D & H_{S1}^\dagger & ... \\ H_{S3} & H_{S2} & H_{S1} & H_D & ... \\ ... & ... & ... & ... & ... \end{bmatrix}. \tag{5.4}$$

In the code SIESTA it is possible to access the integrals of the Hamiltonian in real space between the orbitals in the elementary cell and a supercell. For this reason it is very easy to access elements of the Hamiltonian in Eq. (5.4) when the first layer contains the elementary cell. These are obtained by using Eq. (2.76) but summing only over cells belonging to the appropriate layer. The dimension of each $H_{D/S}$ matrix is defined by the number of atomic orbitals in the elementary cell (meaning the richness of the basis set considered). In SIESTA there is the additional complication of having a non-orthogonal basis. A layered structure equivalent to the one in Eq. (5.4) can be defined for the overlap matrix S. The same notation used for H is used for S as well: S_D contains overlap matrix elements among the same layer, S_{SN} those between layer 1 and layer 1+N.

Importantly, all the elements of $H_{D/S}$ and $S_{D/S}$ depend on $\mathbf{k}_\parallel$. Inside each layer, the periodicity in the direction parallel to the layer (direction x in Figure 5.2) is maintained and $\mathbf{k}_\parallel$ are still good quantum numbers. We do not write explicitly the dependence to avoid extra indexes, but the concept is recovered in the following of the section.

The next step in the study consists in producing an equation where E is fixed and the allowed k (real and complex) vectors are computed. This is the opposite of what it is done in standard band structure calculations, where E is calculated for a set of $\mathbf{k}$. However, the latter approach would be inconvenient in the present situation because k now spans the space $\mathbb{C} = \mathbb{R} \times \mathbb{R}$, which is a much bigger space than that spanned by E (remembering we are interested in real energy only). The idea of the TMM is to project the entire Schrödinger equation (also the wave functions) into a layers structure and to invoke the Bloch theorem to have an eigenvalue problem for the quantity $\exp(ika)$. The quantity a is the thickness of the layer. This is explained in the following. By defining P_j the projector on layer j, which satisfies the relations $I = \sum_j P_j$ (I identity matrix) and $P_j^2 = P_j$, one can write

$$H_{jk} = P_j H P_k, \quad S_{jk} = P_j S P_k, \quad P_n \varphi = \varphi_n. \tag{5.5}$$

Here φ_n is the wave function projected on layer n. The full Schrödinger equation projected on the layer n takes the form

$$P_n H \varphi = E P_n S \varphi, \tag{5.6}$$

which can also be written as

$$P_n H \sum_j P_j^2 \varphi = E P_n S \sum_j P_j^2 \varphi. \tag{5.7}$$

For the sake of clarity, we assume here to have non-zero matrix elements only up to the second layer interaction, although the theory is absolutely general. We can write the Schrödinger equation as

$$\begin{aligned} P_n H(P_{n-2}\varphi_{n-2} + P_{n-1}\varphi_{n-1} + P_n\varphi_n + P_{n+1}\varphi_{n+1} + P_{n+2}\varphi_{n+2}) = \\ E P_n S(P_{n-2}\varphi_{n-2} + P_{n-1}\varphi_{n-1} + P_n\varphi_n + P_{n+1}\varphi_{n+1} + P_{n+2}\varphi_{n+2}, \end{aligned} \tag{5.8}$$

which reduces to

$$\begin{aligned} H_{S2}\varphi_{n-2} + H_{S1}\varphi_{n-1} + H_D\varphi_n + H_{S1}^\dagger\varphi_{n+1} + H_{S2}^\dagger\varphi_{n+2} = \\ E(S_{S2}\varphi_{n-2} + S_{S,1}\varphi_{n-1} + S_D\varphi_n + S_{S,1}^\dagger\varphi_{n+1} + S_{S,2}^\dagger\varphi_{n+2}). \end{aligned} \tag{5.9}$$

Introducing the tautologies $\varphi_l = I\varphi_l$ with $l = n+1, n, n-1$, one obtains:

$$\begin{aligned} &\begin{bmatrix} H_{S2}^\dagger - ES_{S2}^\dagger & H_{S1}^\dagger - ES_{S1}^\dagger & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix} \begin{bmatrix} \varphi_{n+2} \\ \varphi_{n+1} \\ \varphi_n \\ \varphi_{n-1} \end{bmatrix} = \\ &\begin{bmatrix} 0 & ES_D - H_D & ES_{S1} - H_{S1} & ES_{S2} - H_{S2} \\ I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \end{bmatrix} \begin{bmatrix} \varphi_{n+1} \\ \varphi_n \\ \varphi_{n-1} \\ \varphi_{n-2} \end{bmatrix}. \end{aligned} \tag{5.10}$$

The Bloch theorem imposes the following relations between the wave function projectors:

$$\begin{bmatrix} \varphi_{n+2} \\ \varphi_{n+1} \\ \varphi_n \\ \varphi_{n-1} \end{bmatrix} = e^{ika} \begin{bmatrix} \varphi_{n+1} \\ \varphi_n \\ \varphi_{n-1} \\ \varphi_{n-2} \end{bmatrix}. \tag{5.11}$$

This allows to conclude that

$$\begin{aligned} &e^{ika} \begin{bmatrix} H_{S2}^\dagger - ES_{S2}^\dagger & H_{S1}^\dagger - ES_{S1}^\dagger & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix} \begin{bmatrix} \varphi_{n+1} \\ \varphi_n \\ \varphi_{n-1} \\ \varphi_{n-2} \end{bmatrix} = \\ &\begin{bmatrix} 0 & ES_D - H_D & ES_{S1} - H_{S1} & ES_{S2} - H_{S2} \\ I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \end{bmatrix} \begin{bmatrix} \varphi_{n+1} \\ \varphi_n \\ \varphi_{n-1} \\ \varphi_{n-2} \end{bmatrix}. \end{aligned} \tag{5.12}$$

Equation 5.12 is then a generalized eigenvalue problem, whose eigenvalues are $\exp(ika)$, from which the allowed k vectors (both real and complex) for each energy can be obtained.

In principle one could invert $H_{SN}^{\dagger} - ES_{SN}^{\dagger}$ and solve a normal eigenvalue problem, however, most of the times $H_{SN} - ES_{SN}$ is singular and mathematical strategies need to be introduces to remedy at this problem [154, 153]. In the generalized eigenvalue problem instead, the singularity of $H_{SN} - ES_{SN}$ has the only effect to include additional zeros and infinite eigenvalues. Numerical problems have been reported in case the matrix is near-singular [153] as this leads to very large eigenvalues of the Eq. (5.12) and, in turn, to numerical instability in the determination of large ks. We do not encounter such problem in our application and we are anyway interested in small values of k.

A drawback of the TMM is the fact that the dimension of the matrices in Eq. (5.12) scales with a factor of $2L$ compared to the Hamiltonian matrix. L is the number of layers interaction considered. The finite range of the SIESTA basis and the ability of the code to accurately describe systems employing small basis sets is essential for the use of the TMM. The L is regulated by the range of the interaction (that in SIESTA can be controlled with the radius of the basis cutoffs) and the dimension of the elementary cell. For minimal cells and typical values of the basis cutoff radii, we did not encounter any situation where $L > 4$ was needed. In addition to the fact that a basis sets containing about 15 elements per atom are considered to be already pretty accurate, we can conclude that will be unlikely to have situations where one needs to compromise on the accuracy of the calculation due to memory issues connected to the dimensions of the matrices in Eq. (5.12).

Regarding the actual implementation, our code uses the SIESTA's .HSX file contenting the information about the Hamiltonian and the overlap integrals and the ORB_INDX file containing informations about the basis set. The .XV file is also read to extract some information about the structure. The LAPACK [155] routine ZGGEV is used to solved the generalized eigenvalue problem in Eq. (5.12). No restriction to the shape of the calculation cell is required. The method can calculate the CBS in any of the three directions perpendicular to the planes formed by the three lattice vectors. Tools for crystal lattice management (pymatgen [156] for instance) can be used to set a minimal cell exposing a particular surface and basically allow us to explore the CBS in any possible direction. The code is available at `https://github.com/bosonie`.

A typical result of the CBS calculation using our code is showed in Figures 5.3 and 5.4. The two pictures present the same information but in different formats.

In the first picture (Figure 5.3), the energy is on the vertical axis, while the wave vector is on the horizontal one, like in any standard band structure diagram. The picture is divided in three panels: the middle one contains the energy derived from pure real k (normal bands), the left-hand side one shows the allowed energies as function of $\mathfrak{Im}(k)$ when $\mathfrak{Re}(k) = 0$ (pure imaginary bands), while the right-hand side panel reports the same quantity but when $\mathfrak{Re}(k) = \pi/a$ (a=6.96 Bohr - black dots) and $\mathfrak{Re}(k) = 0.29$ 1/Bohr (gray dots). This way of presenting the CBS of a material resembles the familiar band structure and emphasizes nicely the intimate connection between the real and complex bands of a material. This is why it will be used to explain the analytical properties of the complex

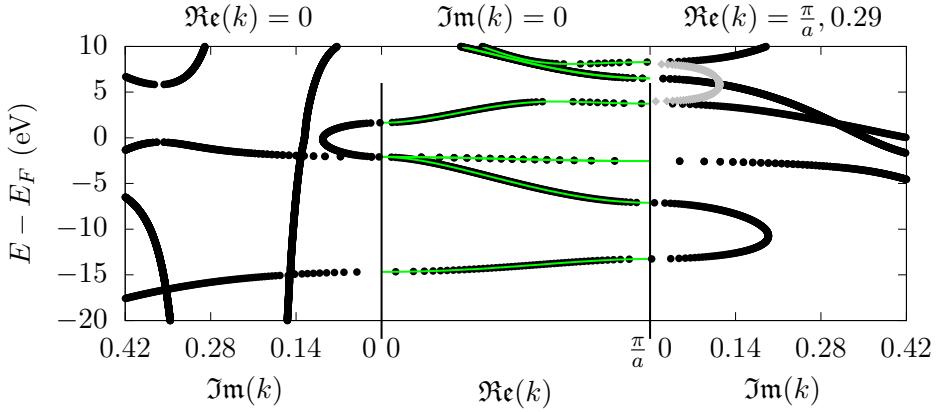


Figure 5.3: Black and gray dots are the prototypical result of a CBS calculation obtained with the TMM as a post processing tool of a SIESTA calculation. In green we plot the real bands calculated by SIESTA. The calculation refers to the CBS of MgS along the [111] direction, at $\mathbf{k}_{\parallel} = (0, 0)$. The k vector is in units of 1/Bohr. Brillouin zone boundaries are marked with the label π/a (a=6.96 Bohr) instead of the actual value in order to facilitate comparison with the models in Section 5.3. The left-hand side panel presents $E(\Im\mathfrak{m}(k))$ when $\Re\mathfrak{e}(k) = 0$. The middle panel presents $E(\Re\mathfrak{e}(k))$ when $\Im\mathfrak{m}(k) = 0$. The right-hand side panel presents $E(\Im\mathfrak{m}(k))$ when $\Re\mathfrak{e}(k) = \pi/a$ (black dots) and when $\Re\mathfrak{e}(k) = 0.29$ 1/Bohr (gray dots).

band structures in the following section. However, this type of graphs lacks of clarity when complex bands with many different real parts (not only 0, π/a and 0.29 1/Bohr as in our example) are present. Moreover, different strategies have to be used when real energies are allowed on lines where the real and imaginary part of k change simultaneously (see reference [157] for an example). This justifies the use of other ways to present the CBS.

In Figure 5.4 the energy is on the horizontal axis while the k vector is on the vertical one. The origin, 0, divides the vertical axis in two part. Below it, the $\Re\mathfrak{e}(k)$ as a function of the energy when $\Im\mathfrak{m}(k) = 0$ is reported. This is the standard band structure. On the positive plane, instead, the imaginary part is drawn as a function of the energy. In addition, the real part of each band is color coded (see color scale on the right-hand side). This is a redundant information for real bands, however makes much more clearer the interpretation of the complex bands as now the information on $\Re\mathfrak{e}(k)$ and $\Im\mathfrak{m}(k)$ are both present on the same band line. Because for spintronics applications the value of $\Im\mathfrak{m}(k)$ and its symmetry around the Fermi energy are the quantity of interest, we will use this way to present the results in Section 5.4.

The CBS is a property of a material and of the direction along which one breaks the translational symmetry. However, it is important to remind that the CBS depends on the $\mathbf{k}_{\parallel}$ as well. In the plane perpendicular to the direction where the translational symmetry is broken, the system is periodic. It is then possible to define a 2D lattice in real space and consequently a 2D reciprocal lattice. We call $\mathbf{q_1}$ and $\mathbf{q_2}$ the reciprocal lattice vectors. In Figure 5.5 we present a comparison between the CBS at two different $\mathbf{k}_{\parallel}$, in particular at $\mathbf{k}_{\parallel} = (0, 0)$ (left-hand side panel) and at $\mathbf{k}_{\parallel} = (0.1\mathbf{q_1}, 0.1\mathbf{q_2})$ (right-hand side panel). Only energies in the close vicinity of the Fermi energy are reported. It can be seen that the CBS is different in the two situations. Especially, we notice that, at the Fermi energy, the minimum $\Im\mathfrak{m}(k)$ is at a value of about 0.12 1/Bohr for $\mathbf{k}_{\parallel} = (0, 0)$, while it is larger

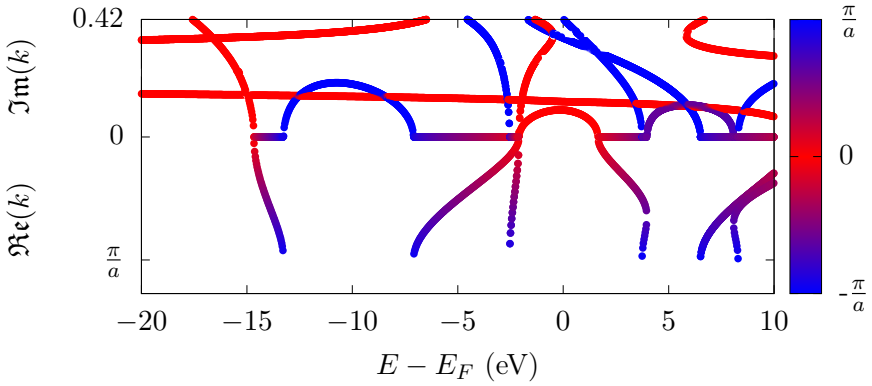


Figure 5.4: Prototypical result of a CBS calculation obtained with the TMM as a post processing tool of a SIESTA calculation. Pure real bands ($\Re\mathfrak{e}(k)$ when $\Im\mathfrak{m}(k)$=0) are reported below the origin of the vertical axis. Above the origin the imaginary bands are reported. The color encodes the $\Re\mathfrak{e}(k)$ part of each line. The calculation refers to the CBS of MgS along the [111] direction, at $\mathbf{k}_{\parallel} = (0, 0)$. The k vector is in units of 1/Bohr. Brillouin zone boundaries are marked with the label π/a (a=6.96 Bohr) instead of the actual value in order to facilitate comparison with the models in Section 5.3.

than 0.15 1/Bohr for the case $\mathbf{k}_{\parallel} = (0.1\mathbf{q_1}, 0.1\mathbf{q_2})$. In spintronics applications the value of the minimum $\Im\mathfrak{m}(k)$ is the important quantity to investigate, because it determines the decay of the evanescent wave function that contributes more to the transport during the tunneling process. For this reason, the identification of the $\mathbf{k}_{\parallel}$ at which the minimum $\Im\mathfrak{m}(k)$ appears is a valuable information. It is then useful to draw maps of the minimum $\Im\mathfrak{m}(k)$ for a certain energy as a function of $\mathbf{k}_{\parallel}$. Figure 5.6 shows how this quantity changes for the material MgS along the [111] direction at the Fermi energy. The x and y axis in Figure 5.6 are in units of the module of the 2D reciprocal lattice vectors, namely $\mathbf{q_1}$ and $\mathbf{q_2}$. In Figure 5.6 the center of the zone presents the minimum $\Im\mathfrak{m}(k)$ among all the $\mathbf{k}_{\parallel}$. This is expected for example in the case of the simple potential barrier of Appendix A, however it is not a general statement.

More phenomenology about the CBS will be presented in Section 5.4, together with some scientifically relevant results. Before that, however, another topic needs to be addressed. In Figure 5.5 one can notice the presence of an almost flat pure imaginary band crossing the entire energy range. This is a feature observed in all the CBS calculations performed with our code. In Figure 5.7 we compare a result of the CBS obtained with our method and the calculation on the same system performed with PWcond, a tool included in the Quantum Espresso package that uses Green's function formalism to compute the CBS [104]. The calculation refers again to MgS, however along the direction [100]. We use a plotting method similar to that of Figure 5.4, however we drop the color coding in order to better compare with PWcond. The matching is almost perfect for both real and complex bands, except for the presence of a large number of flat bands obtained with our method. The flat bands are definitely a feature of our code only. Figure 5.7 is obtained with the TMM on top of a SIESTA calculation employing DZDP basis set. Performing calculations with other basis set choices, we notice that the position of the flat bands changes in a more significant way compared to all the other bands. The analysis of the origin of the flat

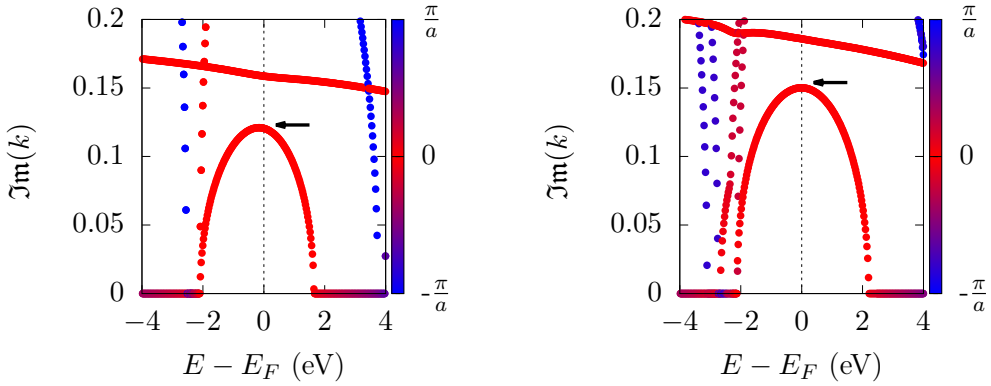


Figure 5.5: CBS obtained with the TMM for MgS along the [111] direction. On the left-hand side panel the case of $\mathbf{k}_{\|} = (0, 0)$ is investigated. On the right-hand side panel we report, instead, the calculation for $\mathbf{k}_{\|} = (0.1\mathbf{q_1}, 0.1\mathbf{q_2})$, where $\mathbf{q_1}$ and $\mathbf{q_2}$ are the 2D reciprocal lattice vectors. The y axis report the imaginary part of the complex band. Color coded is the real part. The k vector is in units of 1/Bohr. Brillouin zone boundaries are marked with the label π/a (a=6.96 Bohr) instead of the actual value.

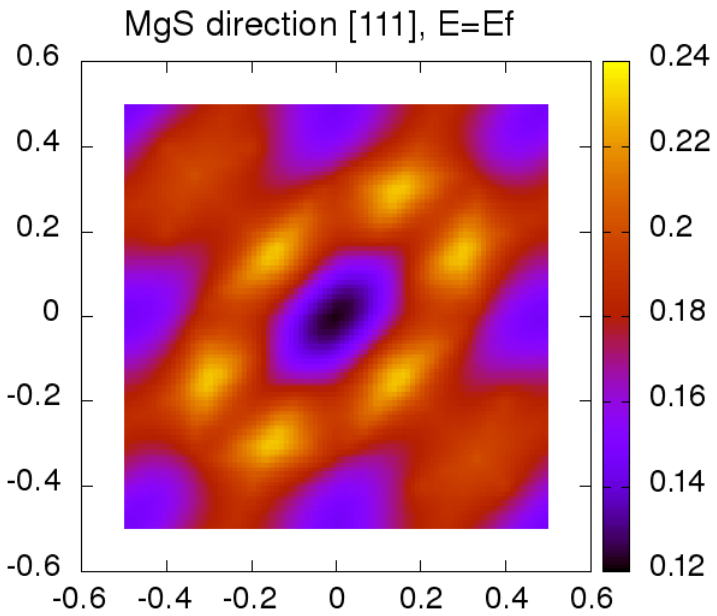


Figure 5.6: Minimum value of $\Im\mathfrak{m}(k)$ associated to the CBS of MgS along [111] at the Fermi energy, as a function of $\mathbf{k}_{\|}$. The x and y axis are in units of the module of reciprocal lattice vectors. The value of the minimum $\Im\mathfrak{m}(k)$, color coded, is in units of 1/Bohr.

bands, their interpretation and the method to deal with them are the subject of the next section. This requires a deep understanding of the analytical properties expected from a CBS.

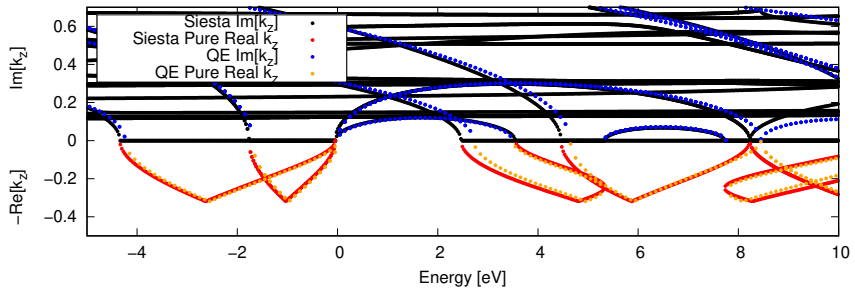


Figure 5.7: CBS of MgS along the direction [100]. Comparison between our code (black lines for $\mathfrak{Im}(k)$ and red lines for the real bands) and PWCond (blue dots for $\mathfrak{Im}(k)$ and yellow dots for the real bands).

5.3 Analytical properties of Complex Band Structure

The focus of this section is the description of the analytical properties expected for the complex quantity $E(\mathbf{k})$ and the "lines of real energy" that collectively form the CBS. The topic was first discussed by Kohn [151] and Heine [150] using the theory of differential equations and arguments of complex analysis. Of great interest for us is the work of Krieger [158], that derived most of the properties for finite-band models and using algebraic arguments only. In the first part of this section we focus, as in Krieger paper, on the situation where the model involves the expansion over an orthonormal basis set. Subsequently the theory is extended to situations where the orthonormality condition is relaxed, like in a SIESTA calculation. This is sufficient to demonstrate the origin of the flat bands and to put the basis for discussing the best way to treat them.

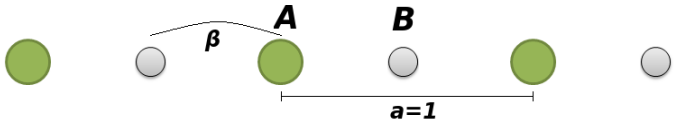


Figure 5.8: Diatomic chain used as a model for the analytic calculation of complex band structures.

Two simple models come to help to better visualize the significance of the analytical properties of the CBS. Both refers to the case of a simple 1D chain composed by two atoms and lattice constant equal to one ($a = 1$ in Figure 5.8). The first model is the empty lattice model, that is taken as an example of a situation where the full Hamiltonian is solved. The second model is a tight-binding study and represents a simple case of a finite-band approximation. The latter model will be treated in both the case of a non-orthogonal basis set and the case of an orthogonal one. The models are detailed in the next section. They are used, firstly, to compared the differences between finite models and the situation where the full Hamiltonian is solved (subsection 5.3.2). Secondly, the tight binding model is used to investigate the difference between the situation when a non-orthogonal basis set is introduced, with respect to the situation when a orthogonal basis set is employed (subsection 5.3.3).

5.3.1 Simple analytical models

The first model is the empty lattice model, where the interaction between electrons and nuclei sites is negligible. The electrons are effectively free electrons and the only effect of the 1D chain (Figure 5.8) is to impose a periodicity in the electron wave function (Bloch theorem). No introduction of basis is required. Assuming $\hbar^2/2m$ equal to one, the dispersion relation in this situation is simply,

$$E_n(k) = (k + 2\pi n)^2, \tag{5.13}$$

with $n = 0, \pm1, \pm2, ...$ If we allow k to be complex, $k = \mathfrak{Re}(k) + i\mathfrak{Im}(k)$, we write:

$$E_n(k) = (\mathfrak{Re}(k) + 2\pi n)^2 - \mathfrak{Im}(k)^2 + 2i\mathfrak{Im}(k)(\mathfrak{Re}(k) + 2\pi n). \tag{5.14}$$

Here E assumes real values when $\mathfrak{Im}(k) = 0$ (normal bands), but also when $\mathfrak{Re}(k) + 2\pi n = 0$. In that case we have the relation

$$E(k) = -\mathfrak{Im}(k)^2. \tag{5.15}$$

A line of real energy is depicted in Figure 5.9. This originates from $E = -\infty$ in the imaginary plane ($\mathfrak{Re}(k) = 0$, left-hand side panel of Figure 5.9) and approaches the real plane ($\mathfrak{Im}(k) = 0$, middle panel Figure 5.9) with $E = -\mathfrak{Im}(k)^2$ dependence. At $E = 0$ it turns by 90° in the real plane where it stays bouncing from $\mathfrak{Re}(k) = 0$ to $\mathfrak{Re}(k) = \pi$ till reaching $+\infty$. Negative values of $\mathfrak{Im}(k)$ (not depicted in 5.9) assumes the same behavior in the imaginary plane and they turn again in the real plane at $E = 0$, but in direction of $\mathfrak{Re}(k) = -\pi$. The reader should in fact consider that a mirror function of the central panel in Figure 5.9 is present between $\mathfrak{Re}(k) = -\pi$ and $\mathfrak{Re}(k) = 0$ (again not depicted). In total there are then two lines of real energy that can be followed continuously from $-\infty$ to $+\infty$. At $\mathfrak{Re}(k) = \pi$ the energy is degenerate. The peculiar properties of this point (and of its equivalents) are detailed in reference [152].

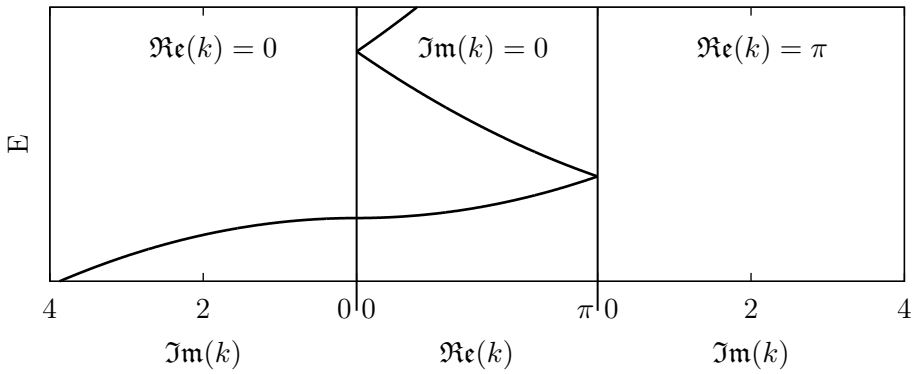


Figure 5.9: Real (middle panel) and complex (right panel) bands for the empty lattice model. In the left panel $\mathfrak{Re}(k) = 0$, thus it reports pure imaginary bands.

This concludes the description of the CBS of the empty lattice model for the first

Brillouin zone $[-\pi,\pi]$. However, one needs to remember the periodicity of the real bands. It is easy to see that the periodicity on the real plane is followed by the lines of real energy as well. In fact, we said that the energy assumes the real values described in Eq. 5.15 when $\mathfrak{Re}(k) + 2\pi n = 0$, not only when $\mathfrak{Re}(k) = 0$. Actually the entire dispersion relation $E(k)$ follow the periodicity in $\mathfrak{Re}(k)$, as discussed in the next subsection.

It would be interesting now to discuss what happens when one opens a gap in this model. This is possible by introducing, for example, the effect of the potential as a perturbation that modifies the bands at the edges of the Brillouin zone [152, 68]. However, the conclusions would not be much different respect to the situation described with the second model. For this reason we ignore this case and we concentrate on the second model.

The second model describes the opposite situation with respect to that introduced by the empty lattice model in terms of how we treat the electrons. It is, in fact, a first nearest-neighbors tight-binding model where the electrons are tightly bonded to the site they belong to. Referring to Figure 5.8, the sites A interacts only with the nearest B sites and *vice-versa* the sites B interacts only with the nearest A sites. In order to solve the problem, a basis set needs to be introduced. Thus the wave function of the system is expressed as a combination of atomic orbitals located on each site (one orbital per site):

$$\varphi(r) = \sum_{i}^{\infty} A_i \phi_{Ai}(r) + B_i \phi_{Bi}(r) \tag{5.16}$$

The index i runs over all the units cells of the linear chain. A unit cell is indicated by "a=1" in Figure 5.8. The elements of the basis $\phi_{A/B}(r)$ are the solutions of the Schrödinger equation for the isolated atom A/B respectively. As we have mentioned already before, for the moment no overlap is considered, meaning that the basis is orthonormal. By expanding the Schrödinger equation over such tight-binding basis, by projecting on the site A and B of a particular elementary cell, and by assuming first-nearest neighbors interaction, one obtains the bands equation

$$\begin{bmatrix} \varepsilon_A & \beta(1+e^{-ik}) \\ \beta(1+e^{ik}) & \varepsilon_B \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = E(k) \begin{bmatrix} A \\ B \end{bmatrix}, \tag{5.17}$$

where ε_A and ε_B are on site energies,

$$\varepsilon_{A/B} = \int \phi^*_{A/Bi}(r) H(r) \phi_{A/Bi}(r) dr, \tag{5.18}$$

and β is the hopping parameter,

$$\beta = \int \phi^*_{Ai}(r) H(r) \phi_{Bi}(r) dr = \int \phi^*_{Bi}(r) H(r) \phi_{Ai}(r) dr. \tag{5.19}$$

The problem can be solved analytically,

$$E^2(k) - E(k)(\varepsilon_A + \varepsilon_B) + \varepsilon_A \varepsilon_B - 2\beta^2(1+\cos(k)) = 0, \tag{5.20}$$

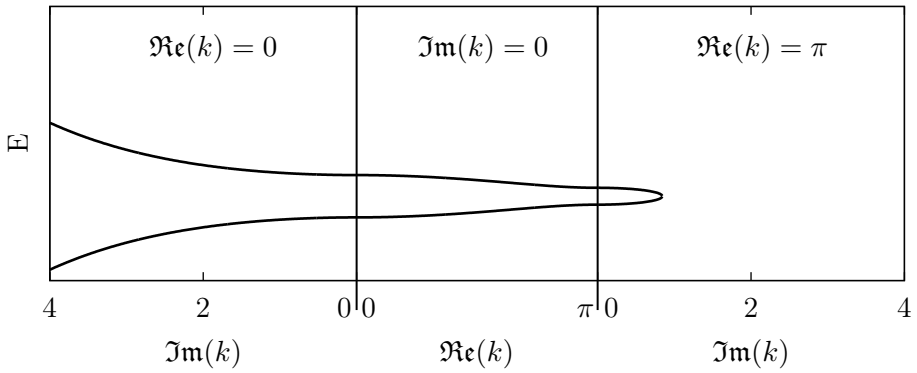


Figure 5.10: Real (middle panel) and complex (right and left panels) bands for the two sites, nearest-neighbor, tight-binding model with choice of parameters $\varepsilon_A = 7$, $\varepsilon_B = 3$, $\beta = 2.3$. In the left panel $\mathfrak{Re}(k) = 0$, thus it reports pure imaginary bands. In the right panel $\mathfrak{Re}(k) = \pi$.

$$E(k) = \Sigma \pm \sqrt{\Xi^2 + 2\beta^2(1 + \cos(k))}, \tag{5.21}$$

where,

$$\Sigma = \frac{\varepsilon_A + \varepsilon_B}{2}, \qquad \Xi = \frac{\varepsilon_A - \varepsilon_B}{2}. \tag{5.22}$$

It is now worth analyzing the values of $\cos(k)$ as a complex variable,

$$\cos(k) = \frac{e^{\mathfrak{Re}(k)+i\mathfrak{Im}(k)} + e^{-\mathfrak{Re}(k)-i\mathfrak{Im}(k)}}{2} \tag{5.23}$$

which is clearly real for $\mathfrak{Im}(k) = 0$, but also when $\mathfrak{Re}(k) = n\pi$ where $n = 0, \pm 1, \pm 2...$ We now distinguish two cases:

- n even

$$\cos(k) = \frac{e^{i\mathfrak{Im}(k)} + e^{-i\mathfrak{Im}(k)}}{2} = \cosh(\mathfrak{Im}(k)), \tag{5.24}$$

- n odd

$$\cos(k) = \frac{-e^{i\mathfrak{Im}(k)} - e^{-i\mathfrak{Im}(k)}}{2} = -\cosh(\mathfrak{Im}(k)). \tag{5.25}$$

The use of this information allows to plot the lines of real energy. One of them for a choice of parameters $\varepsilon_A = 7$, $\varepsilon_B = 3$, $\beta = 2.3$ is presented in Figure 5.10. The first main difference with respect to the empty lattice is the presence of a gap. The line of real energy now at $\mathfrak{Re}(k) = \pi$ leaves the first real band to make another 90° turn in the imaginary plane (right-hand side panel of Figure 5.10). Subsequently, the line reaches a maximum and then it comes back towards $k = \pi + i * 0$ to join to the second real band. Moreover, the finite nature of the model makes impossible for the line to stay in the real plane while $E \to \infty$. It is a pure imaginary band ($\mathfrak{Re}(k) = 0$, left-hand side panel of Figure 5.10) to go to $+\infty$.

As in the case of the empty lattice model, a mirror of the central panel in Figure 5.10 describes the bands between $\mathfrak{Re}(k) = -\pi$ and $\mathfrak{Re}(k) = 0$ (not depicted), and the negative values of $\mathfrak{Im}(k)$ have the same energy of the corresponding positive values. In other words again we have two lines of real energy. A special point in this situation is the maximum of $\mathfrak{Im}(k)$ in the right-hand panel, that is where two energies meet. This discussion gives

a sense of the fact that the real bands are connected to each other even in the case of a band gap. The connections happens, however, at complex k. The full picture would be clearer extending the analysis to the full complex variable $E(k)$, this is done in the next section for the general description of the analytical properties. Finally, also for this model, the periodicity in the real plane is maintained by the complex bands.

We conclude this subsection summarizing the main differences of the similarities of the two models. Both models describe a CBS that is periodic in the real plane with period 2π. Two lines of real energy are present and they can be followed continuously from $E-\infty$ to $E+\infty$. The empty lattice model describes an infinite number of real bands and therefore the line of real energy stays in the real plane while the energy goes to $+\infty$. On the contrary, the tight-binding model has just two real bands and therefore it is a pure imaginary band to go to $+\infty$. This is due to the finiteness of the model. In the case of the empty lattice model, real bands are connected in the real plane. In the case of the tight-binding model, instead, a gap is present and the the real bands are connected through a loop in the complex plane.

5.3.2 Properties of the CBS in the case of orthonormal basis set

The two models analyzed so far have some common features that can be seen as a first indication of general mathematical rules a CBS must obey. However, the models are strictly 1D and the extension to 3D systems where $\mathbf{k}$ is complex in one direction only is not so obvious. We now want to express in the most exhaustive way the full set of analytical properties of the full dispersion relation and the lines of real energy. When possible, a rigorous derivation of the property is presented following the work for finite models of Kreigher [158]. To focus on that derivation is important to understand subsection 5.3.3, when the properties are discussed for models where a non-orthogonal basis set is employed. Only a few properties are not proven by Kreigher [158]. For these the derivation of Heine [150] and Prodan [152] is referenced, however not explained.

Let us start by recollecting some definitions in a mathematical formal way. The Schrödinger equation is

$$H\varphi_{\mathbf{k}}(\mathbf{r}) = E\varphi_{\mathbf{k}}(\mathbf{r}). \tag{5.26}$$

Because of the Bloch theorem, it is always possible to write

$$\varphi_{\mathbf{k}} = e^{i\mathbf{kr}}u_{\mathbf{k}}. \tag{5.27}$$

Consequently one has:

$$H(\mathbf{k})u_{\mathbf{k}}(\mathbf{r}) = Eu_{\mathbf{k}}(\mathbf{r}), \qquad H(\mathbf{k}) = e^{-i\mathbf{kr}}He^{i\mathbf{kr}}. \tag{5.28}$$

From this expression, it can be proven that for a solid, even including spin-orbit interaction that couples spin and lattice, $H(\mathbf{k})$ is always a polynomial function of the components of $\mathbf{k}$. This means that, also when the wave function is expanded on a basis (orthonormal for

the moment), the Schrödinger equation,

$$\overline{H}(\mathbf{k})\mathbf{c} = E\mathbf{c}, \tag{5.29}$$

is a matrix equation where all the matrix elements of $\overline{H}(\mathbf{k})$ are still polynomial functions of the components of $\mathbf{k}$.

We allow one component of $\mathbf{k}$ to be complex, while the other two remains real. If $\overline{H}(\mathbf{k})$ is an $N \times N$ matrix, then there are N solution, $E_i(\mathbf{k})$. The collection of the $E_i(\mathbf{k})$ is the full complex dispersion relation, $E(\mathbf{k})$. In the terminology of complex analysis the $E_i(\mathbf{k})$ can be seen as Riemann surfaces of the complex multi-valued function, $E(\mathbf{k})$.

With this in mind, a list of analytical properties of $E(\mathbf{k})$ can be demonstrated.

1. The $E_i(\mathbf{k})$ are analytic functions of $\mathbf{k}$ except at a set of isolated singular points corresponding to the branch points of $E(\mathbf{k})$. A branch point is a point in the complex plane, where two Riemann surfaces connects. To be more formal we can use the operative definition of Needham [159]. A point q is a branch point if for any close loop in k-space around it, it is not possible to return to the same $E(k)$. For a generic q, if one starts from a point $\overline{k}$ corresponding to $E(\overline{k})$, once a loop around q is concluded, the energy is still $E(\overline{k})$. However, if q is a branch point the statement is not true. We know that $E(k)$ is a multi-valued complex function. Therefore, by looping around a branch point we end up to another value of $E(k)$ for the same k. In the discussion we have dropped the vector symbol for $\mathbf{k}$ as we are focusing on the component that is allowed to be complex. In general q depend on $\mathbf{k}_{\parallel}$. The proof of this first property is given by Kreigher [158] observing that the characteristic polynomial P of Eq. (5.29) is of the form:

$$P = E^N + \gamma_1(\mathbf{k})E^{N-1} + \gamma_2(\mathbf{k})E^{N-2} + ... + \gamma_{N-1}(\mathbf{k})E + \gamma_N(\mathbf{k}) = 0. \tag{5.30}$$

   Because the elements of $\overline{H}(\mathbf{k})$ are polynomial function of $\mathbf{k}$ then the $\gamma_i(\mathbf{k})$ in Eq. (5.30 are as well polynomials in components of $\mathbf{k}$. It is important here to recall the definition of a polynomial: it is an expression consisting of variables and coefficients, that involves only the operations of addition, subtraction, multiplication, and non-negative integer exponents of variables. No division is allowed in the definition, a fact that will be important later. Since the $E_i(\mathbf{k})$ are solutions of an algebraic equation, whose coefficients are polynomials of $\mathbf{k}$, and a polynomial function is analytic at every point in $\mathbb{C}$, one can conclude that the $E_i(\mathbf{k})$ are analytic function for each $\mathbf{k}$, except at a set of singular points where two $E_i(\mathbf{k})$ connects, meaning the branch points of $E(\mathbf{k})$. Kreigher [158] also showed that these branch point must exist, but it is less important for out theory. The take home message at this point is that there are no singularities other than the branch points.

2. E is real for real $\mathbf{k}$ and $E(\mathbf{k}) = E^*(\mathbf{k}^*)$, where the symbol * indicate the complex conjugate. The proof of the first statement is straightforward: since for real $\mathbf{k}$ the matrix $\overline{H}(\mathbf{k})$ is hermitian, its eigenvalues $E_i(\mathbf{k})$ must be real. Moreover, recalling the fundamental theorem of algebra, the characteristic polynomial of Eq. (5.30) can be

written as:

$$P = [E - E_1(\mathbf{k})]\,[E - E_2(\mathbf{k})]\,...\,[E - E_N(\mathbf{k})] = 0. \tag{5.31}$$

Hence the $\gamma_i(\mathbf{k})$ are sums of products of the $E_i(\mathbf{k})$, which are real for real $\mathbf{k}$. This is sufficient to conclude that for any $\mathbf{k}$, $\gamma_i^*(\mathbf{k}) = \gamma_i(\mathbf{k}^*)$. Using the latter equality on the complex conjugate of 5.30 we finally have $E(\mathbf{k}) = E^*(\mathbf{k}^*)$.

3. Along the component of $\mathbf{k}$ that is allowed to be complex, k, in the neighborhood of a branch point $k = q$, the energy behaves like:

$$E(k) = E(q) + \alpha(k - q)^{1/2}, \tag{5.32}$$

where α is a constant. Note that in general q depends on $\mathbf{k}_\parallel$. The proof is rooted in the observation that the polynomial expansion of Eq. (5.30) can be seen as a function of E and k and expanded as a Taylor series in two variables around $k = q$ and $E = E(q)$, namely

$$\begin{aligned} &P(E,k) = P(E(q),q) + \\ &(E - E(q))\left.\frac{\partial P}{\partial E}\right|_{\substack{k=q \\ E=E(q)}} + (k-q)\left.\frac{\partial P}{\partial k}\right|_{\substack{k=q \\ E=E(q)}} + \frac{1}{2}(E - E(q))^2\left.\frac{\partial^2 P}{\partial E^2}\right|_{\substack{k=q \\ E=E(q)}} + ... \end{aligned} \tag{5.33}$$

By looking at Eq. (5.31), $P(E(q), q)$ is zero and also the first derivative with respect to the energy evaluated at $E(q)$ vanishes. In fact q is the branch point, where two $E_i(\mathbf{k})$ are equal, this means that Eq. (5.31) is proportional to $(E - E(q))^2$, whose derivative is zero for $E = E(q)$. In conclusion in first approximation one has

$$P(E,k) = (k-q)\left.\frac{\partial P}{\partial k}\right|_{\substack{k=q \\ E=E(q)}} + \frac{1}{2}(E - E(q))^2\left.\frac{\partial^2 P}{\partial E^2}\right|_{\substack{k=q \\ E=E(q)}} = 0. \tag{5.34}$$

This leads directly to an expression for the energy of the kind in Eq. (5.32).

4. The branch points are always at a complex k and they comes in pairs. In fact, if q is a branch point, also q^* is. The second statement derives directly from 2., the first one can be understood looking at the expansion of $E(k)$. When q approaches the real axis, also q^* does, therefore:

$$E(k) = E(q) + \alpha(k - q)^{1/2}(k - q^*)^{1/2} \sim (k - q) \tag{5.35}$$

that is not a branch point anymore.

5. $E(\mathbf{k})$ is periodic in real k, and of course in the other two components $\mathbf{k}_\parallel$, so that one can write:

$$E(\mathbf{k}) = E(\mathbf{k} + \mathbf{G}) \tag{5.36}$$

where $\mathbf{G}$ is a reciprocal lattice vector, strictly real. The proof is by Heine [150].

The next part of the list concerns the "lines of real energy", namely the lines that form

what we call the CBS of a material. They are the allowed energies with $\mathfrak{Im}E(\mathbf{k}) = 0$. A direct reference to the models in the previous subsection can be used. In this list we focus on the component k only, $\mathbf{k}_{\|}$ is a parameter.

6. Direct consequence of point 2. is that $E(\mathfrak{Re}(k) + i\mathfrak{Im}(k)) = E(\mathfrak{Re}(k) - i\mathfrak{Im}(k))$. This means that, if k gives real energies, also its complex conjugate will do. This is the reason why $-\mathfrak{Im}(k)$ has the same energy of the corresponding positive values in our models.

7. The real bands (energies from k with $\mathfrak{Im}(k) = 0$) are, of course, part of the line of real energy, however, something more can be said. Every time there is a maximum or a minimum in a real band, the energy makes a 90° turn to go into an imaginary plane (a plane where $\mathfrak{Re}(k)$ is constant). More precisely, a maximum or a minimum in the real band is a point where four real energies meet, the two in the real plane (coming from the left and the right of the maximum/minimum) and two in the imaginary plane (one going towards positive $\mathfrak{Im}(k)$ and one going towards negative $\mathfrak{Im}(k)$; both are necessarily present due to point 6.). Moreover, a minimum in the real bands correspond to a maximum in the imaginary plane and vice-versa a maximum in the real bands correspond to a minimum in the imaginary plane. Therefore the point where the four real energies meet is a saddle point for the full dispersion relation $E(k)$. The proof follows from a simple expansion of $E(k)$ around any point in the real plane, k_0

$$E(k) = E(k_0) + (k - k_0) \left.\frac{\partial E}{\partial k}\right|_{k=k_0} + \frac{1}{2}(k - k_0)^2 \left.\frac{\partial^2 E}{\partial k^2}\right|_{k=k_0} + ... \tag{5.37}$$

   For a k just off the real plane, the first derivative must be real to have real $E(k)$, otherwise $(k-k_0)$ would bring an imaginary contribution. So only at maxima or minima one can have real energies coming from complex k. Analyzing now the second order term, $(k - k_0)^2$ is real only if k is real or if $\mathfrak{Re}(k) = k_0$. This concludes the derivation. There is no guarantee that the line stays in the imaginary plane also very far from the real plane. In other words it is possible to have a line of real energy, where the real and imaginary part of k change simultaneously. In our examples the only maxima or minima were at the Brillouin zone boundaries and at $\mathfrak{Re}(k) = 0$. In that points indeed one has four lines merging and forming a saddle point of $E(k)$. However we have seen already the presence of a maximum different from the edges in Figure 5.3.

8. Lines of real energy that leaves the real plane, turns back to the real plane after going around one or more branch point. The line does not necessarily goes into a branch point, but it happens, for example when there is a mirror plane between layers of the layered structure [150, 153]. Proof is not presented here. For the two band model the maximum point reached by the line of real energy in the right panel was a branch point, the situation is a bit more complicated for the empty lattice model as explained by Prodan [152]. The only exception to this statement is the presence of lines that go towards $E = -\infty$ while $\mathfrak{Im}(k)$ goes to $+\infty$ or $-\infty$. They are not mentioned by Kreigher [158], however they are analyzed extensively by Heine for the full Hamiltonian [150]. For strictly 1D system, just two of such lines are present. However in 3D more of

them appears. This can be seen for example from Figure 5.7. In the case of finite band models, our claim is that also energies going to $E = +\infty$ are present (like we see in the case of the two band model). This is certainly due to the finiteness of the model, but we do not have formal proof of this feature at the moment.

9. Heine proved with complex analysis arguments that the lines of real energy do not split began or terminate. Moreover, their chance of crossing is vanishing probable except for energies on the real plane.

A conclusion to this section, an effort to recall the general picture is presented. The full dispersion relation is a multi-valued function in complex plane formed by Riemann surfaces connected by branch points. The lines of real energy forming the CBS of a material are expected to have a peculiar behaviors: starting from $E = -\infty$, one can follow a line of real energy winding up to bigger energies and jumping on and off the real plane in a continuous way. The line leaves the real axis when there is a maximum or a minimum, it comes back to it after going around a branch point. The line keeps this behavior indefinitely for the full Hamiltonian. In finite models, it is an complex band to go to $+\infty$. This is why the concept of "line of real energy" was introduced in the first place. They are effectively lines one can follow in a continuous way.

5.3.3 Overlap and flat bands

It is now time to investigate the situation when a non-orthogonal basis set is introduced. We first consider the simple two-bands model of the previous section, with a first nearest-neighbor overlap between the basis orbitals ϕ_A and ϕ_B. We call it S:

$$S = \int \phi^*_{Ai}(r)\phi_{Bi}(r)dr = \int \phi^*_{Bi}(r)\phi_{Ai}(r)dr. \tag{5.38}$$

The Schrödinger equation is now

$$\begin{bmatrix} \varepsilon_A & \beta(1+e^{-ik}) \\ \beta(1+e^{ik}) & \varepsilon_B \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = E(k) \begin{bmatrix} 1 & S(1+e^{-ik}) \\ S(1+e^{ik}) & 1 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix}, \tag{5.39}$$

leading to energies:

$$E(k) = \frac{\Sigma - 2\beta S(1+\cos(k))}{1-2S^2(1+\cos(k))} \pm \sqrt{\frac{\Xi^2 + 2(1+\cos(k))(\beta - S\varepsilon_A)(\beta - S\varepsilon_B)}{(1-2S^2(1+\cos(k))^2}} \tag{5.40}$$

It is now difficult to find the complex k leading to real energies. However, it is possible to write an equation for $\cos(k)$ as a function of the energy and, by taking the inverse, to obtain an analytic expression for $k(E)$:

$$k(E) = \arccos\left[\frac{(\varepsilon_A - E)(\varepsilon_B - E)}{2(\beta - ES)^2} - 1\right] \tag{5.41}$$

By inverting numerically Eq. (5.41), one finds the left and right panels of Figure (5.11).

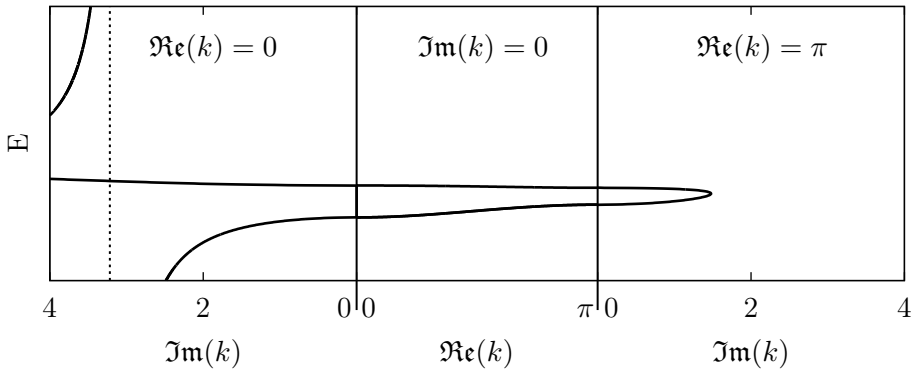


Figure 5.11: Real (middle panel) and complex (right and left panels) bands for the two sites, nearest-neighbor, tight-binding model in presence of an non-orthogonal basis set. The choice of the model parameters is $\varepsilon_A = 7$, $\varepsilon_B = 3$, $\beta = 2.3$, $S = 0.2$. In the left panel $\mathfrak{Re}(k) = 0$, thus it reports pure imaginary bands. In the right panel $\mathfrak{Re}(k) = \pi$.

The first clear difference with respect to the orthonormal basis case, is the presence of a divergence in the line of real energy of the left panel in Figure (5.11). Its value is at $k = \arccos(-1 + 1/2S^2)$. The fact that the overlap matrix needs to be positive defined prevents any real bands divergence (middle panel). However such divergence may appear for complex k.

The major point of concern is the fact that this result is against point 1. of the list introduced above and describing the analytical properties of $E(\mathbf{k})$. It is not true that the dispersion relation is analytic for each point in k-space except the branch points. There is the possibility of a singularity that is a pole.

Looking back at the derivation of Kreigher[158] , we have found that an important point have not been stressed. In Eq. (5.30), $E_i(\mathbf{k})$ are analytic because the $\gamma_i(\mathbf{k})$ are polynomial functions, but also because the coefficient of the leading term, E^N, is one. This is an essential condition. In fact, if the contrary happens, $E_i(\mathbf{k})$ might only be identified as meromorphic functions. An easy example is sufficient to prove the statement. Supposing to have two polynomial functions $a(k)$ and $b(k)$, an equation of the kind $a(k) = \varepsilon b(k)$ leads to a $\varepsilon(k)$ with a pole at q anytime $b(k)$ has a zero at q.

The Schrödinger equation (5.29) in the case of non-orthogonal basis is a generalized eigenvalue problem:

$$\overline{H}(\mathbf{k})\mathbf{c} = E\overline{S}(\mathbf{k})\mathbf{c}. \tag{5.42}$$

The characteristic polynomial is now the determinant of $(\overline{H} - E\overline{S})$, which leads to an equation similar to Eq. (5.30), where the coefficient of the leading power is not one. If this coefficient is $\mathbf{k}$ dependent and has a zero, there will be a probability that the $E_i(\mathbf{k})$ have a pole. We have then proved that the presence of flat bands is admissible in presence of an overlap matrix.

The flat bands are a problematic feature of our implementation of the TMM as a post-processing tool of SIESTA. In fact, they put an obstacle to the automatization of the CBS calculations. We have demonstrated that the flat bands arise in presence of an overlap

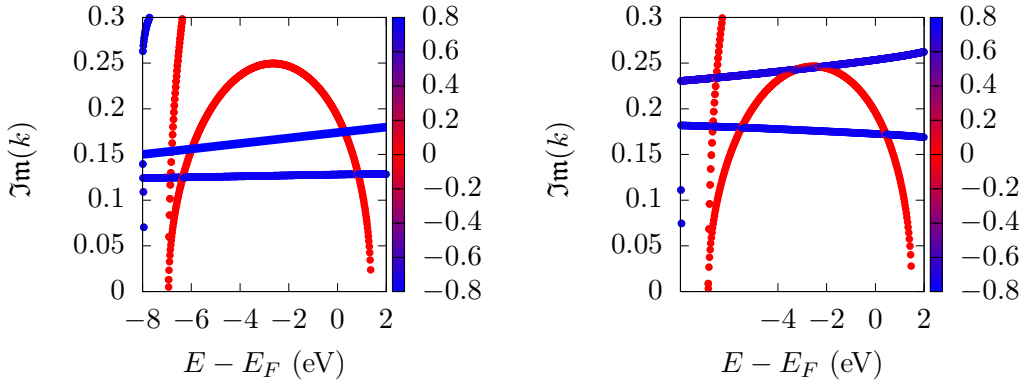


Figure 5.12: CBS obtained with the TMM for MgO along the [111] direction at $\mathbf{k}_{\parallel} = (0, 0)$. On the left-hand side panel we present the case deriving from a DZDP SIESTA calculation. On the right-hand side panel the same calculation is reported but employing a 5% shorter basis set cutoff. The y axis report the imaginary part of the complex band. Color coded is the real part. The k vector is in units of 1/Bohr.

between elements of the basis. In SIESTA, the overlap can be controlled by changing the cutoff of the basis set. We indeed see a variation of the flat bands according to the change in the basis set cutoff. For instance, in Figure 5.12 we report of two calculations of the CBS for MgO in direction [111] and $\mathbf{k}_{\parallel} = (0, 0)$. The left panel shows the situation where the DZDP basis choice is employed, the right panel is the same calculation when the cutoff of the basis is shortened by 5%. The shorter cutoff leads to a smaller overlap and, therefore, the flat bands move up in energy. However, the reduction of the basis cutoff is done at the expenses of the accuracy of the calculation. Hence, it is not a strategy we want to pursue.

We present now a practical procedure aimed to ignore the flat bands. Going back to the properties of the CBS, we recall that all the lines of real energy are connected to the real plane. Each line of real energy keeps jumping on and off the real plane. Also lines coming from $E = -\infty$ reconnect to the real plane and lines going to $E = +\infty$ starts from the real plan too. The TMM allows to obtain the allowed k vectors for each energy, E. Practically, E is a parameter that is increased by discrete values in a range [E_{min},E_{max}]. The strategy is, for each energy, to look if there is a band leaving the real axis ($\Im\mathfrak{m}(k)$ becoming non zero) and to follow that band based on eigenvector similarity. The fact that lines of real energy do not terminate or split, ensures that it is possible to follow the band until it goes back to the real axis or it goes to infinity. In practice, this is done spanning [E_{min},E_{max}] two times. Starting from E_{min}, E is increased and any time $\Im\mathfrak{m}(k)$ passes from $\Im\mathfrak{m}(k) = 0$ to $\Im\mathfrak{m}(k) \neq 0$ a new band is detected and followed. This procedure it is not able to detect lines coming from $E = -\infty$ and reconnecting to the real plane. That is the reason why the procedure is repeated starting from E_{max} and decreasing the energy.

Thanks to the explained procedure, pictures of the kind in Figure 5.13 can be produced. The figure presents in gray the results obtained directly from the TMM method, where many flat bands appear. The black dots are instead obtained following the procedure explained in the previous paragraph. All the bands leaving the axis $\Im\mathfrak{m}(k) = 0$ are detected and tracked. The algorithm proved to be reliable in most cases. Only few exceptions are

observed and they are explained with an example in the next section.

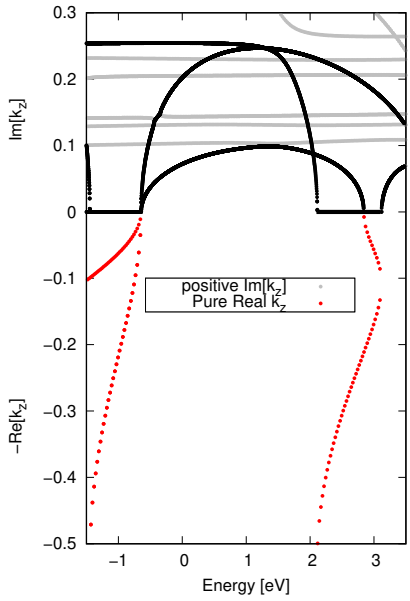


Figure 5.13: CBS of AlP along [111] direction at $\mathbf{k}_{\parallel} = (0, 0)$. Gray dots are obtained with the TMM without any correction. In black we present the result obtained applying the algorithm aimed to ignore the flat bands.

5.4 Complex bands of alkaline oxides/sulphides

In this section we present the results of CBS calculations for selected alkaline oxides and sulphides. These materials are expected to have similar properties of MgO, the prototypical insulator employed in the STT-MRAM technology. We first report a series of results concerning the compound BeS. This specific case aims to show in details the several information we want to gather in this work. Subsequently a table summarizing the results for all the other compounds is presented and the importance of the CBS for STT-MRAMs is recalled.

The electronic structure of BeS is calculated with the code SIESTA employing the pseudopotentials obtained in Section 3.1 and using DZDP as a basis set. We found the use of the DZP basis set insufficient for the compound under scrutiny because for Be, only s orbitals are occupied and this means that only s orbitals have double radial part in the DZP case. With DZDP basis, instead, two radial ζs are associated to p orbitals as well, with significant improvement in the accuracy. The choice of DZDP basis set is employed also for the rest of the calculations presented in this section. The mesh cutoff value is set to 900 Ry and the k-point mesh is 30×30×30. The crystal structure analyzed is the experimental structure obtained from the ICSD database [135] with identifier ICSD_52719. It belongs to the cubic rock-salt type structures (space group 225). According to the result of our calculation, BeS is a semiconductor with indirect Kohn-Sham (KS) gap of 0.5 eV. The minimum of the conduction band is at the high-symmetry point X, while the maximum of the valence is between Γ and K, for this reason we call it a ΓK-X gap.

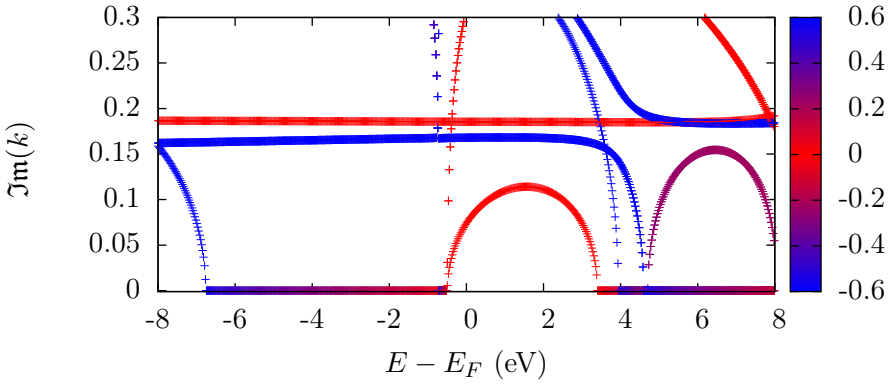


Figure 5.14: CBS of BeS along the direction [111] with parallel component of the wave vector $\mathbf{k}_{\parallel} = (0,0)$. The k are in units of 1/Bohr. On the y axis the $\mathfrak{Im}(k)$ is reported, color coded is the $\mathfrak{Re}(k)$.

Figure 5.14 presents the calculation for the CBS of BeS along the direction [111] with parallel component of the wave vector $\mathbf{k}_{\parallel} = (0,0)$. The picture allows to discuss one problematic situation we rarely encounter in our calculations. By looking at the blue band leaving the energy at $E - E_F \sim 4.5$ eV, it can be seen how it very soon turns into a flat band. This is an effect due to the interaction with another band of the same color, that is coming from energies at $-\infty$. In reality, in fact, the band we are analyzing should behave similarly to the band of the same color leaving the energy axis at $E - E_F \sim 4$ eV, namely going up almost straight in $\mathfrak{Im}(k)$. This band should cross the flat band coming from $+\infty$, not interact with it. We are unsure about the reason of this interaction. The described situations leads to problems in the applicability of the algorithm aimed to ignore the flat bands and explained at the end of Section 5.3. However, most of the times, we are interested only in energies in the close vicinity of the Fermi energy. To limit the energy interval under investigation to $[E_v - \delta E, E_c + \delta E]$ (where E_v is the energy a the top of the valence band, E_c the energy at the bottom of the conduction band and δE an infinitesimal energy) is enough to avoid interferences of the situations above in the judgment of the complex band in the energy gap. The only exception is when a situation of interference happens for the line connecting directly to the valence or conduction band. Moreover, we can recall that Heine proved that the crossing of two bands is vanishing probable outside the real plane. We have no guarantee that the statement still stands for the flat bands, however it seems reasonable. The situation $\mathbf{k}_{\parallel} = (0,0)$ is an high symmetry situation, for points with less symmetry the effect is less probable. This brings us to the next discussion.

In Figure 5.15 we report the calculation for the CBS of BeS along the direction [111] with parallel component of the wave vector $\mathbf{k}_{\parallel} = (0, 0.5\mathbf{q_2})$. Here $\mathbf{q_2}$ is one of the two reciprocal lattice vectors. Few differences respect to the previous picture must be underlined. Firstly, away from $\mathbf{k}_{\parallel} = (0,0)$, the real part of the CBS is not symmetric anymore respect to the real axis. This is why the color coding in Figure 5.15 is no longer the same for positive and negative $\mathfrak{Re}(k)$. Secondly, we see this time a gap much closer to the actual gap of BeS (0.5 eV). The compound has a ΓK-X indirect gap. The point $\mathbf{k}_{\parallel} = (0, 0.5\mathbf{q_2})$ corresponding to direction [111] (usually called L) is equivalent to X. Finally, we encounter

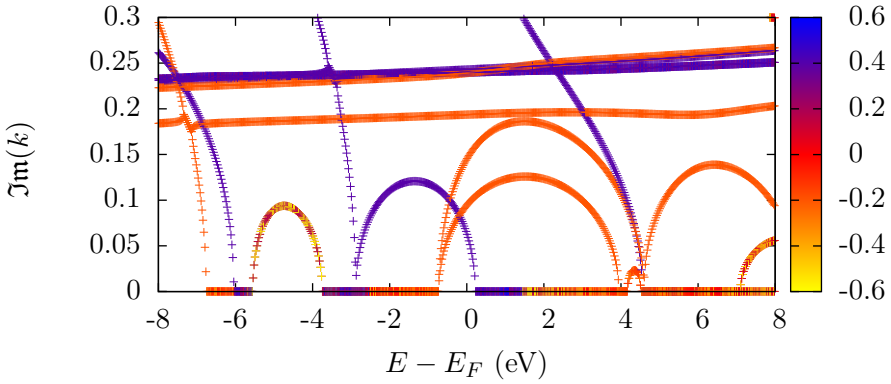


Figure 5.15: CBS of BeS along the direction [111] with parallel component of the wave vector $\mathbf{k}_{\parallel} = (0, 0.5\mathbf{q_2})$. $\mathbf{q_2}$ is one of the two reciprocal vectors of the 2D Brillouin zone perpendicular to the direction [111]. The k are in units of 1/Bohr. On the y axis the $\mathfrak{Im}(k)$ is reported, color coded is the $\mathfrak{Re}(k)$.

for the first time the situation when two different complex bands cross in the middle of the gap. It is important to remind that these two bands do not cross in the complex plane as they have different real part, but it is the imaginary part that matters for the transport. In this peculiar situation, it turns out that the two bands crossing in the gap have the same symmetry, but this might not be always the case. This justify our choice of analyzing the CBS symmetry at three different energy: close to the valence band, close to the conduction band and at the Fermi energy. This will be seen at the end of the section.

Focusing now on the Fermi energy, BeS along [111] is one of the situations when the minimum value of $\mathfrak{Im}(k)$ is not at $\mathbf{k}_{\parallel} = (0, 0)$. In fact, already from the comparison between Figure 5.14 and Figure 5.15, it is clear that at $\mathbf{k}_{\parallel} = (0, 0.5\mathbf{q_2})$ we have a smaller $\mathfrak{Im}(k)$. The general picture can be seen from Figure 5.16. Here the minimum $\mathfrak{Im}(k)$ is reported as a function of $\mathbf{k}_{\parallel}$. The x and y axis of the figure are in units of the module of the reciprocal lattice vectors of the 2D Brillouin zone. In the direction perpendicular to [111], the crystal is periodic and it is possible to define a reciprocal 2D lattice, with reciprocal lattice vectors $\mathbf{q_1}$ and $\mathbf{q_2}$. Figure 5.16 presents the minimum $\mathfrak{Im}(k)$ as a function of $\mathbf{q_1}$ and $\mathbf{q_2}$. The x axis spans the range between -0.5$|\mathbf{q_1}|$ and 0.5$|\mathbf{q_1}|$, the y axis spans the range between -0.5$|\mathbf{q_2}|$ and 0.5$|\mathbf{q_2}|$. In fact, this range delimits the 2D first Brillouin zone. We define M' the point (0.0,0.5 $\mathbf{q_2}$), M the (0.5 $\mathbf{q_1}$,0.0) and R (0.5 $\mathbf{q_1}$,0.5 $\mathbf{q_2}$) as reported in Figure 5.16. M and M' are not in general equivalent, but in case of BeS it seems the case. We notice that the minimum of $\mathfrak{Im}(k)$ is at the edges of the 2D Brillouin zone for the case under investigation.

The discussion so far outlined concerns the situation when the translational symmetry in broken in the direction [111] of the crystal. A similar discussion can be done for any other high-symmetry direction. The results can be very different from direction to direction. Just as an example, we report in Figure 5.17 the calculation at $\mathbf{k}_{\parallel} = (0, 0)$ for the [100] direction of BeS.

The initial idea of the PhD project was to create a fully automatic procedure to obtain, given a crystal lattice, the CBS for all the crystal symmetry directions. This is problematic

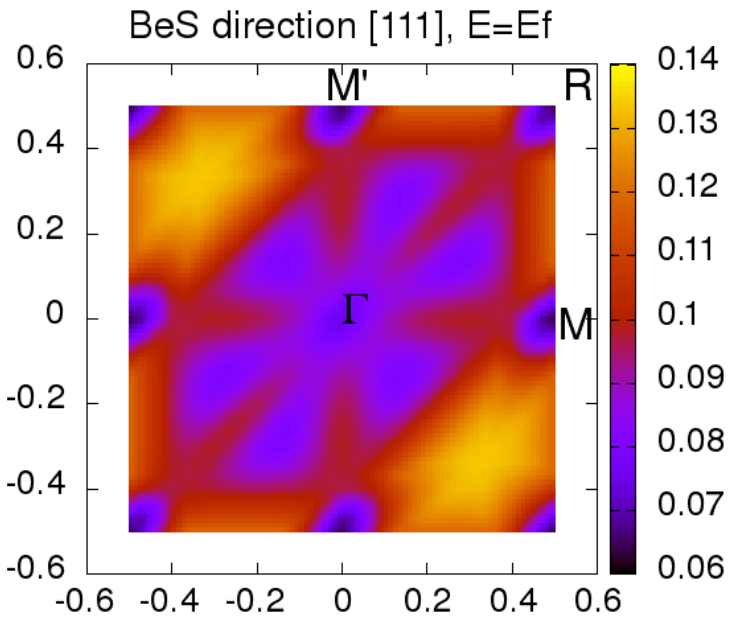


Figure 5.16: Minimum value of $\mathfrak{Im}(k)$ associated to the CBS of BeS along [111] at the Fermi energy. as a function of $\kappa_{\|}$ varying in the 2D Brillouin zone. The x and y axis are in units of the module of the reciprocal lattice vectors $\mathbf{q_1}$ and $\mathbf{q_2}$ respectively. The value of the minimum $\mathfrak{Im}(k)$, color coded, is in units of 1/Bohr.

because of the presence of the flat bands. We have, in fact, a procedure to treat them for one single $\mathbf{k}_{\|}$, but producing plot of the kind in Figure 5.16 is more problematic. In fact, our method to ignore the flat bands is based on the comparison between energy values that are subsequently analyzed from the smallest to the largest (and *vice-versa*). When we face the analysis of the $\mathbf{k}_{\|}$ dependence, instead, we are treating one single energy. However, thanks to a combination of manual check and automatic procedure, we were able to gather information on the CBS along three high-symmetry direction for six compounds. The results are reported in Table 5.1. Each material is calculated for an experimental structure taken from the ICSD database. Firstly, the electronic structure is calculated and the magnitude and nature of the gap is obtained. This information is in the title of each subtable in Table 5.1. All the binaries considered in this study are in cubic structure with space group 225. The CBS is calculated along the high-symmetry directions [100], [110], [111]. Some useful information are then manually extracted. For each element and direction, the point of the 2D Brillouin zone where the $\mathfrak{Im}(k)$ is minimum is identified (second column of each subtable in Table 5.1). For that $\mathbf{k}_{\|}$ the value of $\mathfrak{Im}(k)$ at the Fermi energy is reported (third column of each subtable in Table 5.1). Moreover, it was also possible to identify the symmetry of the complex bands. From the TMM we obtain the eigenvector associated to the complex $\mathbf{k}$. It is of the form reported in Eq. (5.11). The various φ_l composing the eigenvector are projections on layers, meaning they are related

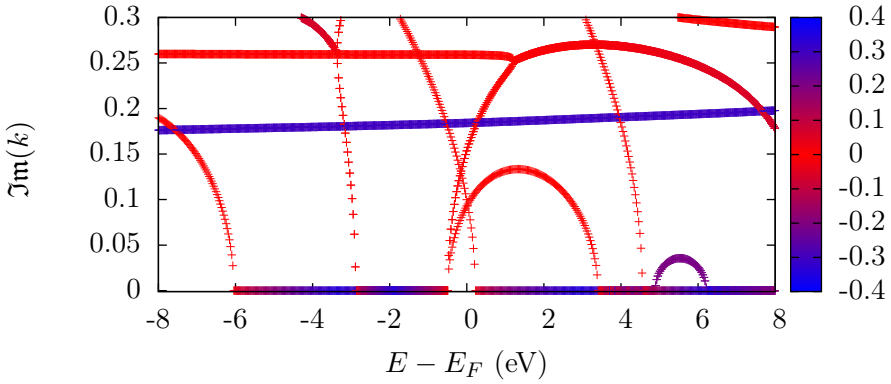


Figure 5.17: CBS of BeS along the direction [100] with parallel component of the wave vector $\mathbf{k}_{\parallel} = (0, 0)$. The k are in units of 1/Bohr. On the y axis the $\mathfrak{Im}(k)$ is reported, color coded is the $\mathfrak{Re}(k)$.

by the identity $\varphi_n = \exp(ika)\varphi_{n-1}$ where $\exp(ika)$ is the corresponding eigenvalue. This information can be used to obtain the projection of the eigenvector on each element of the basis on which the wave function was expanded. Looking at these projections is sufficient for the identification of the symmetry in the simple cases of cubic lattices we are treating. The plan for the future is to formalize this part of the study and achieve a procedure similar to the one explained in Section 4.2 to systematically obtain the band symmetries. The symmetry analysis is performed for the Fermi energy, bur also for an energy close to the valence (fourth column of each subtable in Table 5.1) and an energy close to the conduction band (fifth column of each subtable in Table 5.1). This is because we reckon that it is important two understand if two bands with different symmetries crosses in the gap.

By looking at the results in Table 5.1, we observe that the majority of the compounds under scrutiny presents behavior similar to MgO in terms of symmetry filtering. The only exceptions are BeS and BeO. These two compounds show the minimum of the $\mathfrak{Im}(k)$ at a $\mathbf{k}_{\parallel}$ different from Γ. This is due to the fact that they present an indirect gap. In the case of BeS along [111] we have explicitly shown in Figure 5.14 and Figure 5.15 the comparison between the CBS at Γ and the one at M. At M, the complex band connected to the minimum of the conduction band crosses the Fermi energy, resulting in a smaller $\mathfrak{Im}(k)$ compared to Γ. This is not, however, a general feature for the indirect gap. In fact, MgS has a Γ-X gap as well, but the minimum of $\mathfrak{Im}(k)$ is at Γ. This is because, in the case of MgS, our calculated Fermi energy is very close to the valence band. The CBS at M for MgS is similar to the one of BeS, but the complex band connected to the conduction band crosses the one connected to the valence band at a larger energy compared to E_F. Therefore, in the case of MgS, for the entire 2D Brillouin zone, at the Fermi energy, the minimum $\mathfrak{Im}(k)$ is determined by the band connected to the valence band. By looking at our results, we can say that the indirect gap is a necessary condition to have the minimum of $\mathfrak{Im}(k)$ away from Γ, but it is not sufficient. A second observation relates with the correlation between the value of the minimum $\mathfrak{Im}(k)$ (third column of each subtable in Table 5.1) and the band gap width. Generally, for smaller gaps, the minimum $\mathfrak{Im}(k)$ is smaller. A third feature

Be1O1_163827 E_{gap}=11.25 Ind ΓK-L

111	M/R	0.20 Λ_1	Λ_1	Λ_1
110	R	0.20 Σ_1	Σ_1	Σ_1
100	M	0.22 Δ_1	Δ_1	Δ_1

Be1S1_52719 E_{gap}=0.5 Ind ΓK-X

111	M/R	0.06 Λ_1	Λ_1	Λ_1
110	Γ	0.05 Σ_1	Σ_1	Σ_1
100	R	0.06 Δ_1	Δ_1	Δ_1

Mg1O1_157528 E_{gap}=6.30 Dir Γ

111	Γ	0.20 Λ_1	Λ_1	Λ_1
110	Γ	0.21 Σ_1	Σ_1	Σ_1
100	Γ	0.21 Δ_1	Δ_1	Δ_1

Mg1S1_642787 E_{gap}=2.5 Ind X

111	Γ	0.09 Λ_1	Λ_1	Λ_1
110	Γ	0.10 Σ_1	Σ_1	Σ_1
100	Γ	0.10 Δ_1	Δ_1	Δ_1

Na2O_644917. E_{gap} = 1.43 Dir Γ

111	Γ	0.08 Λ_1	Λ_1	Λ_1
110	Γ	0.09 Σ_1	Σ_1	Σ_1
100	Γ	0.09 Δ_1	Δ_1	Δ_1

Na2S_644962. E_{gap} = 2.34 Dir Γ

111	Γ	0.08 Λ_1	Λ_1	Λ_1
110	Γ	0.08 Σ_1	Σ_1	Σ_1
100	Γ	0.08 Δ_1	Δ_1	Δ_1

Table 5.1: Table reporting a summary of the CBS results for a series of alkaline oxides/sulphides. The title of each subtable reports the material composition, the band gap value and the type of gap. Each subtable reports in the first column the crystal direction, in the second column the point of the 2D Brillouin zone where the minimum $\mathfrak{Im}(k)$ is detected. The third column report the value of such $\mathfrak{Im}(k)$ at the Fermi energy and the corresponding symmetry. The fourth column reports the symmetry of the CBS in the vicinity of the valence band, the fifth column the symmetry of the CBS in the vicinity of the conduction band.

to underline is that, for all the compounds and directions, we report filtering of the bands with spherical symmetry (indicated by the subscript 1 in the symmetry label). This is not surprising as the contribution of s orbitals is expected to be the main contribution to the conduction band for all the materials under investigation. Finally, we can say that the only materials with a band gap sufficiently large for STT-MRAM applications are MgO and BeO. However, as already explained, BeO shows a minimum $\mathfrak{Im}(k)$ away from Γ.

The study is at the moment very limited. This is due to the fact that we still have to systematize the assignment of the correct symmetry to the complex band. As already mentioned, we plan to set up a procedure similar to the one explained in Section 4.2 in order to have a more reliable and systematic way to judge the symmetry of each complex band.

Conclusions

In this PhD thesis we explored the possibility to obtain useful materials information for the improvement of the Spin-Transfer-Torque Magnetic Random Access Memory (STT-MRAM) technology by means of high-throughput Density Functional Theory (DFT) calculations.

The use of high-throughput techniques in computational material science is a fairly new approach and part of the work carried out in this thesis finds its significance in the development of new methodologies for the advancing of the field. In Chapter 3 we discuss two important challenges connected to the task of performing DFT simulations in an high-throughput fashion. The first one is the creation of a set of accurate pseudopotentials and the second is the possibility to correct the DFT gap in a systematic way.

The use of pseudopotentails (PPs) is a common practice in the electronic structure community as it reduces the computational costs of a DFT calculation. Until a few years ago, however, the selection of the PPs for a calculation was not systematic. Only libraries containing untested PPs were available and every researcher had to prove the reliability of the PPs in describing properties of interest before the start of any new project. In recent year we have witnessed, instead, the effort to create a standard protocol to test the PPs accuracy that puts the base for the creation of PP libraries with certified reliability. The use of such libraries is vital for the development of the high-throughput approach as the comparison of calculations performed on large sets of materials is meaningful only if all the calculations have a certain standard of accuracy. In this context, we brought our contribution in creating a computational workflow that is able to generate accurate pseudopotentials in an automatic way. Through the systematic comparison with an all-electron reference, we were able to generate a set of accurate PPs to use for LDA DFT calculations for all the elements of the periodic table between Li and Ge.

The second challenge we faced, associated to high-throughput DFT simulations is the possibility to automatically select the λ to be used in a DFT simulation with Atomic Self Interaction Corrections (ASIC) to describe the correct gap of insulators. Approximate DFT is known to suffer from the self-interaction, the spurious interaction of an electron with the potential generated by itself. ASIC is a DFT formulation aimed to correct this problem. In practice, ASIC reveals itself as a powerful tool to correct the band gap of insulators, a quantity incorrectly described in standard DFT. The ASIC theory, however, requires the introduction of an empirical parameter, λ, that needs to be manually chosen to reproduce the experimental gap. Our idea was to explore, by means of a combination

of experimental data and calculations with the ASIC correction, the possibility to find useful correlations that allow the selection of λ in an automatic way without knowing the experimental gap. The effort was, however, only partially successful. The reason is speculated to be the variability of experimental information in use and a problem in the accuracy of the DFT calculations performed.

Another predominantly methodological part of this thesis was the formalization of a theory and relative implementation for the calculation of the Complex Band Structure (CBS) of a material. We were successful to extend the theory concerning the analytical properties of the CBS to situations where a non-orthogonal basis set in employed for the expansion of the wave function of the system. We used the theory to set up a method able to calculate the CBS as a post process of a DFT calculation employing localized orbitals, known for their computational efficiency. We calculated the CBS for a series of alkaline oxides and alkaline sulfides. This part is described in Chapter 5

The CBS and the symmetry of the complex bands are a quantity of interest for the STT-MRAM technology because it is one of the two ingredients necessary to understand the Tunnel Magneto Resistance (TMR) effect, the principle on which the reading procedure of an STT-MRAM is based on. The second ingredient is the symmetry of the bands around the Fermi energy for a ferromagnet. This was the topic of Chapter 4. We first explained how to rigorously described which quantities are necessary to facilitate the understanding of the symmetry of a band. Secondly, we calculated these quantities for a series of 40 ferromagnets with high Curie temperature. The created database was explored in search for materials with good characteristics for the STT-MRAM technology.

Two are the main scientific outcomes of the study. The first is that all the alkaline oxides and sulphides considered in the study show similar features in terms of symmetry filtering, namely they all select the channel with spherical symmetry. However, only MgO and BeO present a sufficiently large gap for STT-MRAM applications. The difference of BeO with respect to MgO is that it presents a tunneling transport that is predominant at the edges of the 2D Brillouin zone perpendicular to the transport direction. This is true for all the three high-symmetry direction analyzed, namely [111],[110] and [100] of the cubic crystal. The second outcome is the identification of Mn_1Ni_1 as a promising ferromagnet for the STT-MRAM technology. Mn_1Ni_1 is an half metal with Δ_1 symmetry for the states at the Fermi energy along the [100] direction. The latter characteristic allows Mn_1Ni_1 to be paired with a barrier selecting Δ_1 symmetry for an efficient TMR effect.

We hope to make soon publicly available the results of the calculations performed in this thesis as they might be of interest for other applications. For the moments, the softwares developed in this PhD work are available at `https://github.com/bosonie`.

Concerning the future steps of the study, we plan in the following few months to extend the study of the CBS to many more materials. From the theoretical point of view, we still need to formalize a strategy to systematically obtain the symmetry of the complex band. Once this is achieved, the attention can be turned to address technological relevant questions. For instance, it is interesting to look for insulators with cubic symmetry that are able to select the Δ_5 symmetry instead of the Δ_1. The Δ_5 states are degenerate, hence two

channels would be available for the tunneling transport. Moreover, we plane to continue the work on the ASIC and, hopefully, to combine it with the CBS calculations. The correct description of the gap would bring the reliability of the CBS analysis a step further. We do not expect the symmetry of the complex band to be modified by the introduction of the ASIC, but this would ensure a more reliable prediction of the decay rate of the evanescent states. Regarding the part of the work on the ferromagnet selection, we want to expand the study including heavy elements. They were excluded from the study because of the relevant effect the of spin-orbit interaction. The inclusion of spin-orbit in DFT requires the non-collinear treatment for the spin degree of freedom. The theory and tools used in Chapter 4 needs to be reformulated to take into account the complexity deriving from the non collinearity. We also wants to deepen the investigation on Mn_1Ni_1. We want to look for an insulating barrier that could match the lattice constant of Mn_1Ni_1 and perform quantum transport simulations in order to calculate the TMR ratio.

Appendix A

Tunneling transport for a finite potential barrier

In this chapter we summarize the theory of one of the most famous quantum-mechanical exercises, namely the identification of the reflection, R, and the transmission, T, coefficients of an electron tunneling through a potential barrier of thickness a and high V_0. The situation is presented in Figure A.1.

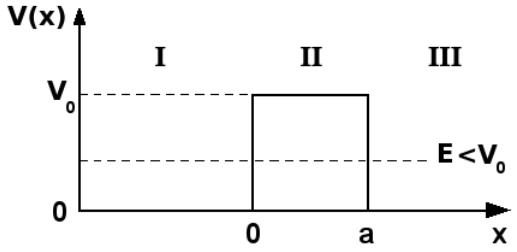


Figure A.1: Potential barrier with high V_0 and thickness a.

We firstly focus on the fully 1D situation. The space can be divided in three regions, indicated in the picture with the symbols **I**, **II** and **III**. Assuming the free electron picture is adequate, the wave function in each portion of the space can be written:

$$\begin{aligned} \varphi_{\mathbf{I}}(x) &= Ae^{ik_1 x} + A'e^{-ik_1 x}, \\ \varphi_{\mathbf{II}}(x) &= Be^{ik_2 x} + B'e^{-ik_2 x}, \\ \varphi_{\mathbf{III}}(x) &= Ce^{ik_1 x} + C'e^{-ik_1 x}, \end{aligned} \tag{A.1}$$

where:

$$k_1 = \frac{1}{\hbar}\sqrt{2mE} \qquad k_2 = \frac{1}{\hbar}\sqrt{2m(E - V_0)} \tag{A.2}$$

As per Figure A.1 we consider energies below the barrier, meaning $\sqrt{E - V_0}$ is a pure imaginary number. This will be used later.

We do not consider the incident particle coming from x = $-\infty$, so $C' = 0$. We define the transmission coefficient $t = C/A$, the reflection one $r = A'/A$. To solve for these two

quantities is equivalent to put $A = 1$ in Eq. A.1.

$$\begin{aligned} \varphi_{\mathbf{I}}(x) &= e^{ik_1 x} + re^{-ik_1 x}, \\ \varphi_{\mathbf{II}}(x) &= Be^{ik_2 x} + B'e^{-ik_2 x}, \\ \varphi_{\mathbf{III}}(x) &= te^{ik_1 x}, \end{aligned} \tag{A.3}$$

The boundary conditions $\varphi_{\mathbf{I}}(0) = \varphi_{\mathbf{II}}(0)$ and $\varphi_{\mathbf{II}}(a) = \varphi_{\mathbf{III}}(a)$ together with the requirements for the corresponding derivatives, $\varphi'_{\mathbf{I}}(0) = \varphi'_{\mathbf{II}}(0)$ and $\varphi'_{\mathbf{II}}(a) = \varphi'_{\mathbf{III}}(a)$ can be used to eliminate B and B' from the equation and lead to:

$$R = |r|^2 = \frac{(k_1^2 - k_2^2)^2 \sin^2(k_2 a)}{4k_1^2 k_2^2 + (k_1^2 - k_2^2)^2 \sin^2(k_2 a)} \tag{A.4}$$

and

$$T = |t|^2 = \frac{4k_1^2 k_2^2}{4k_1^2 k_2^2 + (k_1^2 - k_2^2)^2 \sin^2(k_2 a)}. \tag{A.5}$$

We now go back to the definition of k_2. As already said, it is pure imaginary because $E - V_0 < 0$. It can be written:

$$k_2 = i\kappa \qquad \text{for} \quad \kappa = \frac{1}{\hbar}\sqrt{2m(V_0 - E)} \tag{A.6}$$

κ is now a real number. Considering that $\sinh x = -i \sin(ix)$ (meaning $\sin^2(ix) = -\sinh^2(x)$) and $(ix)^2 = -x^2$, we write:

$$T = |t|^2 = \frac{4k_1^2 \kappa^2}{4k_1^2 \kappa^2 + (k_1^2 + \kappa^2)^2 \sinh^2(\kappa a)}. \tag{A.7}$$

Moreover it can be noticed that $(k_1^2 + \kappa^2)^2$ is a constant equal to $(2mV_0/\hbar^2)^2$. When $\kappa a >> 1$ the relevant term is the $\exp(\kappa a)$ in sinh. Therefore:

$$T = e^{-2\kappa a} \qquad \text{for} \quad \kappa a >> 1. \tag{A.8}$$

We proved that the transmission coefficient goes as $\exp(-2\kappa a)$. This means that when the wave function exits the insulator in a tunneling process, it has a pre-factor $\exp(-2\kappa a)$.

We now turn to the 3D case. The barrier is still along x. The theory is the same as in the 1D case except that the energy is now defined:

$$E = \frac{\hbar^2}{2m}|\mathbf{k}|^2, \qquad \mathbf{k} = (k_x, \mathbf{k}_\parallel). \tag{A.9}$$

The wave vector is in three dimensions and we call $\mathbf{k}_\parallel$ the components perpendicular to the transport direction x. We define $k_\parallel$ its module. So now the definitions in (A.2) are modified to:

$$k_1 = \sqrt{\frac{2mE}{\hbar^2} - k_\parallel^2} \qquad k_2 = \sqrt{\frac{2m(E - V_0)}{\hbar^2} - k_\parallel^2} \tag{A.10}$$

The theory of the 1D case applies exactly as it is. The quantity under the square rood

defining k_2 is still negative, so again we can write:

$$k_2 = i\kappa \qquad \kappa = \sqrt{\frac{2m(V_0 - E)}{\hbar^2} + k_{\|}^2} \tag{A.11}$$

The value of κ defines the decay of the wave function and it clearly assumes the smallest value when $k_{\|} = 0$.

In real situations, a complicated function $f(k_{\|})$ replaces $k_{\|}^2$ and there is no guarantee that the minimum of κ is at $k_{\|} = 0$.

Appendix B

Derivations Chapter 2

B.1 Definition of ρ in second quantization

This section reports a proof of the equality (2.3)= (2.4). The starting point is

$$\langle\Psi|\,\Psi^{\dagger}(\mathbf{r})\Psi(\mathbf{r})\,|\Psi\rangle\,, \tag{B.1}$$

where Ψ is a field operator. Using the definition of field operator we write

$$\langle\Psi|\,\Psi^{\dagger}(\mathbf{r})\Psi(\mathbf{r})\,|\Psi\rangle = \langle\Psi|\sum_{i}^{N}\varphi_i^*(\mathbf{r})\sum_{j}^{N}\varphi_j(\mathbf{r})a_i^{\dagger}a_j\,|\Psi\rangle\,. \tag{B.2}$$

Introducing $N-1$ completeness relations, $\int dr|\mathbf{r}\rangle\langle\mathbf{r}|$, we obtain

$$\begin{aligned}\langle\Psi|\sum_{i}^{N}\varphi_i^*(\mathbf{r})\sum_{j}^{N}\varphi_j(\mathbf{r})a_i^{\dagger}a_j\,|\Psi\rangle = \langle\Psi|\sum_{i,j}^{N}\langle i|\,\mathbf{r}\rangle\langle\mathbf{r}\,|j\rangle\,a_i^{\dagger}a_j\,|\Psi\rangle = \\ \int d^3r_2...\int d^3r_N\,\langle\Psi|\sum_{i,j}^{N}\langle i|\,\mathbf{r}\rangle\langle\mathbf{r}\,|j\rangle\,a_i^{\dagger}a_j\,|\Psi\rangle\,|\mathbf{r_2}\rangle\langle\mathbf{r_2}|...|\mathbf{r_N}\rangle\langle\mathbf{r_N}| = \\ \sum_{i,j}^{N}\langle i|\,j\rangle a_i^{\dagger}a_j\int d^3r_2...\int d^3r_N\Psi^*(\mathbf{r},\mathbf{r_2},...,\mathbf{r_N})\Psi(\mathbf{r},\mathbf{r_2},...,\mathbf{r_N}) =\end{aligned} \tag{B.3}$$

The last step consists in observing that $\sum_{i,j}^{N}\langle i|\,j\rangle a_i^{\dagger}a_j = N$, which lead exactly to the expression in Eq. (2.3).

B.2 Introducing the Bloch theorem for local orbitals

Starting from Bloch theorem, we reach a condition for the coefficients c, the coefficients of the localized orbital expansion. The Bloch theorem can be written as

$$\psi_{\mathbf{k}}(\mathbf{r}+\mathbf{L}) = e^{i\mathbf{k}\mathbf{L}}\psi_{\mathbf{k}}(\mathbf{r}), \tag{B.4}$$

where $\mathbf{L}$ is a vector of the lattice. Expanding the wave function on a localized basis

$$\psi_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{R}}^{cells} \sum_{\nu}^{orb} c_{\nu}^{\mathbf{R}} \phi_{\nu}^{\mathbf{R}}(\mathbf{r}). \tag{B.5}$$

Here $\phi_{\nu}^{\mathbf{R}}(\mathbf{r})$ are the elements of the basis and the sum has been separated in the sum over cells pf the lattice, $\mathbf{R}$, and elements of the basis inside the elementary cell, ν. Using Eq. (B.5) inside Eq. (B.4), we can write:

$$\begin{aligned}
\sum_{\mathbf{R}}^{cells} \sum_{\nu}^{orb} c_{\nu}^{\mathbf{R}} \phi_{\nu}^{\mathbf{R}}(\mathbf{r}+\mathbf{L}) =& e^{i\mathbf{kL}} \sum_{\mathbf{R}}^{cells} \sum_{\nu}^{orb} c_{\nu}^{\mathbf{R}} \phi_{\nu}^{\mathbf{R}}(\mathbf{r}) \\
\sum_{\mathbf{R}}^{cells} \sum_{\nu}^{orb} c_{\nu}^{\mathbf{R}} \phi_{\nu}^{\mathbf{R}-\mathbf{L}}(\mathbf{r}) =& e^{i\mathbf{kL}} \sum_{\mathbf{R}}^{cells} \sum_{\nu}^{orb} c_{\nu}^{\mathbf{R}} \phi_{\nu}^{\mathbf{R}}(\mathbf{r}). \\
\sum_{\mathbf{R}}^{cells} \sum_{\nu}^{orb} c_{\nu}^{\mathbf{R}+\mathbf{L}} \phi_{\nu}^{\mathbf{R}}(\mathbf{r}) =& e^{i\mathbf{kL}} \sum_{\mathbf{R}}^{cells} \sum_{\nu}^{orb} c_{\nu}^{\mathbf{R}} \phi_{\nu}^{\mathbf{R}}(\mathbf{r})
\end{aligned} \tag{B.6}$$

Few points have been used. Firstly, a function centered in $\mathbf{R}$ and evaluated in $\mathbf{r}+\mathbf{L}$ is equivalent to a function centered in $\mathbf{R}-\mathbf{L}$ and evaluated in $\mathbf{r}$. Secondly, $c_{\nu}^{\mathbf{R}} \phi_{\nu}^{\mathbf{R}-\mathbf{L}} = c_{\nu}^{\mathbf{R}+\mathbf{L}} \phi_{\nu}^{\mathbf{R}}$. This is just a change of variable. More explicitly, it would be $\mathbf{R}' = \mathbf{R}-\mathbf{L}$, from which $\mathbf{R} = \mathbf{R}'+\mathbf{L}$. To sum over $\mathbf{R}'+\mathbf{L}$ is equivalent to sum over $\mathbf{R}'$ when the sum extends on the entire space. The sum in Eq. (B.6) is truncated to *cells* for the finite range of the interaction, but it derives from a sum over the entire space. The argument is still valid. Finally $\mathbf{R}'$ is back to be called $\mathbf{R}$.

From Eq. (B.6) we reads $c_{\nu}^{\mathbf{R}+\mathbf{L}} = e^{i\mathbf{kL}} c_{\nu}^{\mathbf{R}}$ for each $\mathbf{R}$ and $\mathbf{L}$, in particular: $c_{\nu}^{\mathbf{R}} = e^{i\mathbf{kR}} c_{\nu}^{\mathbf{0}}$. This concludes the proof.

Appendix C

Band gap experimental data

We report the full list of experimental band gaps utilized in Section 3.2. The data are collected from various references: [123, 124, 125, 126, 127, 128, 129, 131, 132, 133]. For each entry of the list, the corresponding reference is reported in the last column. Together with the compound composition, information about the ICSD identifier (prototype) and the space group of the structure (SG) are reported in the table. The association between the experimental data and the structure has been done *a-posteriori* as explained in Section 3.2. The column "Type" reports the experimental method used to measure the band gap, O means optical (absorption edge, two phonon absorption, photo-conductivity,etc.), T means transport (gap extracted from conductivity, resistivity, hall coefficient dependence on the temperature), U means unknown and it is used for situations where the experimental method was not clear. The column "Temp" reports the temperature, in K, at which the experiment was performed. The experimental gap (exp_gap) is in eV, the experimental density (exp_dens) is in g/cm^3.

prototype	compound	SG	exp_dens	exp_gap	Temp	Type	Ref
Ag1Br1_ICSD_65062	Ag1Br1	225	6.470	2.700	2	O	[124]
Ag1Cl1_ICSD_56540	Ag1Cl1	225	5.560	3.200	2	O	[124]
Ag1F1_ICSD_18008	Ag1F1	225	5.850	2.800	5	O	[123]
Ag1I1_ICSD_161579	Ag1I1	216	5.680	2.820	300	O	[123]
Ag1I1_ICSD_56553	Ag2I2	186	NaN	3.000	2	O	[123]
Ag1I1_ICSD_1899	Ag4I4	186	NaN	3.000	2	O	[123]
Ag1I1_ICSD_52361	Ag4I4	216	5.680	2.820	300	O	[123]
Ag2S1_ICSD_182916	Ag8S4	14	7.234	0.850	300	O	[123]
Al1As1_ICSD_606008	Al1As1	216	3.729	2.153	300	O	[124]
Al1P1_ICSD_52649	Al1P1	216	2.400	2.500	2	O	[123]
Al1Sb1_ICSD_609290	Al1Sb1	216	4.260	1.630	300	O	[124]
Al1N1_ICSD_44107	Al2N2	186	3.255	6.130	300	O	[123]
Al2O3_ICSD_89664	Al4O6	167	3.980	9.900	300	O	[124]
As2Ge1_ICSD_610599	As16Ge8	55	NaN	1.060	300	O	[123]
As4S4_ICSD_185788	As16S16	14	3.506	2.400	300	O	[123]
As1B1_ICSD_184568	As1B1	216	5.220	0.670	300	O	[123]
As1Ga1_ICSD_41674	As1Ga1	216	5.316	1.420	300	O	[124]
As2O3_ICSD_183098	As8O12	227	3.870	4.500	300	O	[124]
As2Se3_ICSD_2600	As8Se12	14	NaN	2.150	10	O	[123]
Au1Cs1_ICSD_58427	Au1Cs1	221	7.065	2.500	300	O	[123]
Au1Rb1_ICSD_58428	Au1Rb1	221	NaN	2.500	300	O	[123]
B1N1_ICSD_77271	B1N1	216	3.480	6.200	300	O	[123]
B1P1_ICSD_602964	B1P1	216	2.970	2.100	300	O	[123]
Ba1F2_ICSD_64717	Ba1F2	225	4.830	9.100	300	O	[124]
Ba1O1_ICSD_26961	Ba1O1	225	5.680	4.400	0	T	[123]
Be1Te1_ICSD_290008	Be1Te1	216	NaN	2.800	300	O	[123]
Be1O1_ICSD_62734	Be2O2	186	3.010	10.580	300	O	[123]
Bi2O3_ICSD_169686	Bi16O24	14	9.370	2.800	300	O	[124]
Bi2Se3_ICSD_20385	Bi2Se3	166	7.680	0.160	77	O	[123]
Bi2Te3_ICSD_158366	Bi2Te3	166	7.860	0.130	300	O	[123]
Bi2O3_ICSD_616890	Bi8O12	14	9.370	2.800	300	O	[124]
Br1Cs1_ICSD_61516	Br1Cs1	225	4.440	6.900	300	O	[124]

Br1Cu1_ICSD_78275	Br1Cu1	216	4.770	3.000	80	O	[124]
Br1K1_ICSD_44282	Br1K1	225	2.750	7.600	300	O	[124]
Br1Li1_ICSD_52236	Br1Li1	225	3.464	7.900	300	O	[124]
Br1Na1_ICSD_61672	Br1Na1	225	3.203	7.500	300	O	[124]
Br1Rb1_ICSD_18017	Br1Rb1	225	3.350	7.250	300	O	[124]
Br1Tl1_ICSD_181756	Br1Tl1	221	7.450	2.650	2	O	[123]
Br1Tl1_ICSD_61519	Br1Tl1	225	7.557	3.100	300	O	[124]
Br2Hg2_ICSD_23721	Br2Hg2	139	7.680	2.600	300	O	[124]
Br2Hg2_ICSD_36196	Br4Hg4	139	7.680	2.600	300	O	[124]
C1Si1_ICSD_181128	C1Si1	216	3.210	2.360	300	O	[126]
Ca1F2_ICSD_60559	Ca1F2	225	3.180	10.000	300	O	[124]
Ca1O1_ICSD_90486	Ca1O1	225	3.300	6.900	75	O	[123]
Cd7P10_ICSD_200596	Cd14P20	43	5.010	1.730	300	O	[123]
Cd1Cl2_ICSD_62202	Cd1Cl2	166	4.047	5.700	300	O	[124]
Cd1F2_ICSD_183499	Cd1F2	225	6.640	6.000	300	O	[124]
Cd1Se1_ICSD_620439	Cd1Se1	216	5.810	1.740	300	O	[123]
Cd1Te1_ICSD_290011	Cd1Te1	216	5.870	1.480	300	O	[123]
Cd1I2_ICSD_38116	Cd2I4	186	5.670	3.900	300	O	[124]
Cd1S1_ICSD_60629	Cd2S2	186	4.820	2.480	300	O	[123]
Cd1Se1_ICSD_41826	Cd2Se2	186	5.470	1.700	300	O	[124]
Cd1I2_ICSD_42198	Cd4I8	186	5.670	3.900	300	O	[124]
Cd1I2_ICSD_37378	Cd8I16	186	5.670	3.900	300	O	[124]
Cl1Cs1_ICSD_52274	Cl1Cs1	225	3.900	7.400	300	O	[124]
Cl1Cu1_ICSD_60711	Cl1Cu1	216	4.140	3.300	80	O	[124]
Cl1K1_ICSD_187219	Cl1K1	225	1.984	8.500	300	O	[124]
Cl1Li1_ICSD_52235	Cl1Li1	225	2.068	9.300	300	O	[124]
Cl1Na1_ICSD_240599	Cl1Na1	225	2.165	9.000	300	O	[124]
Cl1Tl1_ICSD_29107	Cl1Tl1	221	7.010	3.200	4	O	[123]
Cl1Tl1_ICSD_61518	Cl1Tl1	225	7.604	3.600	300	O	[124]
Cl2Hg2_ICSD_65441	Cl2Hg2	139	7.150	3.900	300	O	[124]
Cl3Gd2_ICSD_9320	Cl6Gd4	12	NaN	0.850	300	U	[123]
Cl2Pb1_ICSD_43344	Cl8Pb4	62	5.850	4.860	80	O	[123]
Co1Sb3_ICSD_55452	Co4Sb12	204	NaN	0.050	300	O	[129]
Cr1Si2_ICSD_626776	Cr3Si6	180	NaN	0.160	300	T	[123]
Cs1F1_ICSD_44288	Cs1F1	225	4.638	10.000	80	O	[124]
Cs1I1_ICSD_61517	Cs1I1	225	4.510	6.200	300	O	[124]
Cs1F1_ICSD_53832	Cs4F4	225	4.638	10.000	80	O	[124]
Cu1I1_ICSD_163428	Cu1I1	216	5.680	3.100	80	O	[124]
Cu2O1_ICSD_628619	Cu4O2	224	6.110	2.170	4	O	[123]
Eu1O1_ICSD_631456	Eu1O1	225	NaN	1.120	300	O	[123]
Eu1S1_ICSD_631599	Eu1S1	225	NaN	3.400	300	O	[123]
Eu1Se1_ICSD_631659	Eu1Se1	225	NaN	1.800	300	O	[123]
Eu1Te1_ICSD_33612	Eu1Te1	225	NaN	2.000	300	O	[123]
Eu1H2_ICSD_631320	Eu4H8	62	NaN	1.850	300	O	[123]
F1K1_ICSD_64686	F1K1	225	2.480	10.900	300	O	[124]
F1Li1_ICSD_181799	F1Li1	225	2.635	13.600	300	O	[124]
F1Na1_ICSD_41410	F1Na1	225	2.588	10.500	300	O	[124]
F1Rb1_ICSD_53828	F1Rb1	225	NaN	10.400	80	O	[124]
F2Pb1_ICSD_86738	F2Pb1	225	7.700	5.600	4	O	[123]
F2Sr1_ICSD_262348	F2Sr1	225	4.240	9.400	300	O	[124]
F2Mg1_ICSD_94270	F4Mg2	136	3.180	10.800	300	O	[124]
F2Mn1_ICSD_68736	F4Mn2	136	4.478	10.200	300	O	[124]
Fe1Se2_ICSD_42115	Fe2Se4	58	NaN	0.310	300	T	[123]
Fe1Te2_ICSD_633870	Fe2Te4	58	NaN	0.920	600	T	[123]
Fe1S2_ICSD_633287	Fe4S8	205	5.000	0.950	300	O	[123]
Fe1Se2_ICSD_633475	Fe4Se8	205	NaN	0.200	300	T	[123]
Ga1N1_ICSD_157511	Ga1N1	216	NaN	3.170	300	O	[123]
Ga1P1_ICSD_77088	Ga1P1	216	4.130	2.260	300	O	[124]
Ga1N1_ICSD_67769	Ga2N2	186	6.109	3.500	300	O	[123]
Ga1Se1_ICSD_635372	Ga4Se4	187	5.030	2.000	300	O	[123]
Gd2O3_ICSD_162247	Gd2O3	12	8.330	2.900	933	O	[124]
Gd2O3_ICSD_636103	Gd6O9	12	8.330	2.900	933	O	[124]
Ge1Mg2_ICSD_81735	Ge1Mg2	225	3.090	0.550	300	O	[123]
Ge1Te1_ICSD_43202	Ge1Te1	160	6.190	0.200	300	O	[123]
Ge1O2_ICSD_9162	Ge2O4	136	6.239	5.600	300	O	[124]
Ge1S1_ICSD_637787	Ge4S4	62	4.010	1.650	300	O	[123]
Ge1Se1_ICSD_637854	Ge4Se4	62	5.520	1.075	300	O	[123]
H2Mg1_ICSD_161962	H4Mg2	136	NaN	5.160	300	O	[125]
Hf1S2_ICSD_601164	Hf1S2	164	NaN	1.960	300	O	[123]
Hf1S3_ICSD_638846	Hf2S6	11	NaN	3.100	300	O	[123]
Hf1O2_ICSD_638744	Hf4O8	14	10.110	5.500	300	O	[124]
Hg2I2_ICSD_36189	Hg2I2	139	7.680	2.400	300	O	[124]
Hg2I2_ICSD_36197	Hg4I4	139	7.680	2.400	300	O	[124]
Hg1O1_ICSD_15890	Hg4O4	62	11.080	2.200	300	O	[123]
I3Sb1_ICSD_30906	I12Sb4	14	5.640	1.780	300	O	[123]
I1K1_ICSD_44283	I1K1	225	3.120	6.200	300	O	[124]
I1Li1_ICSD_53820	I1Li1	225	4.076	6.000	300	O	[124]
I1Na1_ICSD_44279	I1Na1	225	3.667	5.900	300	O	[124]
I1Rb1_ICSD_22168	I1Rb1	225	3.550	5.830	300	O	[124]
I1Tl1_ICSD_61520	I1Tl1	221	7.090	2.800	80	O	[123]
I2Pb1_ICSD_24263	I4Pb2	186	6.160	2.550	4	O	[123]

I2Pb1_ICSD_60327	I8Pb4	186	6.160	2.550	4	O	[123]
In1P1_ICSD_53105	In1P1	216	4.800	1.344	300	O	[124]
In1S1_ICSD_660105	In4S4	58	5.180	1.890	300	O	[123]
In1Se1_ICSD_640503	In4Se4	187	5.550	1.170	300	O	[123]
Ir1Se2_ICSD_640985	Ir8Se16	62	NaN	0.170	20-510	T	[133]
K3Sb1_ICSD_246648	K3Sb1	225	NaN	1.400	300	O	[123]
K3Sb1_ICSD_26886	K6Sb2	194	NaN	1.100	300	O	[123]
La2O3_ICSD_56166	La2O3	164	6.574	2.900	530	O	[124]
Lu2O3_ICSD_33659	Lu16O24	206	9.426	3.900	733	O	[124]
Mg1O1_ICSD_157528	Mg1O1	225	3.580	7.800	300	O	[124]
Mg1S1_ICSD_642787	Mg1S1	225	2.860	4.500	77	O	[123]
Mg1Se1_ICSD_159398	Mg1Se1	216	NaN	4.050	300	O	[123]
Mg1Te1_ICSD_159402	Mg1Te1	216	NaN	3.490	300	O	[123]
Mg2Si1_ICSD_163708	Mg2Si1	225	1.880	0.770	0	T	[123]
Mn1S1_ICSD_158652	Mn1S1	225	NaN	2.800	300	O	[123]
Mn1Se1_ICSD_643568	Mn1Se1	225	NaN	2.500	300	O	[123]
Mo1S2_ICSD_43560	Mo1S2	160	NaN	1.900	300	O	[123]
Mo1S2_ICSD_31067	Mo2S4	194	NaN	1.971	77	O	[123]
Mo1Se2_ICSD_644346	Mo2Se4	194	NaN	1.600	300	O	[123]
Mo1Te2_ICSD_24155	Mo2Te4	194	NaN	1.100	300	O	[123]
N5Ta3_ICSD_76460	N10Ta6	63	NaN	2.100	300	O	[132]
N4Si3_ICSD_187736	N8Si6	176	3.240	5.000	300	O	[124]
Na3Sb1_ICSD_26882	Na6Sb2	194	NaN	1.100	300	O	[123]
Ni1S2_ICSD_22369	Ni4S8	205	NaN	0.270	300	O	[123]
O5V2_ICSD_60767	O10V4	59	NaN	2.300	300	O	[128]
O2Te1_ICSD_30222	O16Te8	61	6.000	3.000	300	O	[124]
O1Sr1_ICSD_26960	O1Sr1	225	4.750	5.220	300	O	[123]
O3Sc2_ICSD_41264	O24Sc16	206	NaN	6.300	300	O	[127]
O3Y2_ICSD_26190	O24Y16	206	NaN	5.600	300	O	[124]
O1Pb1_ICSD_62847	O2Pb2	129	9.400	2.100	300	O	[123]
O2Th1_ICSD_77691	O2Th1	225	9.860	5.700	300	O	[124]
O1Zn1_ICSD_67849	O2Zn2	186	5.670	3.440	6	O	[123]
O2Zr1_ICSD_89429	O2Zr1	225	5.640	5.000	300	O	[124]
O1Pb1_ICSD_15402	O4Pb4	57	8.000	2.800	300	O	[124]
O2Ti1_ICSD_33846	O4Ti2	136	4.260	3.500	300	O	[124]
O2Zr1_ICSD_9993	O4Zr2	137	5.820	4.990	300	O	[124]
Os1S2_ICSD_647750	Os4S8	205	NaN	2.000	300	O	[123]
P2Si1_ICSD_43098	P16Si8	55	2.700	1.890	300	O	[123]
P2Zn3_ICSD_26876	P16Zn24	137	4.500	1.350	300	O	[123]
Pb1S1_ICSD_183239	Pb1S1	225	7.500	0.420	300	O	[124]
Pb1Se1_ICSD_187699	Pb1Se1	225	8.100	0.278	300	O	[124]
Pb1Te1_ICSD_600843	Pb1Te1	225	8.200	0.311	300	O	[124]
Pd1S2_ICSD_648753	Pd4S8	61	NaN	0.700	300	T	[123]
Pd1Se2_ICSD_170327	Pd4Se8	61	NaN	0.400	300	T	[123]
Pd1S1_ICSD_648749	Pd8S8	84	6.600	0.500	300	T	[123]
Pt1S2_ICSD_41375	Pt1S2	164	NaN	0.950	300	O	[123]
Pt1S1_ICSD_649541	Pt2S2	131	10.040	0.800	300	O	[123]
Re1S2_ICSD_75459	Re4S8	2	7.506	1.330	300	O	[123]
Re1Se2_ICSD_26256	Re4Se8	2	9.237	1.150	300	O	[123]
Re1S2_ICSD_81814	Re8S16	2	7.506	1.330	300	O	[123]
Rh2S3_ICSD_15344	Rh8S12	60	6.400	0.800	300	T	[123]
Ru2Si3_ICSD_2344	Ru16Si24	60	NaN	0.700	300	T	[123]
Ru1S2_ICSD_604472	Ru4S8	205	NaN	1.800	300	O	[123]
Ru1Se2_ICSD_657508	Ru4Se8	205	NaN	1.000	300	O	[123]
Ru1Te2_ICSD_659137	Ru4Te8	205	NaN	0.250	300	T	[123]
S3Sn2_ICSD_31995	S12Sn8	62	NaN	0.950	300	O	[123]
S1Yb1_ICSD_651441	S1Yb1	225	NaN	0.980	300	O	[123]
S1Zn1_ICSD_651451	S1Zn1	216	4.090	3.700	300	O	[123]
S2Sn1_ICSD_42566	S2Sn1	164	4.500	2.070	300	O	[123]
S2Ti1_ICSD_603788	S2Ti1	164	3.220	0.700	300	O	[123]
S2W1_ICSD_651393	S2W1	160	NaN	2.070	300	O	[123]
S2Zr1_ICSD_651465	S2Zr1	164	3.870	1.700	300	O	[123]
S1Sn1_ICSD_52108	S4Sn4	62	5.080	1.400	300	O	[123]
S3Ti1_ICSD_42072	S6Ti2	11	NaN	0.900	300	O	[123]
S3Zr1_ICSD_651463	S6Zr2	11	NaN	2.800	300	O	[123]
S2Tc1_ICSD_81816	S8Tc4	2	NaN	1.000	NaN	U	[131]
Se1Yb1_ICSD_652203	Se1Yb1	225	NaN	1.500	300	U	[123]
Se1Zn1_ICSD_652223	Se1Zn1	216	5.260	2.800	6	O	[123]
Se2Sn1_ICSD_651910	Se2Sn1	164	6.010	0.970	300	O	[123]
Se2Zr1_ICSD_601542	Se2Zr1	164	NaN	1.200	300	O	[123]
Se1Sn1_ICSD_16997	Se4Sn4	62	NaN	0.900	300	O	[123]
Se2Tl2_ICSD_104189	Se4Tl4	140	8.150	0.730	300	O	[123]
Se2W1_ICSD_84182	Se4W2	194	NaN	1.780	300	O	[123]
Se3Zr1_ICSD_652229	Se6Zr2	11	NaN	1.250	300	O	[123]
Sn1Te1_ICSD_652741	Sn1Te1	225	6.440	0.200	300	O	[123]
Te1Yb1_ICSD_106154	Te1Yb1	225	NaN	1.800	300	U	[123]
Te1Zn1_ICSD_653205	Te1Zn1	216	5.630	2.350	300	O	[123]

Bibliography

[1] D. Weller and A. Moser. Thermal effect limits in ultrahigh-density magnetic recording. *IEEE Transactions on Magnetics*, 35(6):4423–4439, 1999.

[2] Sciencedirect. Superparamagnetism. `https://www.sciencedirect.com/topics/materials-science/superparamagnetism`.

[3] Semiconductor Industry Association (SIA). 2015 International Technology Roadmap for Semiconductors (ITRS). `https://www.semiconductors.org/resources/2015-international-technology-roadmap-for-semiconductors-itrs/`.

[4] D. Ielmini. Resistive switching memories based on metal oxides: mechanisms, reliability and scaling. *Semiconductor Science and Technology*, 31(6):063002, 2016.

[5] M. Wuttig, and N. Yamada. Phase-change materials for rewriteable data storage. *Nature Materials*, 6(11):824–832, 2007.

[6] Wikipedia. 3D XPoint. `https://en.wikipedia.org/wiki/3D_XPoint`.

[7] N. F. Mott. The electrical conductivity of transition metals. *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences*, 153(880):699–717, 1936.

[8] M. N. Baibich, J. M. Broto, A. Fert, F. Nguyen Van Dau, F. Petroff, P. Etienne, G. Creuzet, A. Friederich, and J. Chazelas. Giant Magnetoresistance of (001)Fe/(001)Cr Magnetic Superlattices. *Physical Review Letters*, 61(21):2472–2475, 1988.

[9] G. Binasch, P. Grünberg, F. Saurenbach, and W. Zinn. Enhanced magnetoresistance in layered magnetic structures with antiferromagnetic interlayer exchange. *Physical Review B*, 39(7):4828–4830, 1989.

[10] L. Berger. Emission of spin waves by a magnetic multilayer traversed by a current. *Physical Review B*, 54(13):9353–9358, 1996.

[11] J.C. Slonczewski. Current-driven excitation of magnetic multilayers. *Journal of Magnetism and Magnetic Materials*, 159(1):L1–L7, 1996.

[12] MRAM-Info. Everspin starts sampling 256Mb ST-MRAM chips, plans 1Gb chips by the end of 2016. `https://www.mram-info.com/everspin-starts-sampling-256mb-st-mram-chips-plans-1gb-chips-end-2016`.

[13] MRAM-Info. Stt-mram. `https://www.mram-info.com/stt-mram`.

[14] M. O. A. Ellis, M. Stamenova, and S. Sanvito. Multiscale modeling of current-induced switching in magnetic tunnel junctions using ab-initio spin-transfer torques. *Physical Review B*, 96(22):224410, 2017.

[15] J.M.D. Coey. *Magnetism and Magnetic Materials*. Knovel Library. Cambridge University Press, 2010.

[16] N. Majlis. *The Quantum Theory of Magnetism*. World Scientific, 2007.

[17] R. M. Martin. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, 2004.

[18] S. Yuasa, and D. D. Djayaprawira. Giant tunnel magnetoresistance in magnetic tunnel junctions with a crystalline MgO(001) barrier. *Journal of Physics D: Applied Physics*, 40(21):R337–R354, 2007.

[19] K. Sunaga, M. Tsunoda, K. Komagaki, Y. Uehara, and M. Takahashi. Inverse tunnel magnetoresistance in magnetic tunnel junctions with an Fe4N electrode. *Journal of Applied Physics*, 102(1):013917, 2007.

[20] QuantumWise. Spin Transfer Torque. `https://docs.quantumwise.com/tutorials/spin_transfer_torque/spin_transfer_torque.html`.

[21] D.C. Ralph and M.D. Stiles. Spin transfer torques. *Journal of Magnetism and Magnetic Materials*, 320(7):1190–1216, 2008.

[22] L.D. Landau and E.M. Lifshitz. On the Theory of the Dispersion of Magnetic Permeability in Ferromagnetic Bodies. *Physikalische Zeitschrift der Sowjetunion*, 8(1):153–164, 1935.

[23] T.L. Gilbert. Classics in Magnetics A Phenomenological Theory of Damping in Ferromagnetic Materials. *IEEE Transactions on Magnetics*, 40(6):3443–3449, 2004.

[24] B. Dieny, and N, Locatelli. Introductory course on Magnetic Random Access Memory: lectures presentations. `https://www.inmram.com`.

[25] Wikipedia. Tunnel Magnetoresistance. `https://en.wikipedia.org/wiki/Tunnel_magnetoresistance`.

[26] S. Yuasa, T. Nagahama, A. Fukushima, Y. Suzuki, and K. Ando. Giant room-temperature magnetoresistance in single-crystal Fe/MgO/Fe magnetic tunnel junctions. *Nature Materials*, 3(12):868–871, 2004.

[27] W. H. Butler, X.-G. Zhang, T. C. Schulthess, and J. M. MacLaren. Spin-dependent tunneling conductance of Fe | MgO | Fe sandwiches. *Physical Review B*, 63(5):054416, 2001.

[28] J. Mathon and A. Umerski. Theory of tunneling magnetoresistance of an epitaxial Fe/MgO/Fe(001) junction. *Physical Review B*, 63(22):220403, 2001.

[29] William H Butler. Tunneling magnetoresistance from a symmetry filtering effect. *Science and technology of advanced materials*, 9(1):014106, 2008.

[30] J. Atulasimha, and S. Bandyopadhyay. *Nanomagnetic and Spintronic Devices for Energy-Efficient Memory and Computing*. Wiley Publishing, 1st edition, 2016.

[31] M. D. Stiles, and A. Zangwill. Anatomy of spin-transfer torque. *Physical Review B*, 66(1):014407, 2002.

[32] A. S. Núñez, and A. H. MacDonald. Theory of spin transfer phenomena in magnetic metals and semiconductors. *Solid State Communications*, 139(1):31–34, 2006.

[33] M. D. Stiles. Exchange coupling in magnetic heterostructures. *Physical Review B*, 48(10):7238–7258, 1993.

[34] A. R. Rocha, V. M. García-Suárez, S. Bailey, C. Lambert, J. Ferrer, and S. Sanvito. Spin and molecular electronics in atomically generated orbital landscapes. *Physical Review B*, 73(8):085414, 2006.

[35] C. Heiliger, and M. D. Stiles. Ab-Initio Studies of the Spin-Transfer Torque in Magnetic Tunnel Junctions. *Physical Review Letters*, 100(18):186805, 2008.

[36] S. Iida. The difference between gilbert's and landau-lifshitz's equations. *Journal of Physics and Chemistry of Solids*, 24(5):625–630, 1963.

[37] M. Nakayama, T. Kai, N. Shimomura, M. Amano, E. Kitagawa, T. Nagase, M. Yoshikawa, T. Kishi, S. Ikegawa, and H. Yoda. Spin transfer switching in TbCoFe/CoFeB/MgO/CoFeB/TbCoFe magnetic tunnel junctions with perpendicular magnetic anisotropy. *Journal of Applied Physics*, 103(7):07A710, 2008.

[38] S. Ikeda, K. Miura, H. Yamamoto, K. Mizunuma, H. D. Gan, M. Endo, S. Kanai, J. Hayakawa, F. Matsukura, and H. Ohno. A perpendicular-anisotropy CoFeB–MgO magnetic tunnel junction. *Nature Materials*, 9(9):721–724, 2010.

[39] R. F. L. Evans, W. J. Fan, P. Chureemart, T. A. Ostler, M. O. A. Ellis, and R. W. Chantrell. Atomistic spin model simulations of magnetic nanomaterials. *Journal of Physics: Condensed Matter*, 26(10):103202, 2014.

[40] F. S. M. Guimarães, J. R. Suckert, J. Chico, J. Bouaziz, M. dos Santos Dias, and S. Lounis. Comparative study of methodologies to compute the intrinsic Gilbert damping: interrelations, validity and physical consequences. *Journal of Physics: Condensed Matter*, 31(25):255802, 2019.

[41] G. H. O. Daalderop, P. J. Kelly, and M. F. H. Schuurmans. First-principles calculation of the magnetocrystalline anisotropy energy of iron, cobalt, and nickel. *Phys. Rev. B*, 41:11919–11937, 1990.

[42] S. Lounis, and P. H. Dederichs. Mapping the magnetic exchange interactions from first principles: Anisotropy anomaly and application to fe, ni, and co. *Phys. Rev. B*, 82:180404, 2010.

[43] F. Bonell, T. Hauet, S. Andrieu, F. Bertran, P. Le Fèvre, L. Calmels, A. Tejeda, F. Montaigne, B. Warot-Fonrose, B. Belhadji, A. Nicolaou, and A. Taleb-Ibrahimi. Spin-Polarized Electron Tunneling in bcc FeCo / MgO / FeCo (001) Magnetic Tunnel Junctions. *Physical Review Letters*, 108(17):176602, 2012.

[44] D. D. Djayaprawira, K. Tsunekawa, M. Nagai, H. Maehara, S. Yamagata, N. Watanabe, S. Yuasa, Y. Suzuki, and K. Ando. 230% room-temperature magnetoresistance in CoFeBMgOCoFeB magnetic tunnel junctions. *Applied Physics Letters*, 86(9):092502, 2005.

[45] J. Hayakawa, S. Ikeda, Y.M. Lee, R. Sasaki, T. Meguro, F. Matsukura, H. Takahashi, and H. Ohno. Current-Driven Magnetization Switching in CoFeB/MgO/CoFeB Magnetic Tunnel Junctions. *Japanese Journal of Applied Physics*, 44(41):L1267–L1270, 2005.

[46] S. Mangin, Y. Henry, D. Ravelosona, J. A. Katine, and Eric E. Fullerton. Reducing the critical current for spin-transfer switching of perpendicularly magnetized nanomagnets. *Applied Physics Letters*, 94(1):012502, 2009.

[47] H. Sato, E. C. I. Enobio, M. Yamanouchi, S. Ikeda, S. Fukami, S. Kanai, F. Matsukura, and H. Ohno. Properties of magnetic tunnel junctions with a MgO/CoFeB/Ta/CoFeB/MgO recording structure down to junction diameter of 11nm. *Applied Physics Letters*, 105(6):062403, 2014.

[48] P. Hohenberg and W. Kohn. Inhomogeneous Electron Gas. *Physical Review*, 136(3B):B864–B871, 1964.

[49] W. Kohn and L. J. Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review*, 140(4A):A1133–A1138, 1965.

[50] R.G. Parr and W. Yang. *Density-Functional Theory of Atoms and Molecules*. International Series of Monographs on Chemistry. Oxford University Press, 1994.

[51] J. P. Perdew, K. Burke, and Y. Wang. Generalized gradient approximation for the exchange-correlation hole of a many-electron system. *Physical Review B*, 54(23):16533–16539, 1996.

[52] J. P. Perdew and A. Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Physical Review B*, 23(10):5048–5079, 1981.

[53] U. von Barth and L. Hedin. A local exchange-correlation potential for the spin polarized case. i. *Journal of Physics C: Solid State Physics*, 5(13):1629–1642, 1972.

[54] K. Capelle and G. Vignale. Nonuniqueness of the Potentials of Spin-Density-Functional Theory. *Physical Review Letters*, 86(24):5546–5549, 2001.

[55] N. I. Gidopoulos. Potential in spin-density-functional theory of noncollinear magnetism determined by the many-electron ground state. *Physical Review B*, 75(13):134408, 2007.

[56] J. Kubler, K. H. Hock, J. Sticht, and A. R. Williams. Density functional theory of non-collinear magnetism. *Journal of Physics F: Metal Physics*, 18(3):469–483, 1988.

[57] G.P. Kerker. Non-singular atomic pseudopotentials for solid state applications. *Journal of Physics C: Solid State Physics*, 13(9):L189, 1980.

[58] D. R. Hamann, M. Schlüter, and C. Chiang. Norm-Conserving Pseudopotentials. *Physical Review Letters*, 43(20):1494–1497, 1979.

[59] G. B. Bachelet, D. R. Hamann, and M. Schlüter. Pseudopotentials that work: From H to Pu. *Physical Review B*, 26(8):4199–4228, 1982.

[60] N. Troullier and J. L. Martins. Efficient pseudopotentials for plane-wave calculations. *Physical Review B*, 43(3):1993–2006, 1991.

[61] N. Troullier and J. L. Martins. Efficient pseudopotentials for plane-wave calculations. II. Operators for fast iterative diagonalization. *Physical Review B*, 43(11):8861–8869, 1991.

[62] L. Kleinman and D. M. Bylander. Efficacious Form for Model Pseudopotentials. *Physical Review Letters*, 48(20):1425–1428, 1982.

[63] S. G. Louie, S. Froyen, and M. L. Cohen. Nonlinear ionic pseudopotentials in spin-density-functional calculations. *Physical Review B*, 26(4):1738–1742, 1982.

[64] G. B. Bachelet and M. Schlüter. Relativistic norm-conserving pseudopotentials. *Physical Review B*, 25(4):2103–2108, 1982.

[65] L. Fernández-Seivane, M. A. Oliveira, S. Sanvito, and J. Ferrer. On-site approximation for spin–orbit coupling in linear combination of atomic orbitals density functional methods. *Journal of Physics: Condensed Matter*, 18(34):7999–8013, 2006.

[66] D. Vanderbilt. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Physical Review B*, 41(11):7892–7895, 1990.

[67] P. E. Blöchl. Projector augmented-wave method. *Physical Review B*, 50(24):17953–17979, 1994.

[68] C. Kittel. *Introduction to solid state physics*. Wiley, 2005.

[69] N. W. Ashcroft and N. D. Mermin. *Solid state physics*. Holt, Rinehart and Winston, 1976.

[70] C. M. Goringe, D. R. Bowler, and E. Hernández. Tight-binding modelling of materials. *Reports on Progress in Physics*, 60(12):1447–1512, 1997.

[71] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal. The SIESTA method for ab-initio order-N materials simulation. *Journal of Physics: Condensed Matter*, 14(11):2745–2779, 2002.

[72] O. F. Sankey and D. J. Niklewski. Ab-initio multicenter tight-binding model for molecular-dynamics simulations and other applications in covalent systems. *Physical Review B*, 40(6):3979–3995, 1989.

[73] D. Sánchez-Portal, E. Artacho, and J. M. Soler. Analysis of atomic orbital basis sets from the projection of plane-wave results. *Journal of Physics: Condensed Matter*, 8(21):3859–3880, 1996.

[74] V. Heine. *Group theory in quantum mechanics : an introduction to its present usage.* Dover Publications, 2007.

[75] M. S. Dresselhaus, G. Dresselhaus, and A. (Ado) Jorio. *Group theory : application to the physics of condensed matter.* Springer-Verlag, 2010.

[76] V. C. S. Barrett. *Structure of metals. Crystallographic methods, principles, and data.*, volume 5. McGraw-Hill Publishing Company Ltd., 1954.

[77] O. Kovalev. *Irreducible representations of the space groups.* Gordon and Breach, New York, 1965.

[78] S. Miller. *Tables of irreducible representations of space groups and co-representations of magnetic space groups.* Pruett Press, Boulder Colo., 1967.

[79] Bilbao Crystallographic Server. `http://www.cryst.ehu.es/`.

[80] L. J. Sham and M. Schlüter. Density-functional theory of the band gap. *Physical Review B*, 32(6):3883–3889, 1985.

[81] R.M. Dreizler and E.K.U. Gross. *Density Functional Theory: An Approach to the Quantum Many-Body Problem.* Springer Berlin Heidelberg, 2012.

[82] S. Fabris, S. de Gironcoli, S. Baroni, G. Vicario, and G. Balducci. Taming multiple valency with density functionals:A case study of defective ceria. *Physical Review B*, 71(4):041102, 2005.

[83] P. M. Grant. Electronic properties of rocksalt copper monoxide: a proxy structure for high temperature superconductivity. *Journal of Physics: Conference Series*, 129:012042, 2008.

[84] R. M. Wentzcovitch, W. W. Schulz, and P. B. Allen. VO 2 : Peierls or Mott-Hubbard? A view from band theory. *Physical Review Letters*, 72(21):3389–3392, 1994.

[85] V. I. Anisimov, J. Zaanen, and O. K. Andersen. Band theory and Mott insulators: Hubbard U instead of Stoner I. *Physical Review B*, 44(3):943–954, 1991.

[86] M. Cococcioni and S. de Gironcoli. Linear response approach to the calculation of the effective interaction parameters in the LDA + U method. *Physical Review B*, 71(3):035105, 2005.

[87] B. Himmetoglu, A. Floris, S. de Gironcoli, and M. Cococcioni. Hubbard-corrected DFT energy functionals: The LDA+U description of correlated systems. *International Journal of Quantum Chemistry*, 114(1):14–49, 2014.

[88] A. J. Cohen, P. Mori-Sánchez, and W. Yang. Insights into current limitations of density functional theory. *Science*, 321(5890):792–4, 2008.

[89] R. A. Heaton, J. G. Harrison, and C. C. Lin. Self-interaction correction for density-functional theory of electronic energy bands of solids. *Physical Review B*, 28:5992, 1983.

[90] J. G. Harrison, R. A. Heaton, and C. C. Lin. Self-interaction correction to the local density Hartree-Fock atomic calculations of excited and ground states. *Journal of Physics B: Atomic and Molecular Physics*, 16:2079–2091, 1983.

[91] M. R. Pederson, R. A. Heaton, and C. C. Lin. Local-density Hartree–Fock theory of electronic states of molecules with self-interaction correction. *The Journal of Chemical Physics*, 80(5):1972–1975, 1984.

[92] A. Filippetti and N. A. Spaldin. Self-interaction-corrected pseudopotential scheme for magnetic and strongly-correlated systems. *Physical Review B*, 67(12):125109, 2003.

[93] C. D. Pemmaraju, T. Archer, D. Sánchez-Portal, and S. Sanvito. Atomic-orbital-based approximate self-interaction correction scheme for molecules and solids. *Physical Review B*, 75(4):045101, 2007.

[94] A. Filippetti, C. D. Pemmaraju, S. Sanvito, P. Delugas, D. Puggioni, and V. Fiorentini. Variational pseudo-self-interaction-corrected density functional approach to the ab initio description of correlated solids and molecules. *Phys. Rev. B*, 84:195127, 2011.

[95] Yourgenome. Third generation sequencing. `https://www.yourgenome.org/stories/third-generation-sequencing`.

[96] S. Anthony. The history of supercomputers. `http://www.extremetech.com/extreme/125271-the-history-of-supercomputers`.

[97] S. Curtarolo, G. L. W. Hart, M. Buongiorno-Nardelli, N. Mingo, S. Sanvito, and O. Levy. The high-throughput highway to computational materials design. *Nature Materials*, 12(3):191–201, 2013.

[98] G. Pizzi, A. Cepellotti, R. Sabatini, N. Marzari, and B. Kozinsky. AiiDA: automated interactive infrastructure and database for computational science. *Computational Materials Science*, 111:218–230, 2016.

[99] J. Junquera, Ó. Paz, D. Sánchez-Portal, and E. Artacho. Numerical atomic orbitals for linear-scaling calculations. *Physical Review B*, 64(23):235111, 2001.

[100] K. Lejaeghere, V. Van Speybroeck, G. Van Oost, and S. Cottenier. Error Estimates for Solid-State Density-Functional Theory Predictions: An Overview by Means of the Ground-State Elemental Crystals. *Critical Reviews in Solid State and Materials Sciences*, 39(1):1–24, 2014.

[101] K. Lejaeghere *et al.*. Reproducibility in density functional theory calculations of solids. *Science*, 351:aad3000, 2016.

[102] G. Prandini, A. Marrazzo, I. E. Castelli, N. Mounet, and N. Marzari. Precision and efficiency in solid-state pseudopotential calculations. *npj Computational Materials*, 4(1):72, 2018.

[103] M.J. van Setten, M. Giantomassi, E. Bousquet, M.J. Verstraete, D.R. Hamann, X. Gonze, and G.-M. Rignanese. The PseudoDojo: Training and grading a 85 element optimized norm-conserving pseudopotential table. *Computer Physics Communications*, 226:39–54, 2018.

[104] P. Giannozzi *et al.*. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *Journal of Physics: Condensed Matter*, 21(39):395502, 2009.

[105] M. J.T. Oliveira and F. Nogueira. Generating relativistic pseudo-potentials with explicit incorporation of semi-core states using APE, the Atomic Pseudo-potentials Engine. *Computer Physics Communications*, 178(7):524–534, 2008.

[106] A. García, M. J. Verstraete, Y. Pouillon, and J. Junquera. The psml format and library for norm-conserving pseudopotential data curation and interoperability. *Computer Physics Communications*, 227:51–71, 2018.

[107] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler. Ab initio molecular simulations with numeric atom-centered orbitals. *Computer Physics Communications*, 180(11):2175–2196, 2009.

[108] A. García. Atom. `https://departments.icmab.es/leem/siesta/Pseudopotentials/Code/downloads.html`.

[109] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized Gradient Approximation Made Simple. *Physical Review Letters*, 77(18):3865–3868, 1996.

[110] F. Birch. Finite Elastic Strain of Cubic Crystals. *Physical Review*, 71(11):809–824, 1947.

[111] E. Kucukbenli, M. Monni, B. I. Adetunji, X. Ge, G. A. Adebayo, N. Marzari, S. de Gironcoli, and A. Dal Corso. Projector augmented-wave and all-electron calculations across the periodic table: a comparison of structural and energetic properties. Preprint at arXiv:1404.3015, 2014.

[112] F. Jollet, M. Torrent, and N. Holzwarth. Generation of Projector Augmented-Wave atomic data: A 71 element validated table in the XML format. *Computer Physics Communications*, 185(4):1246–1254, 2014.

[113] J. Lee, A. Seko, K. Shitara, K. Nakayama, and I. Tanaka. Prediction model of band gap for inorganic compounds by combination of density functional theory calculations and machine learning techniques. *Physical Review B*, 93(11):115104, 2016.

[114] J. Liang and X. Zhu. Phillips-Inspired Machine Learning for Band Gap and Exciton Binding Energy Prediction. *The Journal of Physical Chemistry Letters*, 10(18):5640–5646, 2019.

[115] Y. Zhuo, A. Mansouri Tehrani, and J. Brgoch. Predicting the Band Gaps of Inorganic Solids by Machine Learning. *The Journal of Physical Chemistry Letters*, 9(7):1668–1673, 2018.

[116] P. Dey, J. Bible, S. Datta, S. Broderick, J. Jasinski, M. Sunkara, M. Menon, and K. Rajan. Informatics-aided bandgap engineering for solar materials. *Computational Materials Science*, 83:185–195, 2014.

[117] G. Pilania, A. Mannodi-Kanakkithodi, B. P. Uberuaga, R. Ramprasad, J. E. Gubernatis, and T. Lookman. Machine learning bandgaps of double perovskites. *Scientific Reports*, 6(1):19375, 2016.

[118] C. Ricca, I. Timrov, M. Cococcioni, N. Marzari, and U. Aschauer. Self-consistent site-dependent DFT+ U study of stoichiometric and defective SrMnO 3. *Physical Review B*, 99(9):094102, 2019.

[119] K. Tao, J. Zhou, Q. Sun, Q. Wang, V. S. Stepanyuk, and P. Jena. Self-consistent determination of Hubbard U for explaining the anomalous magnetism of the Gd 13 cluster. *Physical Review B*, 89(8):085103, 2014.

[120] T. Hamada and T. Ohno. A new constraint DFT technique for self-consistent determination of U values. *Journal of Physics: Condensed Matter*, 31(6):065501, 2019.

[121] H. J. Kulik and N. Marzari. A self-consistent Hubbard U density-functional theory approach to the addition-elimination reactions of hydrocarbons on bare FeO+. *The Journal of Chemical Physics*, 129(13):134314, 2008.

[122] G. Moynihan, G. Teobaldi, and D. D. O'Regan. A self-consistent ground-state formulation of the first-principles Hubbard U parameter validated on one-electron self-interaction error. Preprint at arXiv:1704.08076, 2017.

[123] O. Madelung. *Semiconductors: Data Handbook*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[124] M. J. Weber. *Handbook of optical materials*. CRC Press, 2003.

[125] P. Vajeeston, P. Ravindran, A. Kjekshus, and H. Fjellvåg. Pressure-Induced Structural Transitions in M g H 2. *Physical Review Letters*, 89(17):175506, 2002.

[126] New Semiconductor Materials Archive. NSM Archive - Silicon Carbide (SiC). `http://www.ioffe.ru/SVA/NSM/Semicond/SiC/`.

[127] F. Ren and J. C. Zolper. *Wide Energy Bandgap Electronic Devices*. WORLD SCIENTIFIC, 2003.

[128] A. Chakrabarti, K. Hermann, R. Druzinic, M. Witko, F. Wagner, and M. Petersen. Geometric and electronic structure of vanadium pentoxide:A density functional bulk and surface study. *Physical Review B*, 59(16):10583–10590, 1999.

[129] E. Z. Kurmaev, A. Moewes, I. R. Shein, L. D. Finkelstein, A. L. Ivanovskii, and H. Anno. Electronic structure and thermoelectric properties of skutterudite compounds. *Journal of Physics: Condensed Matter*, 16(6):979–987, 2004.

[130] A. O'Hara, T. N. Nunley, A. B. Posadas, S. Zollner, and A. A. Demkov. Electronic and optical properties of NbO2. *Journal of Applied Physics*, 116(21):213705, 2014.

[131] Y. Jiao, L. Zhou, F. Ma, G. Gao, L. Kou, J. Bell, S. Sanvito, and A. Du. Predicting Single-Layer Technetium Dichalcogenides (TcX2 , X = S, Se) with Promising Applications in Photovoltaics and Photocatalysis. *ACS Applied Materials & Interfaces*, 8(8):5385–5392, 2016.

[132] C. M. Fang, E. Orhan, G. A. de Wijs, H. T. Hintzen, R. A. de Groot, R. Marchand, J.-Y. Saillard, and G. de With. The electronic structure of tantalum (oxy)nitrides TaON and Ta3N5. *Journal of Materials Chemistry*, 11(4):1248–1252, 2001.

[133] C.H. Liang, Y.H. Chan, K.K. Tiong, Y.S. Huang, D.O. Dumcenco, and P.C. Liao. Synthesis and characterization of needle-like IrSe2 microrods. *Journal of Alloys and Compounds*, 480(1):70–72, 2009.

[134] J. R. Rumble. *CRC handbook of chemistry and physics*. CRC Press, 2019.

[135] FIZ Karlsruhe - Leibniz Institute for Information Infrastructure. Inorganic Crystal Structure Database (ICSD). `https://icsd.products.fiz-karlsruhe.de/`.

[136] ICMAB - Institut de Ciència de Materials de Barcelona. LDA Pseudopotential Database. `https://departments.icmab.es/leem/siesta/Databases/Pseudopotentials/periodictable-lda-abinit.html`.

[137] Abinit. Pseudo potentials for abinit. `https://www.abinit.org/psps_abinit`.

[138] T. Hastie, R. Tibshirani, and J.H. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer series in statistics. Springer, 2009.

[139] scikit-learn: machine learning in Python. `https://scikit-learn.org/stable/`.

[140] F. Di Quarto, C. Sunseri, , S. Piazza, and M. C. Romano. Semiempirical Correlation between Optical Band Gap Values of Oxides and the Difference of Electronegativity of the Elements. Its Importance for a Quantitative Use of Photocurrent Spectroscopy in Corrosion Studies. *J. Phys. Chem.* , 101(14):2519–2525, 1997.

[141] R. S. Mulliken. Electronic Population Analysis on LCAO–MO Molecular Wave Functions. I. *The Journal of Chemical Physics*, 23(10):1833–1840, 1955.

[142] A. Edwards, A. Pineda, P. Schultz, M. Martin, A. Thompson, H. Hjalmarson, and C. Umrigar. Electronic structure of intrinsic defects in crystalline germanium telluride. *Physical Review B*, 73(4):045210, 2006.

[143] R. Rausch and W. Nolting. The Curie temperature of thin ferromagnetic films. *Journal of Physics: Condensed Matter*, 21(37):376002, 2009.

[144] P.J. Jensen, H. Dreyssé, and K.H. Bennemann. Thickness dependence of the magnetization and the Curie temperature of ferromagnetic thin films. *Surface Science*, 269:627–631, 1992.

[145] J. Nelson and S. Sanvito. Predicting the Curie temperature of ferromagnets using machine learning. *Phys. Rev. Materials* 3:104405, 2019.

[146] Aflow - automatic flow for materials discovery. `http://aflowlib.org`.

[147] R. H. Taylor, F. Rose, C. Toher, O. Levy, K. Yang, M. Buongiorno Nardelli, and S. Curtarolo. A RESTful API for exchanging materials data in the AFLOWLIB.org consortium. *Computational Materials Science*, 93:178–192, 2014.

[148] Y. Hinuma, G. Pizzi, Y. Kumagai, F. Oba, and I. Tanaka. Band structure diagram paths based on crystallography. *Computational Materials Science*, 2017.

[149] C. Felser and A. Hirohata. *Heusler Alloys: Properties, Growth, Applications.* Springer Series in Materials Science. Springer International Publishing, 2015.

[150] V. Heine. On the General Theory of Surface States and Scattering of Electrons in Solids. *Proceedings of the Physical Society*, 81(2):300–310, 1963.

[151] W. Kohn. Analytic Properties of Bloch Waves and Wannier Functions. *Physical Review*, 115(4):809–821, 1959.

[152] E. Prodan. Analytic structure of Bloch functions for linear molecular chains. *Physical Review B*, 73(3):035128, 2006.

[153] M. G. Reuter. A unified perspective of complex band structure: interpretations, formulations, and applications. *Journal of Physics: Condensed Matter*, 29(5):053001, 2017.

[154] J. K. Tomfohr and O. F. Sankey. Complex band structure, decay lengths, and Fermi level alignment in simple molecular electronic systems. *Physical Review B*, 65(24):245105, 2002.

[155] Netlib. LAPACK—Linear Algebra PACKage . `http://www.netlib.org/lapack/`.

[156] Paymatgen. `https://pymatgen.org/`.

[157] Y.C. Chang. Complex band structures of zinc-blende materials. *Physical Review B*, 25(2):605–619, 1982.

[158] J. B. Krieger. Some Analytic Properties of Finite-Band Models in Solids. *Physical Review*, 156(3):776–781, apr 1967.

[159] T. Needham. *Visual Complex Analysis*. Clarendon Press, 1998.