TESIS DOCTORAL

Mejora de los modelos generalizados de segundos-principios en sistemas de estado sólido: Avances en modelos electrónicos y acoplamiento electrón-red

PH.D. THESIS

Improvement of generalized second-principles models for solid-state systems: Advances in electron models and electron-lattice coupling

Nayara Carral Sainz

Tesis presentada para la obtención del título de
Doctora en Física
por la Universidad de Cantabria

Dirigida por los Prof. Javier Junquera Quintana
y Prof. Pablo García Fernández

Declaración

D. Javier Junquera Quintana, Catedrático de la Universidad de Cantabria y Pablo García Fernández, Profesor Titular de la Universidad de Cantabria, certifican:

Que la presente Memoria, titulada “Improvement of generalized second-principles models for solid-state systems: Advances in electron models and electron-lattice coupling” ha sido realizada, bajo su dirección, por Nayara Carral Sainz y emiten su conformidad para que dicha memoria sea presentada y tenga lugar, posteriormente, su correspondiente defensa para optar al título de Doctora por la Universidad de Cantabria.

Y para que así conste, se firma en Santander, el de de 2025

Fdo.: Javier Junquera Quintana

Fdo.: Pablo García Fernández

Abstract

In this thesis we present a systematic, quasi-automated methodology for generating electronic models in the framework of second-principles density functional theory (SPDFT). Second-principles electronic models employ an effective multiscale approach for treating the relevant electronic degrees of freedom. In this framework they are represented by a tight-binding model corrected by electron–electron interactions. The Hamiltonian matrix elements are expressed in a basis of Wannier functions which are obtained from the band manifolds of interest in the problem. As a result, the first objective of this thesis was to study the methodology for constructing a Wannier function basis, which provides an exact tight-binding representation of the Hamiltonian matrix elements. This involved a comprehensive examination of both the theoretical foundation and its computational implementation using WANNIER90.

At this point, the main goal of this thesis has been focused on the development of the methodology to fit automatically the parameters involved in the second-principles model considering one-electron Hamiltonians and including properly the electron-lattice coupling and the electron-electron interaction, enabling accurate modeling of structural and electronic responses. The developed approach derives all necessary parameters from first-principles calculations on a carefully designed training set with the objective of maintaining a high level of accuracy and predictive power similar to that of first-principles methods. Notably, this methodology relies entirely on theoretical input, without incorporating any experimental data.

This thesis places particular emphasis on the study and calculation of electron–lattice interaction parameters, aiming to clarify their physical meaning and establish reliable methods for their determination. Electron–vibration interaction plays a crucial role in the understanding of different solid-state properties such a as charge and energy transport in polarons or structural distortions as the Jahn-Teller effect. Alongside a comprehensive review of state-of-the-art approaches to vibronic coupling within the framework of density functional theory, we present a detailed theoretical analysis that connects these methodologies with the electron–lattice parameters used in second-principles models. Special attention is given to recent developments, including finite-difference techniques, reciprocal-space formulations, and Wannier function-based approaches, which serve as a foundation for the accurate parametrization of second-principles models.

Accordingly, the chosen procedure for obtaining the electron–lattice interaction parameters involved their

direct evaluation using the finite-difference method. In contrast, the determination of electron–electron interaction parameters was carried out by fitting these parameters to first-principles calculations through the minimization of a suitably defined goal function. A key feature of our method is the enforcement of space group symmetries, which reduces both the number of independent parameters and the required computational effort.

Finally, we apply the methodology to $SrTiO_3$ and LiF, materials representative of transition-metal perovskites and wide-band-gap insulators, respectively. In both cases, the resulting models reproduce DFT reference data with high fidelity across various atomic configurations and charge states. Our results validate the robustness of the approach and highlight its potential for simulating complex phenomena such as polarons and excitons.

Resumen

En esta tesis se presenta una metodología sistemática y cuasi-automatizada para la generación de modelos electrónicos en el marco de la teoría del funcional de la densidad de segundos-principios (Second-Principles Density Functional Theory, SPDFT). Los modelos electrónicos de segundos-principios emplean un enfoque multiescala efectivo para tratar los grados de libertad electrónicos relevantes, representándolos mediante un modelo de enlace fuerte (*tight-binding*) corregido por interacciones electrón–electrón. Los elementos de matriz del Hamiltoniano se expresan en una base de funciones de Wannier obtenidas a partir de los conjuntos de bandas de interés en el sistema en estudio. Como resultado, el primer objetivo de esta tesis ha sido estudiar la metodología para la construcción de una base de funciones de Wannier, la cual proporciona una representación exacta del Hamiltoniano en el marco del modelo de enlace fuerte. Este análisis implica una revisión exhaustiva tanto de los fundamentos teóricos como de su implementación computacional mediante el código Wannier90.

A partir de este punto, el objetivo principal de la tesis se centra en el desarrollo de una metodología para ajustar automáticamente los parámetros del modelo, considerando Hamiltonianos a un electrón e incorporando adecuadamente tanto la interacción entre electrones y las vibraciones de la red cristalina como los grados de libertad electrón–electrón. Esto permite una descripción precisa de las respuestas estructurales y electrónicas del sistema. La metodología desarrollada permite derivar todos los parámetros necesarios a partir de cálculos *ab initio* realizados sobre un conjunto de entrenamiento cuidadosamente diseñado, con el fin de mantener un alto nivel de precisión y capacidad predictiva comparable con la de los métodos de primeros-principios. Cabe destacar que esta metodología no está basada en datos experimentales.

Esta tesis pone especial énfasis en el estudio y la determinación de los parámetros que capturan la interacción electrón–vibración, con el objetivo de clarificar su significado físico y establecer métodos fiables para su cálculo. La interacción electrón–vibración desempeña un papel crucial en la comprensión de diversas propiedades del estado sólido, tales como el transporte de carga y energía en polarones, o las distorsiones estructurales asociadas al efecto Jahn–Teller. Además de una revisión detallada del estado del arte en el acoplamiento vibriónico en el contexto de la teoría del funcional de la densidad, en esta tesis se presenta un análisis teórico que conecta dichas metodologías con los parámetros de interacción electrón–vibración empleados en los modelos de segundos-principios. Se presta especial atención a desarrollos recientes, incluyendo técnicas de diferencias finitas, formulaciones en el espacio recíproco y

métodos basados en funciones de Wannier, que constituyen la base para una parametrización precisa de los modelos SPDFT.

El procedimiento adoptado para la obtención de los parámetros de interacción electrón–red se ha basado en su evaluación directa mediante el método de diferencias finitas. Por otro lado, los parámetros de interacción electrón–electrón han sido determinados ajustando su valor a cálculos de primeros-principios, mediante la minimización de una función objetivo asociada al error cuadrático. Además, una característica clave del método propuesto es la imposición de las simetrías del grupo espacial del sistema, lo que permite reducir tanto el número de parámetros independientes como el esfuerzo computacional requerido.

Finalmente, se ha aplicado la metodología desarrollada a los materiales $SrTiO_3$ y LiF, representativos de perovskitas con metales de transición y aislantes con gran ancho de banda, respectivamente. En ambos casos, los modelos resultantes reproducen con alta fidelidad los datos de referencia obtenidos por DFT en distintas configuraciones atómicas y estados electrónicos. Nuestros resultados validan la solidez del enfoque propuesto y destacan su potencial para la simulación de fenómenos complejos como polaronas y excitones.

Agradecimientos

Este camino ha sido, sin duda, uno de los más inspiradores, duros e inolvidables de mi vida. He aprendido muchísimo sobre física, pero, sobre todo, he comprendido que el tiempo es un tesoro incalculable. Aquí dejo un agradecimiento sincero a todas las personas que han pasado por mi vida en estos cuatro años, y que de una u otra forma han dejado su huella en este viaje.

En primer lugar, quiero darle las gracias a mis directores de tesis, Javier Junquera y Pablo García. A Javier, gracias por estar siempre atento a cada paso de mi recorrido, por desentrañar conmigo los laberintos de la incipiente vida investigadora y por enseñarme a ser sistemática en la búsqueda de soluciones, como son también ahora los modelos de segundos-principios. Por esas visitas matutinas en las que cada duda encontraba su espacio y, sobre todo, por tu inquebrantable dedicación. A Pablo, gracias por ser compañero en cada batalla que Modelmaker nos ha planteado, por estar siempre dispuesto a resolver mis interminables preguntas y por preocuparte no solo por los problemas de la tesis, sino por cómo calaban en mí. Por recordarme, en los momentos de inseguridad, que la fortaleza no se encuentra, se construye. También quiero agradecer a los profesores Toño y Miguel. Cada encuentro con ellos ha sido un paréntesis para desconectar con sus historias tan llenas de vida. Y, sobre todo, gracias por enseñarme que el sentido crítico es un faro, tanto en la ciencia como en la vida. Je tiens également à remercier le professeur Philippe Ghosez pour m'avoir accueilli avec bienveillance lors de mon bref séjour à Liège, m'avoir présenté à toute son équipe et m'avoir montré l'importance de l'autre facette des méthodes de second principe — les modèles de réseau.

También, quiero darle las gracias a cada una de las personas que forman parte del CITIMAC y a la Universidad de Cantabria por brindarme la oportunidad de haber llevado a cabo esta tesis doctoral con una beca Concepción Arenal. También, agradecer al proyecto PGC2018-096955-B-C41 financiado por MICIU/AEI /10.13039/501100011033 y por FEDER, UE y al proyecto PID2022-139776NB-C63 financiado por MICIU/AEI /10.13039/501100011033 y por FEDER, UE.

En estas líneas, quiero expresar mi enorme agradecimiento a una de las personas más especiales que estos años de tesis me han regalado: Inés, mi “experta en efecto Jahn-Teller”, mi inseparable compañera de “misiones” y una gran confidente de la crónica social. Aquí, le quiero agradecer su infinita paciencia, su generosidad sin límites y cada instante compartido mientras librábamos juntas la batalla contra la tesis. Sin duda, ella ocupa un asiento privilegiado en el primer vagón del tren del que soy la locomotora (sé que ella comprenderá el significado). Gracias a Toraya, por ser mi compañera de navegación en ese mar de desesperación y obstáculos que la tesis nos impuso, y por ser quien mejor ha comprendido cada una de mis frustraciones. Agradezco a Fernando su generosidad constante, su amabilidad genuina y la disposición siempre positiva que ha mostrado cada vez que he recurrido a su ayuda. También, gracias a mis dos primeros compañeros de despacho, Raúl y Guille, que aunque su estancia fue breve, estuvo llena de buenos momentos como los "operativos antiarácnidos".

Sin duda, una de las mejores pausas del día ha sido compartir conversaciones con mis queridos "compañeros del café". Cada encuentro alrededor de sus tazas y mi fruta era un soplo de aire fresco, una ventana abierta a las risas y el desahogo. En especial, gracias a David, Pedro y Pablo, por su apoyo incondicional y constante. A medida que avanzaban las jornadas, rodeada de grandes personas, los momentos de comida se han convertido en un descanso muy necesario. Palmerina, única, con sus historias y su caudal de cultura, ha logrado alejarme por un rato de estas páginas que ahora lleno. Junto a ella, Miguel ha sido la voz que traía cordura, la bondad sincera y grandes instantes tan divertidos. También, gracias a Marina, por su sonrisa siempre dispuesta, y ser un ejemplo de perseverancia en la ciencia. En especial, quiero agradecer a Begoña, por teñir nuestras mañanas con sus historias, sembrar enseñanzas entre palabras, regalarnos anécdotas y, sobre todo, por su valiosa compañía. Thank you so much to Simone, Esther, Steve, Paul, Bill, Farouk, and Emmanuel. I thoroughly enjoyed learning about the incredible science you're doing at your universities, but above all, I loved listening about your cultures, traditions, and company. For me, it has been an unforgettable experience. No quiero olvidar, a mis compañeros de máster: a Miguel, María, Jesús, Nuria, Álvaro, Aitor, Ezequiel, Antonio ... Gracias por tantos buenos momentos.

Y, más allá de las paredes de la universidad, está mi círculo cercano, ese refugio que me ha sostenido cuando la soledad frente a esta pantalla se hacía, a ratos, más pesada. A Carmen, mi compañera incondicional, gracias por convertir los fines de semana de estudio en momentos de risas y por estar siempre dispuesta a caminar a mi lado en este sendero tan arduo. Por llenar cada espacio con tu esencia única, por brindarme tu ayuda sin reservas, por ser tú, auténtica. Tu compañía y tu escucha han sido un ancla imprescindible para que estas horas interminables en la biblioteca fueran mucho menos pesadas. A Celia, por estar siempre pendiente, por ser ese oído atento en los momentos de desahogo y por ofrecerme la mirada más objetiva cuando más la necesitaba. Gracias por tu presencia constante y tu comprensión. A Fer, por escuchar cada uno de mis problemas con la tesis y en general. A Raquel, que pese a la distancia, siempre ha sabido acortar kilómetros con cada reencuentro. Porque cada vez que volvemos a vernos, es como si nunca hubiésemos dejado de hablar. Gracias por estar presente desde lejos, por escucharme con paciencia infinita y por entender mis dificultades con esa empatía tan tuya. Gracias a Georgi, por ayudarme a entender lo que realmente importa en la vida. Gracias al grupo de rutas y a mis compañeros y compañeras de baile, por su comprensión y por todos los momentos divertidos que me han regalado. Después de este parón, confío en que seguirán repitiéndose.

Y en general, quiero dar las gracias a todas las personas que han sido parte de este recorrido y, especialmente, en los últimos meses. A quienes, con su dedicación permanente, han estado ahí, interesándose siempre por mis progresos, preguntándome un millón de veces: ¿qué tal la tesis?, regalándome su compañía incondicional y brindándome su apoyo constante.

Y, finalmente, gracias a mi familia. Gracias a mi madre, por su infinita paciencia en estos meses, por estar siempre ahí, tendiéndome su mano en cada obstáculo. Por darme alas para elegir libremente y por enseñarme que, aunque a veces toque nadar contra la corriente, siempre se puede alcanzar la cima. Gracias por valorarme tanto. Y a Thalía, mi hermana, mi cómplice, mi mejor aliada. Sin duda, la persona más importante de mi vida. A pesar de que en la última etapa de la tesis mi carácter haya sido más “insoportable”, como dices tú, sé cuánto has hecho por intentar entenderme y escucharme. Eres la persona con la que más me gusta estar, la que siempre busco, aunque a veces me vuelva un poco pesada. Estos últimos meses, compartir espacio de trabajo ha sido un regalo inesperado. Las mañanas y los regresos a casa se han convertido en momentos llenos de charlas, bromas y complicidad. Gracias por ser siempre ese refugio al que acudir. Gracias a Juan, o como cariñosamente le llamo, “Juanito”. Con su presencia, ha llenado mis domingos de muchísimos momentos divertidos y alguna que otra preocupación con sus bromas. Aunque a veces me he quejado, sin duda ha convertido los domingos en los días más esperados de la semana. Y, por supuesto, gracias al resto de mi familia: a mis tíos y a mis primos, por su cariño constante, por hacerme sentir valorada y por saber comprenderme siempre, sin necesidad de palabras.

Mi dedicatoria

A las dos personas que sostienen mi vida, quienes están por encima de todo y por debajo de nada:
mi madre y mi hermana.

Glossary

The acronyms used throughout this thesis are presented below in alphabetical order:

AFD	Antiferrodistortive
AFE	Antiferrolectric
APES	Adiabatic Potential Energy Surface
ASR	Acoustic Sum Rule
BO	Born-Oppenheimer
BZ	Brillouin Zone
DFT	Density Functional Theory
DFPT	Density Functional Perturbation Theory
FE	Ferroelectric
GF	Goal Function
JT	Jahn-Teller
ML	Machine Learning
MLWF	Maximally localized Wannier functions
MO	Molecular Orbital
PES	Potential Energy Surface
PJT	Pseudo Jahn-Teller
RAG	Reference Atomic Geometry
SAT	Symmetry Adapted Terms
SP	Second-principles
SPDFT	Second-principles Density Functional Theory
STO	$SrTiO_3$
TB	Tight Binding
TS	Training Set
$\widetilde{H}_{mn\vec{k}}$	Hamiltonian matrix elements expressed in a basis of smooth Bloch-like functions $\{\lvert\widetilde{\psi}_{n\vec{k}_j}\rangle\}$ obtained through a unitary transformation of the original *ab initio* Bloch eigenstates, $\{\lvert\psi_{n\vec{k}_j}\rangle\}$, eliminating their nonanalytic behavior.
T_C	Curie temperature

Contents

List of Publications

The following is a list of publications produced during the course of this doctoral research. Notably, the thesis is primarily based on the third publication.

- *Compatibility of DFT+U with non-collinear magnetism and spin-orbit coupling within a framework of numerical atomic orbitals.*
  Fernando Gómez-Ortiz, Nayara Carral-Sainz, James Sifuna, Virginia Monteseguro, Ramón Cuadrado, Pablo García-Fernández, Javier Junquera.
  Comput. Phys. Commun. **286**, 108684 (2023).

- *Structural and electronic properties of bulk Li_2O_2: first-principles simulations based on numerical atomic orbitals.*
  Paul M Masanja, Toraya Fernández-Ruiz, Esther J Tarimo, Nayara Carral-Sainz, P V Kanaka Rao, Vijay Singh, Bernard Mwankemwa, Juan María García-Lastra, Pablo García-Fernández and Javier Junquera.
  J. Phys.: Condens. Matter **37**, 165502 (2025)

- *Systematic generation of electron models for second-principles density functional theory methods.*
  Nayara Carral-Sainz, Toraya Fernández-Ruiz, Jorge Íniguez-González, Javier Junquera, and Pablo García-Fernández.

Motivation, most relevant questions and main results

Advancements in materials science research have catalysed the development and innovation of technological devices. Whether it's a positron emission tomography machine to detect a tumor in a hospital or the large hadron collider at CERN, modern equipment exists because material properties are pushed to their limits. Looking into the very near future, the advent of artificial intelligence (AI), the extensive use of big-data analysis, and the increasing connectivity of electronic devices in the so-called Internet of Things (IoT) impose an ever-growing demand for energy and natural resources. For instance, training a single AI model can consume as much energy as five cars in their lifetimes [1]. Cooling large data centers is also a significant energy drain, often accounting for 40% of the total energy consumption of these facilities [2]. Another example is the emergence of autonomous vehicles, which will require orders of magnitude higher levels of computing with sustainable power consumption. This trend of designing new electronic devices with the ability to "learn" and improve without human intervention necessitates novel and smart microelectronic components that push the limits of existing materials. Just to mention a few of the requirements that the new materials should gather: low energy consumption, increased storage density for the myriad of produced data and, ideally, energy harvesting capabilities.

Moreover, the rapid advancement of technological capabilities is driving an unprecedented exploitation of natural resources. Humanity now faces challenges never before encountered, many of which are highlighted in the Sustainable Development Goals [3]. Achieving these goals often hinges on discovering more efficient, eco-friendly materials that maintain or enhance functionality.

However, developing such materials is often a lengthy process that requires substantial effort. The transition from laboratory research to commercial application is typically characterized by trial-and-error experimentation, with successful materials taking an average of 15 to 20 years to reach the market [4].

This is where computational simulations can make a significant impact. Modern computational materials science enables the prediction of many properties before materials are synthesized, resulting in substantial savings in materials optimization and significant economic benefits across several key industries. In this regard, the field of materials science and engineering is undergoing a paradigm shift. The reliance on traditional experimental methods is giving way to an era where computational modeling plays a crucial role in guiding knowledge-based decisions, accelerating development, and significantly reducing costs.

The ability to utilize workflows and high-throughput simulations to pre-screen materials with specific desired properties “in silico”, filtering promising candidates based on existing knowledge, is fueled by the rapid and continuous growth of computational power (hardware) and the significant development of more efficient algorithms (software).

At the foundational level, hardware improvements stem from relentless research and innovation in high-speed, smaller, and more energy-efficient microelectronic devices [5; 6], consistent with Moore’s Law (1965) [7], which predicts that the number of transistors on an integrated circuit doubles approximately every two years. These advanced chips are connected in parallel [8] within powerful supercomputers, capable of delivering performance measured in hundreds of petaflops (1 petaflop = 1.000 teraflops), with leading systems surpassing 250 Pflops [9]. New architectures, such as those based on GPU computing [10], are also emerging. Additionally, the era of quantum computing applied to material science challenges is on the horizon [11].

From the software point of view, today it is possible to accurately describe material properties using methods directly grounded in the fundamental laws of quantum mechanics and electrostatics. While studying complex systems often requires practical approximations, these methods are free from empirically adjustable parameters, which is why they are known as “first-principles” or “*ab initio*” techniques. These methods offer atomic-level resolution, enabling highly precise calculations of a system’s electronic and atomic structures, even in the absence of prior experimental data, making them fully predictive. Additionally, they provide a comprehensive understanding of material properties, fostering knowledge-driven research.

Indeed, first-principles methods are often the ultimate computational tools for the design of functional materials. However, despite their widespread use, two major limitations hinder the practical application of first-principles methods for designing materials-enabled products. First, the computational cost of first-principles simulations remains high, restricting their use to relatively small systems (up to a few thousand atoms for basic properties) and short time scales (up to a fraction of a nanosecond in small systems). Second, first-principles simulations are typically confined to zero or very low temperatures, meaning they cannot model realistic operating conditions, such as temperature, mechanical constraints, or external fields, which are crucial for industrial applications.

To address these limitations, we need to climb the different steps of the multiscale ladder (Fig. 1), both in terms of the spatial and temporal scales that simulation methods can address. To achieve this goal, approximate models have been developed, where the electronic and/or atomic degrees of freedom are further simplified. Where to put the focus (on the lattice or the electronic properties) depend on the specific problem. For instance, several lattice-dynamical properties of materials such as structural and mechanical features, dielectric and piezoelectric responses or lattice thermal transport are dominated by the structural degrees or freedom. Then, the treatment of the electrons in the simulations can be avoided. In order to tackle these problems, different methods have been developed. In particular, it is possible to create effective models that capture the phenomena of interest in a "simple" mathematical form allowing for much faster calculations.

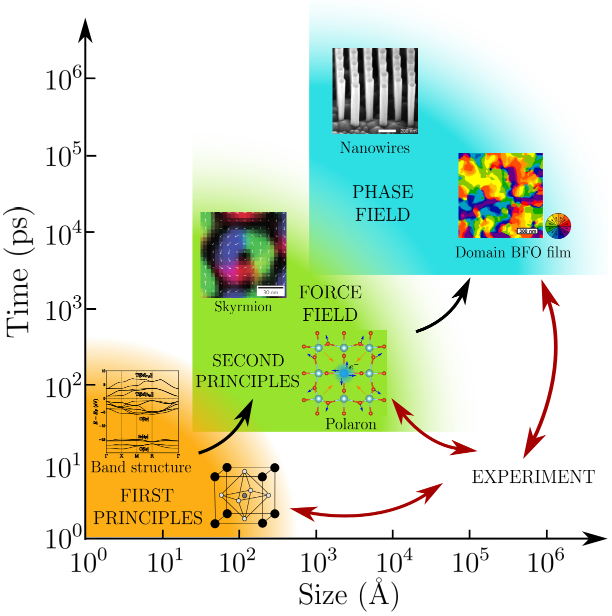


Figure 1: Range of length and time scales accessible through the different theoretical approaches discussed in this introduction. Black arrows highlight how these methods interconnect. First-principles approaches, which capture atomic-level details feed second-principles models that focus on fewer degrees of freedom. Force-field methods model atomic interactions using empirical or semi-empirical potential functions based on classical mechanics. These methods describe systems at the atomistic level, where bonded interactions are modeled similarly to springs. In contrast, phase field models capture mesoscale phenomena, simulating phase transitions and domain dynamics at larger spatial scales without the need to model individual atoms. This makes them particularly suitable for studying microstructure evolution and domain dynamics over extended spatial and temporal scales. The red arrows emphasize the interactions between experimental data and theoretical models across these different scales. Adapted from Ref. [12].

The most frequently used methods have originated within the fields of chemical physics, material science, and biophysics. Atomistic force-field methods utilize empirically determined interatomic potentials to calculate system energy as a function of atomic positions. That is the case of the Lennard-Jones potential [13] that depends only on the magnitude of the separation between a pair of atoms, or the bond order potentials (such as the one proposed by Tersoff [14]) where the potential is also written as a simple pair potential depending on the distance between two atoms, but the strength of this bond is modified by the environment of the atom via the bond order (the strength of a chemical bond depends on the bonding environment, including the number of bonds and possibly also angles and bond lengths). The inclusion of these connection-dependent terms in the force-field description opened the door to treat reactive events with reactive force fields [15]. However, systematically extending these models to enhance precision is generally not well-defined or feasible. Another successful lattice model is the so-called effective Hamiltonian approach, initially proposed by Rabe and Joannopoulos [16] and latter applied to compute the temperature phase diagram of $BaTiO_3$ [17].

The models presented until now describe the Born–Oppenheimer ground energy surface in terms of structural degrees of freedom. However, this could be a constriction in some cases, specially, in those problems where the relevant physics is dominated by the electronic degrees of freedom. Consequently, different methods were also proposed to reintroduce explicitly the treatment of the meaningful electronic degrees of freedom. Between the methodologies that have been proposed, the tight-binding approximation stands out.

In tight-binding (TB) methods [18], electronic interactions are efficiently described using a Hamiltonian expressed in a basis of localized atomic orbitals or Wannier functions. Significant advances have been made in developing these methods. The TB methodology allows for the adaptation of interaction models to suit the specific problem, ranging from effective free-electron descriptions to partially self-consistent solutions, while also incorporating effects such as magnetism, electronic correlations, spin-orbit coupling, and other subtle phenomena. As a result, tight-binding methods have been successfully applied to a wide range of materials, including carbon nanostructures, biological molecules, and simple oxides. Additionally, the TB framework facilitates the calculation of atomic forces, making it a powerful tool for studying structural and lattice-dynamical properties, as well as processes involving bond breaking and formation.

However, these methods face limitations when addressing situations where the key interactions involve very small energy differences, on the order of meV per atom, requiring high precision, or where a fully atomistic description of the material is necessary. Additionally, many approximate approaches struggle to treat both electronic and lattice degrees of freedom with comparable accuracy and completeness. Most existing methods tend to focus predominantly on either electronic properties [19; 20; 21] or lattice properties [16; 22; 23; 24; 25]. Moreover, the few approaches that aim to treat both types of variables simultaneously often rely on highly coarse-grained representations [26; 27; 28].

In the last few years, machine-learning (ML) models in condensed matter physics and chemistry have proliferated [29; 30]. Universal deep neural network approaches to represent the DFT Hamiltonian of crystalline materials, aiming to bypass the computationally demanding self-consistent field iterations of DFT and substantially improve the efficiency of ab initio electronic-structure calculations, are already available [31; 32]. Machine learning and artificial intelligence models have proven to be highly accurate in condensed matter physics, capable of predicting material properties, phase transitions, and electronic structures in large systems with remarkable precision. However, the complexity of these models, often involving deep neural networks or high-dimensional feature spaces, makes them inherently opaque. Unlike traditional physics-based approaches, where analytical equations offer clear insights into underlying mechanisms, ML-driven models function as “black boxes”, providing results without a transparent explanation of the reasoning behind them. This lack of interpretability poses a challenge, as scientists may obtain precise predictions but struggle to extract fundamental physical principles or causal relationships from the model’s inner workings. Moreover, ML-based approaches are often tailored to specific properties (such as energy gap predictions [33]) or materials (for instance, the mechanical properties of zeolites [34]).

Here, we propose a second-principles approach [35], which involves developing effective models using a multiscale perspective. This method, implemented in the SCALE-UP code, directly derives parameters from first-principles data, with the goal of maintaining a high level of accuracy and predictive power similar to that of first-principles methods [36]. This might be a constriction in some cases, specially, in those problems where the relevant physics is dominated by the electronic degrees of freedom. Recently, a method has also been proposed to *reintroduce explicitly the treatment of the meaningful electronic degrees of freedom in the form of a tight-binding model* [37], while avoiding double-counting with the effective atomic potentials. Since there is not input coming from the experiment, our method retains full predictive power, and that is why it is coined as second-principles simulations. In contrast, second-principles density-functional theory offers a physics-driven framework that applies to a broader range of systems and properties, making it a valuable complement to data-driven methodologies. The problem before this thesis was how to fit automatically the parameters in the tight-binding model including properly the interaction of the electronic and the lattice, and the electron-electron degrees of freedom. This has been precisely the topic of this work.

In this thesis we present a systematic, quasi-automated methodology for *generating electronic models* in the framework of second-principles density functional theory. This approach enables the construction of accurate and computationally efficient models by deriving all necessary parameters from first-principles calculations on a carefully designed training set. A key feature of our method is the enforcement of space group symmetries, which reduces both the number of independent parameters and the required computational effort. The formalism includes improved treatments of one-electron Hamiltonians, electron-lattice coupling—through both linear and quadratic terms—and electron-electron interactions, enabling accurate modeling of structural and electronic responses. It has been implemented in a python code, called MODELMAKER that bridges a first-principles code producing the band-structures for different configurations (SIESTA [38]) and a code that computes tight-binding matrix elements in real space in a basis of Wannier functions (WANNIER90 [39; 40]). We apply the methodology to $SrTiO_3$ and LiF, materials representative of transition-metal perovskites and wide-band-gap insulators, respectively. In both cases, the resulting models reproduce DFT reference data with high fidelity across various atomic configurations and charge states. Our results validate the robustness of the approach and highlight its potential for simulating complex phenomena such as polarons and excitons. This work lays the foundation for extending second-principles density functional theory to real-time simulations of optoelectronic properties and further integration with machine-learning methods.

The thesis will be organized as follows. In Chapter 1, we present the theoretical approaches behind the second-principles methods. Specifically, this chapter outlines the tight-binding approach which constitutes the foundation for modeling the electronic degrees of freedom. In addition, it introduces the localized Wannier functions, utilized as the basis set for the electronic part in the second-principles implementation of SCALE-UP. Chapter 2 details the practical procedure for constructing tight-binding-like interactions within a Wannier function basis using the WANNIER90 code. These interactions are subsequently employed to extract the parameters of the second-principles Hamiltonian. In Chapter 3, we provide an overview of the current state of the art in the study of electron-phonon coupling, emphasizing its crucial role in the accurate description of a wide range of physical phenomena. We outline the underlying theoretical framework and the methodologies used to incorporate electron-phonon interactions within first-principles approaches. Furthermore, we place particular emphasis on the implementation of this interaction within the second-principles formalism and its connection to first-principles methods. In Chapter 4, we describe the theoretical framework underlying the electronic component of the second-principles methodology as implemented in SCALE-UP. The procedure for generating the electronic models which feed SCALE-UP calculations is described in Chapter 5. This chapter discusses the different methodologies developed and employed in this work to obtain the model parameters with the MODELMAKER software. In addition, it outlines the key aspects involved in the calculation of these variables, including the construction of training sets and the application of symmetry constraints. Finally, the electronic models constructed using the methodology implemented in MODELMAKER are applied to the $SrTiO_3$ perovskite and the LiF crystal systems. In Chapter 6, we present the results obtained by comparing the band structures and quadratic errors across different atomic configurations and electronic states. These comparisons are made between second-principles calculations (which employ the generated electronic models) and reference results obtained from density functional theory, in order to assess the robustness and accuracy of the proposed approach.

Chapter 1

Computational Methods

1.1 Effective Hamiltonian Approach

One important step forward in the computational study of functional material properties using a coarse-grain for the lattice degrees of freedom comes from the ferroelectric community. Within this area, the method coined as first-principles-based *effective Hamiltonian* (H_{eff}) approach was developed [24]. It can be considered as a generalization of a scheme originally proposed by Rabe and Joannopoulos for GeTe [16].

The development of the effective Hamiltonian relies on the fact that the ferroelectric phase transitions of some typical ferroelectric perovskite oxides (such as $BaTiO_3$) implies *small* atomic displacements and macroscopic strains with respect to a reference prototype structure. This aristotype structure in the perovskites is usually taken as the high-symmetry cubic phase, with symmetry $Pm\bar{3}m$, stable at high-temperature. Moreover, the structural phase transition that yields to the relaxed ferroelectric structure is driven mostly from the *condensation of a single unstable (soft) normal mode* [41], that presents an imaginary frequency in the high-symmetry cubic structure. This soft mode is a local cooperative pattern of atomic distortion associated to each unit cell. The displacements of the atoms can be obtained from the soft-mode eigendisplacement vectors computed in the high-symmetry phase. For instance, the overlap between the total distortion of the tetragonal phase of $BaTiO_3$ in the optimized tetragonal phase and the eigendisplacements associated with the soft-mode of the optimized cubic phase might be as large as 99.3% [42]. Since essentially one mode is involved in the phase transition, and the final displacements are small, then the Born-Oppenheimer energy surface can be parametrized as a low-order Taylor expansion within the restricted subspace defined by the local soft mode (lattice part) and macroscopic strains (elastic part, that is also coupled to the lattice). The method explicitly treats the long-range dipole-dipole interactions and restricts to *local* anharmonic terms. All the parameters, both the ones required for the Taylor expansion and for the electrostatic treatment, are directly fitted to first-principles DFT calculations.

The approximations introduced in the effective Hamiltonian method reduce computational cost on two levels. First, at the level of *identifying the relevant degrees of freedom*, instead of accounting for the $(3 \times N_{\mathrm{at}})$ degrees of freedom per unit cell plus strain, the model focuses on just one soft local mode. This leads to a computational cost reduction by a factor of 2 to 5. Second, at the level of the *simple parametric form of the energy*, rather than performing self-consistent DFT calculations to compute energy and forces for the simulation box, the method evaluates a minimal Taylor expansion of the energy in terms of soft modes and strains. This results in a substantial time savings, reducing computational effort by many orders of magnitude. These models have been coupled to classical [24] or even quantum [43; 44] Monte Carlo or Molecular Dynamic simulations to sample the configurational space at finite temperatures.

Since the first-proposal, the effective Hamiltonian approach proved to be extremely successful, predicting the right sequence of phase transitions in $BaTiO_3$ as a function of temperature (cubic → tetragonal → orthorhombic → rhombohedral) [24; 25]. In practice, the phase transition temperatures T_c are systematically and significantly underestimated using the effective Hamiltonian approach, but good qualitative description of the systems can be achieved when rescaling T_c on experimental data.

As it has been previously cited, the initial simulations were carried out for GeTe [16] and $BaTiO_3$ [24]. But effective Hamiltonians have been fitted to other bulk ferroelectrics, such as $PbTiO_3$ [45], $KNbO_3$ [46], or $NaNbO_3$ [47; 48]. More recently, they have been extended to treat oxygen octahedra rotations, paving the way to deal with antiferrodistortive phase transitions, for instance in quantum paraelectrics like $SrTiO_3$ [17; 43] or in prototypical multiferroics like $BiFeO_3$ [28], where the two primary order parameters (polarization and oxygen octahedra rotations) coexist.

Although the effective Hamiltonian approach is physically-motivated, computationally very efficient, and its precision can be improved, to some extent, in a well-defined way (just increasing the number of terms in the Taylor expansion), their usability is restricted to a very small set of problems. In particular, this approach has not been adopted in other fields beyond the original ferroelectric community. The main reason behind this "failure" is the coarse-graining step involved in the construction of the potential. In general, only one mode is retained in the model. Although this was supported by a plethora of former simulations in the field of the ferroic perovskite oxides, it will be difficult to generalise to other materials. On general grounds, it may be unclear how to choose a subset of the relevant degrees of freedom. In other words, which modes might play an important role in determining the properties of a given material. That is the case, for instance, of Pnma perovskites combining sizable in-phase and out-of-phase oxygen rotations and antipolar motions, or hybrid improper ferroelectrics.

Based on these premises, a more sophisticated approach that retains many of the good features of the effective-Hamiltonian method developed within the ferroelectrics community and avoids its most serious limitations has been developed. In short, this method creates models that describe the energetics of a material by Taylor expanding the potential energy surface around a suitably chosen reference structure as a function of *all the atomic degrees of freedom* [36]. This approach has been implemented in the SCALE-UP [37] and MULTIBINIT [49] codes. One of the most important fingerprint of these methods is that the parameters required for the description of the Born-Oppenheimer energy surface in terms of the structural degrees of freedom are determined directly from first-principles data. Thus, they keep most of the accuracy and predictive power of first-principles approaches. For these reasons, they have been called *second-principles methods* [35].

1.2 Second-principles method for the lattice. The SCALE-UP code.

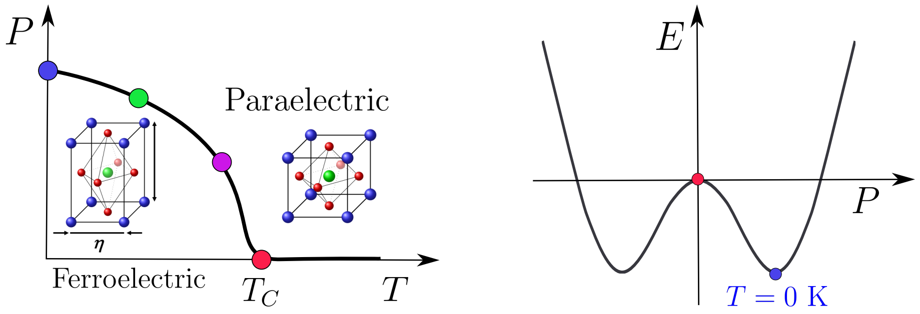
In this section the rudiments of the second-principles SCALE-UP code will be presented. We shall only focus on those aspects that will be relevant to the explanation about how the electrons will be included in the lattice model. A more detailed explanation on all the technicalities behind the technique can be found in Ref. [36].

1.2.1 Reference structure

In order to build a second-principles effective model for a selected system, the first step consists on selecting a reference structure, denominated as Reference Atomic Geometry (RAG). This structure is a particular configuration of the nuclei which will be employed as reference to describe any other configuration of the system. In principle, no restrictions are imposed on the choice of the RAG. However, it is usually convenient to employ the ground state structure or, alternatively, a properly chosen high-symmetry configuration. Note that these choices correspond to critical points (a local minimum, a local maximum, or a saddle point) of the potential energy surface (PES), where all gradients vanish, so that the corresponding forces on the atoms and stresses on the cell are zero. Depending on the purpose for which the model is built, the suitable reference structure might be a local minimum if the aim of the model is focused on studying the anharmonic behavior of a given phase or a local maximum in order to describe phase transitions, competition between phases or domain structures [36; 37].

As an illustrative case, let us consider typical ferroelectric perovskites such as $\mathrm{PbTiO_3}$. The polarization of the system along the (001) direction, which is directly related to the strain, exhibits a temperature-dependent behavior. At $T = 0\,\mathrm{K}$, the system adopts a tetragonal lattice elongated in the (001) direction with atomic displacements leaving the centrosymmetric positions, leading to a ferroelectric phase that corresponds to a minimum in the energy curve (indicated by a blue dot in Fig. 1.1). However, as the temperature increases and reaches the Curie temperature (T_C), the system transitions into a paraelectric state with a cubic phase (marked by a red dot in Fig. 1.1).

According to Fig. 1.1(a), the two most natural structural configurations to consider are the cubic phase, in which the strain is zero, and the ferroelectric phase, which corresponds to the minimum of the energy curve. In these systems, the distortions from the cubic phase to the ferroelectric phase are relatively small. Therefore, by considering the RAG as the cubic phase, a perturbative approach can be applied to analyze the system. The cubic lattice represents a high-symmetry configuration, enabling the generation of other configurations through small lattice distortions. Thus, to facilitate the study of various phases in these perovskites, the cubic phase is the most suitable choice for the RAG.



(a) (b)

Figure 1.1: (a) Polarization as a function of temperature for a prototypical perovskite system. The polarization, which is intrinsically linked to the strain, reveals that for temperatures below the Curie temperature (T_{C}), the system adopts a tetragonal Bravais lattice. This is marked by atomic displacements from their centrosymmetric positions, giving rise to ferroelectric phases. At $T = 0\,\mathrm{K}$, the system attains its energy minimum. Upon reaching T_{C}, the structure transitions to a paraelectric phase characterized by a cubic lattice. (b) Energy as a function of polarization. The system's energy minimum corresponds to the ferroelectric phase, which is reached at $T = 0\mathrm{K}$ and arises from small distortions relative to the cubic structure.

In the second-principles method it is convenient to see a material as the periodic repetition of the unit cell of its reference structure. To describe the atomic and electronic configuration of the system we shall adopt a notation similar to that of Ref. [37]. The magnitudes related to the atomic structure will be labeled by Greek letters. Considering periodic three dimensional infinite crystal, the different cells are denoted by uppercase letters $\{\Lambda, \Delta, \ldots\}$ and the corresponding atoms in the cell as lowercase letters, $\{\lambda, \delta, \ldots\}$. According to this, the lattice vector of the cell Λ is $\vec{R}_\Lambda$. In order to employ a more compact notation, a cell/atom pair shall be represented as a bold lowercase index, i.e., $\vec{R}_\Lambda \lambda \leftrightarrow \boldsymbol{\lambda}$.

The positions in the RAG are defined as

$$\vec{\tau}^{(0)}_{\boldsymbol{\lambda}} = \vec{R}_\Lambda + \vec{\tau}^{(0)}_{\lambda}. \tag{1.1}$$

Any possible crystal configuration can be specified by expressing the atomic positions, $\vec{\tau}_{\boldsymbol{\lambda}}$, as a distortion with respect to the RAG as

$$\vec{\tau}_{\boldsymbol{\lambda}} = \left(\mathbb{1} + \overleftrightarrow{\eta}\right)\left(\vec{R}_\Lambda + \vec{\tau}^{(0)}_{\lambda}\right) + \vec{u}_{\boldsymbol{\lambda}}, \tag{1.2}$$

where $\mathbb{1}$ is the identity matrix, $\overleftrightarrow{\eta}$ is the homogeneous strain tensor, and $\vec{u}_{\boldsymbol{\lambda}}$ is the absolute displacement of atom λ in cell Λ with respect to the strained reference structure. In Eq. (1.2), the strain tensor, denoted by $\overleftrightarrow{\eta}$, refers to the the homogeneous strain tensor since the inhomogeneous strains (or local strains) are described by the $\vec{u}_{\boldsymbol{\lambda}}$ atomic displacements in the proposed methodology. The homogeneous strain tensor $\overleftrightarrow{\eta}$ contains both the symmetric and antisymmetric components. However, for the purposes of the following discussion, it will be restricted to its symmetric part, such that $\eta_{\alpha\beta} = \eta_{\beta\alpha}$. Under this assumption, the Voigt notation for strain, represented as η_a with $a = 1\ldots6$, can be employed, leading to the following transformation,

$$\begin{pmatrix} \eta_{11} & \eta_{12} & \eta_{13} \\ \eta_{21} & \eta_{22} & \eta_{23} \\ \eta_{31} & \eta_{32} & \eta_{33} \end{pmatrix} \rightarrow \begin{pmatrix} \eta_1 & \eta_6 & \eta_5 \\ \eta_6 & \eta_2 & \eta_4 \\ \eta_5 & \eta_4 & \eta_3 \end{pmatrix}. \tag{1.3}$$

This simplifies the notation for use in subsequent equations.

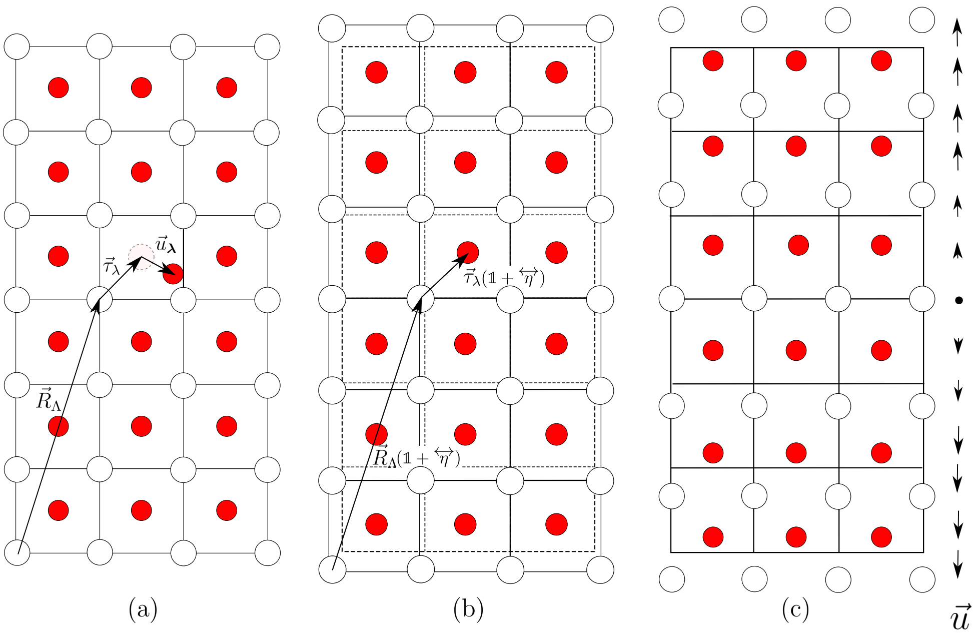


Figure 1.2: Illustration of various distortions experienced by a lattice, described from the reference atomic geometry (RAG). (a) RAG structure with atomic displacements occurring in the absence of homogeneous strain. (b) Homogeneous strain applied without any accompanying atomic displacements. (c) Inhomogeneous (local) strain represented by specific atomic displacement patterns. Figure adapted from Ref. [36].

The various contributions to lattice distortions are illustrated in Fig. 1.2. First, Fig. 1.2(a) depicts the positions of the nuclei in the reference structure, $\vec{R}_\Lambda + \vec{\tau}_\lambda$, along with the atomic displacements, $\vec{u}_\lambda$, relative to the reference atomic geometry (RAG). Figure 1.2(b) illustrates the system's response to a homogeneous strain, $\overleftrightarrow{\eta}$, which results in a uniform deformation of the lattice. Lastly, Fig. 1.2(c) highlights the lattice distortion caused by an inhomogeneous strain. In this case, the effect of local strain is represented through atomic displacements.

With this notation, a given snapshot describing a particular configuration of the system is completely determined by $(\{\vec{u}_\lambda\}, \overleftrightarrow{\eta})$, where the curled brackets represent the collection of the displacements for all the atoms in the system. This section has been written according to Ref. [37].

In Chapter 4, once the electronic part will be introduced, a similar notation will be used. The elements related to electrons will be labeled by latin subscripts. In particular, the Wannier functions will group both the cell and its discrete label in a contracted bold index $\boldsymbol{a} \leftrightarrow \vec{R}_A a$.

1.2.2 Lattice-models

1.2.2.1 Energy contributions

In this subsection, the general scheme for the construction of effective model potential for lattice-dynamical simulations is presented. This part is based on Refs. [36; 50].

The total energy of the second-principles model is represented as E_{eff}, employing the subscript "eff" to differentiate between the energy that the *effective* potential gives for atomic state $(\{\vec{u}_{\boldsymbol{\lambda}}\}, \overleftrightarrow{\eta})$ and the exact energy $E(\{\vec{u}_{\boldsymbol{\lambda}}\}, \overleftrightarrow{\eta})$ that would be obtained from a first-principles simulation of the same configuration. The energy E_{eff} is split in two contributions. The first one represents the energy of the reference structure, E_{RAG} , for which the atomic positions verify $\vec{r}_{\boldsymbol{\lambda}} = \vec{R}_{\Lambda} + \vec{\tau}_{\lambda}$ i.e. $(\{\vec{u}_{\boldsymbol{\lambda}}\}, \overleftrightarrow{\eta}) = (0,0)$. The second term, $\Delta E_{\text{eff}} = E_{\text{eff}} - E_{\text{RAG}}$, which describes the energy of the effective polynomial, captures the energy variations around the reference atomic geometry (RAG) and is a function of the atomic positions and the strain in the crystal after undergoing a distortion.

A natural approach to handling the energy ΔE_{eff} is to decompose it into three terms, distinguishing those associated with atomic displacements and strain dependence,

$$\Delta E_{\text{eff}}(\{\vec{u}_{\boldsymbol{\lambda}}\}, \overleftrightarrow{\eta}) = E_{\text{p}}(\{\vec{u}_{\boldsymbol{\lambda}}\}) + E_{\text{s}}(\overleftrightarrow{\eta}) + E_{\text{sp}}(\{\vec{u}_{\boldsymbol{\lambda}}\}, \overleftrightarrow{\eta}). \tag{1.4}$$

Here, $E_{\text{p}}(\{\vec{u}_{\boldsymbol{\lambda}}\})$ denotes the phonon contribution to the total energy when the RAG is distorted by atomic displacements, while the energy change due to a strain applied on the RAG is given by E_{s}. The energy variations resulting from the phonon-strain coupling, arising from the simultaneous presence of both distortions are captured by E_{sp}.

PHONON CONTRIBUTION

The phonon contribution to the energy captures the variation of the energy as consequence of atomic distortions $\{\vec{u}_{\boldsymbol{\lambda}}\}$. This energy is expanded in a Taylor series around the RAG as

$$E_{\text{p}}(\{\vec{u}_{\boldsymbol{\lambda}}\}) = \frac{1}{2} \sum_{\boldsymbol{\lambda}\alpha\boldsymbol{\delta}\beta} \left. \frac{\partial^2 E_{\text{eff}}}{\partial u_{\boldsymbol{\lambda}\alpha} \partial u_{\boldsymbol{\delta}\beta}} \right|_{\text{RAG}} u_{\boldsymbol{\lambda}\alpha} u_{\boldsymbol{\delta}\beta} + \frac{1}{6} \sum_{\boldsymbol{\lambda}\alpha\boldsymbol{\delta}\beta\boldsymbol{\pi}\gamma} \left. \frac{\partial^3 E_{\text{eff}}}{\partial u_{\boldsymbol{\lambda}\alpha} \partial u_{\boldsymbol{\delta}\beta} \partial u_{\boldsymbol{\pi}\gamma}} \right|_{\text{RAG}} u_{\boldsymbol{\lambda}\alpha} u_{\boldsymbol{\delta}\beta} u_{\boldsymbol{\pi}\gamma} + \dots \tag{1.5}$$

where α, β, γ represent Cartesian axis. The first term refers to the harmonic contribution and it is associated to the second order coefficient which represents the inter-atomic force constants. The higher-order derivatives account for the anharmonic terms. Equation (1.5) is written under the assumption that the RAG is a stationary point on the potential energy surface (PES), implying that the first derivative is zero. It is essential to recognize the preceding derivatives are not independent, as they must satisfy the acoustic sum rules (ASRs), ensuring that the energy remains invariant under rigid translations of the system.

STRAIN CONTRIBUTION

The expression for the elastic energy $E(\overleftrightarrow{\eta})$ is also expanded as a Taylor series around the RAG. In order to simplify the notation, the Voigt criterion to represent the symmetric homogeneous strain tensor shall be used, Eq. (1.3). In this way,

$$E_{\mathrm{s}}(\{\overleftrightarrow{\eta}\}) = \frac{N}{2}\sum_{ab} \left.\frac{\partial^2 E_{\mathrm{eff}}}{\partial \eta_a \partial \eta_b}\right|_{\mathrm{RAG}} \eta_a \eta_b + \frac{N}{6} \left.\frac{\partial^3 E_{\mathrm{eff}}}{\partial \eta_a \partial \eta_b \partial \eta_c}\right|_{\mathrm{RAG}} \eta_a \eta_b \eta_c + \dots, \tag{1.6}$$

where N represents the number of reference unit cells in the simulation box of the crystal. The second order coefficient is the elastic force constant, the elastic tensor for fixed atomic positions and can be obtained from finite differences or DFPT calculations. As before, the first order term is null since we have assumed that the RAG has been selected as a stationary point of the PES.

STRAIN-PHONON CONTRIBUTION

The strain-phonon coupling contribution to the energy is also expanded in a Taylor series around the RAG. This term depends on both the atomic displacement and the strain of the system with respect to the reference structure,

$$\begin{aligned} E_{\mathrm{sp}}(\{\vec{u}_\lambda\}, \overleftrightarrow{\eta}) =& \frac{1}{2}\sum_a \sum_{\lambda\alpha} \left.\frac{\partial^2 E_{\mathrm{eff}}}{\partial \eta_a \partial u_{\lambda\alpha}}\right|_{\mathrm{RAG}} \eta_a u_{\lambda\alpha} + \\ &+ \frac{1}{6}\sum_a \sum_{\lambda\alpha\delta\beta} \left.\frac{\partial^3 E_{\mathrm{eff}}}{\partial \eta_a \partial u_{\lambda\alpha} \partial u_{\delta\beta}}\right|_{\mathrm{RAG}} \eta_a u_{\lambda\alpha} u_{\delta\beta} + \frac{1}{6}\sum_{ab} \sum_{\lambda\alpha} \left.\frac{\partial^3 E_{\mathrm{eff}}}{\partial \eta_a \partial \eta_b \partial u_{\lambda\alpha}}\right|_{\mathrm{RAG}} \eta_a \eta_b u_{\lambda\alpha} + \dots \end{aligned} \tag{1.7}$$

The lowest-order coupling term represents the force-response internal strain tensor. This constant can be computed employing DFPT or finite differences considering mixed second order derivatives with respect to homogeneous strain and atomic distortions.

This term describes the forces exerted on the atoms under a specific strain state or the stresses within the unit cell resulting from a given set of atomic displacements $\{\vec{u}\}$. In certain cases, such as the cubic perovskite structure, which serves as the reference for the second-principles model developed in this work, the lowest-order coupling term vanishes due to symmetry. For the phonon-coupling term, the aforementioned coefficients must comply with the ASRs.

The parameters of the lattice model are the coefficients shown in Eqs. (1.5)-(1.7). As they have been presented so far, they are functions of the displacements of the atoms with respect the positions of the RAG. However, they can also be rewritten in terms of *displacement differences*, so that the potential is explicitly compliant with the acoustic sum rule at all orders of the expansion [36].

In addition, employing the points symmetry and the lattice-translations which characterized the reference structure around which we develop the polynomial expansion, it is possible to build symmetry-based relationships among the different coefficient reducing the number of independent parameters. This symmetry allows for the construction of irreducible sets of anharmonic coefficients, and can be understood as restrictions in the system. Consequently, the terms in the Taylor series can be grouped attending to these symmetry relations stabilizing *symmetry-adapted terms* (SAT), which group together symmetry-equivalent contributions.

1.2.2.2 Building lattice models

As it was explained at the end of Sec. 1.1, the lattice models discussed in this thesis encompass all lattice degrees of freedom, in contrast to effective Hamiltonian models which consider only a subset of selected modes. The coefficients appearing in the previous Eqs. (1.5)-(1.7) represent the parameters of these lattice models. The problem now is how to give values to these coefficients and construct the model for a particular system. Since the original proposal described in Ref. [36], different strategies have been designed [51].

These construction models share some common backgrounds. In particular: (i) they must be obtained from a set of first-principles calculations on a *training set* (TS) of relevant lattice-dynamical and structural data. The TS should contain a representative "unbiased" sampling of the potential energy surface, and can be generated in many different ways. The most used one is to take snapshots of first-principles molecular dynamic runs at different temperatures, augmented by an automatic generation of distorted structures from random occupations of stable phonon modes. (ii) The parameters are chosen to minimize a positive defined *goal function* (GF). This GF is a simple function of the model parameters that serve as a metric of the "distance" between the energy, forces, and stresses predicted by the model and those obtained in the first-principles results.

The most important differences in the construction models come from the fact about whether all the parameters are fitted, or some of them are computed. In particular for the lattice atomistic model all the second order energy derivatives can be directly *calculated* by linear response within the Density Functional Perturbation Theory (DFPT) [52; 53; 54], providing exactly all the *harmonic terms*. The terms required to compute the long-range dipole-dipole interactions, such as the Born effective charges or the electronic (optical) dielectric tensor are also directly accessible from all the modern first-principles packages. Then, *only the anharmonic parts* for the phonon contributions and the phonon-strain coupling are fitted to reproduce the energy, forces, and stresses of the first-principles simulations included in the training set, keeping only the most relevant symmetry adapted terms. This is the approach followed in MULTIBINIT [49].

Other strategy is to fit all the parameters, including the harmonic ones, following the scheme proposed by Escorihuela, Wojdeł and Íñiguez [51]. This is the schema followed in SCALE-UP. Ultimately, this process involves solving a system of linear equations, while recognizing that there are far too many independent coefficients to determine without imposing constraints. To address this, a limited subset of important terms is selected, and their coefficients are fitted accordingly. Defining a complete lattice expansion is impractical due to the exponential increase in the number of SAT coefficients with increasing order of expansion. Then, to prioritize the most significant anharmonic coefficients, a series of cutoff distances are introduced, ensuring that only atomic pairs separated by distances below the cutoff are included in the expansion. Special attention is paid to the fact the model must be bounded, avoiding terms that would produce divergences favoring nonphysically long atomic displacements.

In this Section, we have presented the second-principles model for the lattice. The goal of this thesis is to extend this methodology to incorporate an accurate treatment of the relevant electronic degrees of freedom. In the following Sec. 1.3, we introduce the tight-binding formalism that is at the basis of the new implementation. A comprehensive description of its integration into the second-principles code is provided in Chapter 4.

1.3 Treatment of the electronic degrees of freedom: the tight-binding approach

The previous effective Hamiltonian approach presented in Sec. 1.1 does not consider explicitly the electronic degrees of freedom. Thus, if the electrons are essential to understand the physics of the problem, and in agreement with the spirit followed for the lattice, a simplified and efficient scheme for their treatment must be constructed.

Elaborating on the conditions suggested by Pople and Beveridge [55], and beautifully summarized in Ref. [56], the features that this method should fulfill are: (i) the method should be simple enough to allow for its application to large systems with a reasonable computational effort; (ii) the approximations introduced should not be that severe that they modify the physical forces that determine structural and dynamical properties; (iii) the approximate wave function should be as unbiased as possible; (iv) the method should account for all the chemically active electrons; and (v) the approach should be sufficiently general to allow for systematic improvements, from the simplest possible up to *ab initio* levels.

The tight-binding approach has been proposed as a way to cope with all of these requirements, in order to simulate the *electronic properties* of the material with arbitrarily high accuracy, and at a modest computational cost. In the following we shall summarize the most important points of such approximation.

The first issue to be introduced is the concept of the *basis set* to expand the Hilbert space of the electronic eigenfunctions of the Hamiltonian. The starting point of the tight-binding approach is to consider that, in a first order approximation, electrons are localized in a single atom, but they have the possibility to jump to neighboring atoms [56]. In order to have an efficient method, the range of localization for these hopping should be small. Therefore, based on the two previous assumptions, a basis set made of localized atomic orbitals is a sensible choice. We select this to explain the basics of the approximation, although the version we shall work on in the final code is based on the use of Wannier functions, that shall be introduced in Sec. 1.4.

Let's start our discussion with a set of local orbital basis $\phi_\alpha(\vec{r}-\vec{R}_I)$, each associated with an atom in the unit cell at position $\vec{R}_I$. In principle, α might run over all the atomic orbitals of a given atom ($1s$, $2s$, $2p$, $3d$, ...). In principle, the atomic orbitals considered are those occupied in the isolated atom, and it will be assumed that these orbitals hybridize to give rise to energy bands. Therefore, for each atomic species we consider a minimal basis set of atomic valence orbitals, i.e. one orbital for each valence state occupied in the isolated atom [56]. In practice, we restrict the atomic orbitals that will be explicitly included in the simulation to a given subset, typically only the atomic orbitals that participate in the description of the states at the energy window of interest are retained. For instance, if we are interested in the description of the lowest conduction bands of $SrTiO_3$, we can retain in the basis set only the $t_{2\mathrm{g}}$ orbitals of Ti ($3d_{xy}$, $3d_{yz}$, $3d_{xz}$) as in Ref. [57].

In order to simplify the notation, we will let μ denote both α and the site I, so that μ runs from 1 to N_{basis} (the total number of atomic orbitals retained in the basis set in a unit cell). With this new notation, the atomic orbital can also be written as $\phi_\mu(\vec{r}-\vec{R}_\mu)$. In a crystal, the atoms in a unit cell are at positions $\vec{\tau}_{\kappa,j}$, where $\vec{\tau}_{\kappa,j}$ is the position of the $j=1,\ldots,n_\kappa$ atoms of type κ. The composite index $\{\kappa, j, \alpha \to \mu\}$ allows the entire basis to be specified by $\phi_\mu\left[\vec{r}-\left(\vec{\tau}_\mu+\vec{T}\right)\right]$, where $\vec{T}$ is a translation vector.

We know that the eigenstates of a Hamiltonian in a periodic potential must comply with the Bloch theorem,

$$\psi_{i\vec{k}}(\vec{r}+\vec{T}) = e^{i\vec{k}\cdot\vec{T}}\psi_{i\vec{k}}(\vec{r}), \tag{1.8}$$

where i is a band index and $\vec{k}$ is a wavevector in the first Brillouin zone. A deeper discussion of the Bloch theorem and its consequences on our method will be described in Sec. 2.1.2. A good starting point to represent the eigenstate $\psi_{i\vec{k}}$ in a basis of localized atomic orbitals is to define a basis that complies also with the Bloch theorem, i.e. by defining a basis state with wavevector $\vec{k}$

$$\phi_{\mu\vec{k}}(\vec{r}) = A_{\mu\vec{k}}\sum_{\vec{T}} e^{i\vec{k}\cdot\vec{T}}\phi_\mu(\vec{r}-\vec{\tau}_\mu-\vec{T}). \tag{1.9}$$

where $\vec{\tau}_\mu$ is the position of the atom within the unit cell to which orbital ϕ_μ belongs, and $A_{\mu\vec{k}}$ is a normalization factor.

1.3.1 Hamiltonian matrix elements

The next step is the computation of the Hamiltonian matrix elements in these basis of numerical atomic orbitals. From the expressions given in Eq. (1.9)

$$\begin{aligned}\langle\phi_\mu(\vec{k})|H|\phi_\nu(\vec{k}')\rangle &= A^*_{\mu\vec{k}}\sum_{\vec{T}} e^{-i\vec{k}\cdot\vec{T}}\int_{\text{all space}} d\vec{r}\phi^*_\mu(\vec{r}-\vec{\tau}_\mu-\vec{T})H\left(A_{\nu\vec{k}'}\sum_{\vec{T}'} e^{i\vec{k}'\cdot\vec{T}'}\phi_\nu(\vec{r}-\vec{\tau}_\nu-\vec{T}')\right)\\ &= A^*_{\mu\vec{k}}A_{\nu\vec{k}'}\sum_{\vec{T}}\left(\sum_{\vec{T}'} e^{i\left(\vec{k}'\cdot\vec{T}'-\vec{k}\cdot\vec{T}\right)}\int_{\text{all space}} d\vec{r}\phi^*_\mu(\vec{r}-\vec{\tau}_\mu-\vec{T})H\phi_\nu(\vec{r}-\vec{\tau}_\nu-\vec{T}')\right).\end{aligned} \tag{1.10}$$

In the sum in brackets, the translation vector $\vec{T}$ is fixed. We can therefore make the following change of variables in the integral $\vec{r}^{\,\prime} = \vec{r} - \vec{T}$,

$$\begin{aligned}
\langle \phi_\mu(\vec{k})|H|\phi_\nu(\vec{k}^{\,\prime})\rangle &= A^*_{\mu\vec{k}} A_{\nu\vec{k}^{\prime}} \sum_{\vec{T}} \left(\sum_{\vec{T}^{\prime}} e^{i\left(\vec{k}^{\prime}\cdot\vec{T}^{\prime} - \vec{k}\cdot\vec{T}\right)} \int_{\text{all space}} d\vec{r}^{\,\prime}\, \phi^*_\mu(\vec{r}^{\,\prime} + \vec{T} - \vec{\tau}_\mu - \vec{T}) H \phi_\nu(\vec{r}^{\,\prime} + \vec{T} - \vec{\tau}_\nu - \vec{T}^{\prime}) \right) \\
&= A^*_{\mu\vec{k}} A_{\nu\vec{k}^{\prime}} \sum_{\vec{T}} \left(\sum_{\vec{T}^{\prime}} e^{i\left(\vec{k}^{\prime}\cdot\vec{T}^{\prime} - \vec{k}\cdot\vec{T}\right)} \int_{\text{all space}} d\vec{r}^{\,\prime}\, \phi^*_\mu(\vec{r}^{\,\prime} - \vec{\tau}_\mu) H \phi_\nu \left[\vec{r}^{\,\prime} - \vec{\tau}_\nu - \left(\vec{T}^{\prime} - \vec{T}\right)\right] \right).
\end{aligned} \tag{1.11}$$

Now, making the change in variables $\vec{T}^{\prime} - \vec{T} = \vec{T}^{\prime\prime}$,

$$\begin{aligned}
\langle \phi_\mu(\vec{k})|H|\phi_\nu(\vec{k}^{\,\prime})\rangle &= A^*_{\mu\vec{k}} A_{\nu\vec{k}^{\prime}} \sum_{\vec{T}} \left(\sum_{\vec{T}^{\prime\prime}} e^{i\left(\vec{k}^{\prime}\cdot\left(\vec{T}^{\prime\prime}+\vec{T}\right) - \vec{k}\cdot\vec{T}\right)} \int_{\text{all space}} d\vec{r}^{\,\prime}\, \phi^*_\mu(\vec{r}^{\,\prime} - \vec{\tau}_\mu) H \phi_\nu(\vec{r}^{\,\prime} - \vec{\tau}_\nu - \vec{T}^{\prime\prime}) \right) \\
&= A^*_{\mu\vec{k}} A_{\nu\vec{k}^{\prime}} \left[\sum_{\vec{T}} e^{i\left(\vec{k}^{\prime} - \vec{k}\right)\cdot\vec{T}} \right] \left(\sum_{\vec{T}^{\prime\prime}} e^{i\vec{k}^{\prime}\cdot\vec{T}^{\prime\prime}} \int_{\text{all space}} d\vec{r}^{\,\prime}\, \phi^*_\mu(\vec{r}^{\,\prime} - \vec{\tau}_\mu) H \phi_\nu(\vec{r}^{\,\prime} - \vec{\tau}_\nu - \vec{T}^{\prime\prime}) \right) \\
&= N\delta_{\vec{k},\vec{k}^{\prime}} A^*_{\mu\vec{k}} A_{\nu\vec{k}^{\prime}} \left(\sum_{\vec{T}^{\prime\prime}} e^{i\vec{k}^{\prime}\cdot\vec{T}^{\prime\prime}} \int_{\text{all space}} d\vec{r}^{\,\prime}\, \phi^*_\mu(\vec{r}^{\,\prime} - \vec{\tau}_\mu) H \phi_\nu(\vec{r}^{\,\prime} - \vec{\tau}_\nu - \vec{T}^{\prime\prime}) \right),
\end{aligned} \tag{1.12}$$

where we have used Eq. (1.13a) [see Ref. [58], Eq. (F.4)]

$$\sum_{\vec{T}} e^{i\vec{k}\cdot\vec{T}} = N\delta_{\vec{k},\vec{0}}, \tag{1.13a}$$

$$\sum_{\vec{k}} e^{i\vec{k}\cdot\vec{T}} = N\delta_{\vec{T},\vec{0}}, \tag{1.13b}$$

where $\vec{T}$ runs through the N sites of the Bravais lattice in Eq. (1.13a), and the sum on $\vec{k}$ in Eq. (1.13b) runs through all the sites in the first Brillouin zone consistent with the Born-von Kármán boundary condition.

The integral within the last bracket of Eq. (1.12) requires a more deep study. It represents the interaction between two orbitals in real space. If the orbital μ and the orbital ν are the same and are located on the same atom ($\vec{T} = 0$), this integral is often recalled as the *on-site* energy term. These depend on the type of orbital, and would be the only matrix elements that would be non-zero in the limiting case where the atoms are sufficiently far away.

When the atoms are brought together, there is a possibility for the electrons to jump from one orbital to its neighbors. This is represented by the *off-diagonal* terms in the Hamiltonian (where $\mu \neq \nu$ and/or the orbitals are not centered on the same atom, $\vec{T} \neq 0$). These integrals are often called as *hopping-integrals*. The hopping parameters depend on the type of orbitals between which the electron jumps, and on the distance between the atoms. The hopping parameters can be interpreted in terms of the band width, and fitted to experimental or first-principles calculated values.

Now, we can denote the matrix element of the hamiltonian of a orbital μ in the unit cell at the origin and an orbital ν in the cell labelled by translation vector $\vec{T}^{\prime\prime}$ as

$$H_{\mu,\nu}(\vec{T}^{\prime\prime}) = \int_{\text{all space}} d\vec{r}^{\,\prime}\, \phi^*_\mu(\vec{r}^{\,\prime} - \vec{\tau}_\mu) H \phi_\nu(\vec{r}^{\,\prime} - \vec{\tau}_\nu - \vec{T}^{\prime\prime}). \tag{1.14}$$

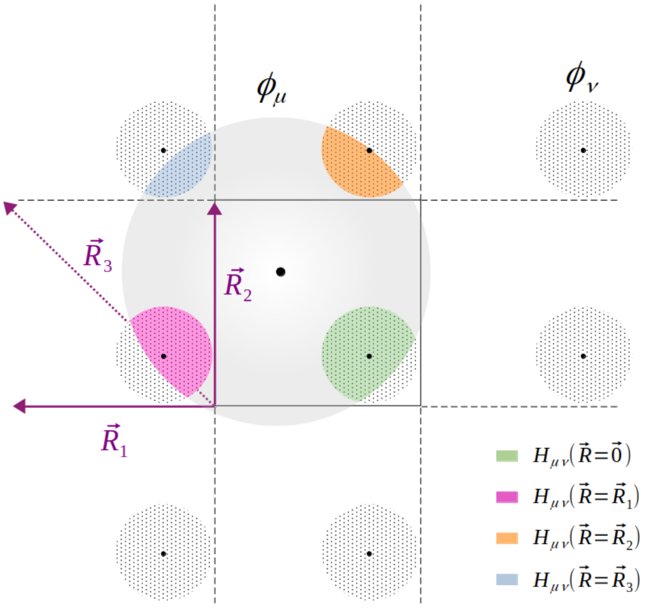


Figure 1.3: Interactions between an orbital placed at the home unit cell, $\phi_\mu(\vec{r}' - \vec{\tau}_\mu)$, and the periodic images of another orbital ϕ_ν. The Hamiltonian matrix elements will vanish if there is not overlap between the orbitals.

For the purpose of computing the matrix elements of the Hamiltonian, it is shown the interaction between two different atomic-like orbitals in a periodic crystal, Fig. 1.3. Firstly, we select an orbital $\phi_\mu(\vec{r}' - \vec{\tau}_\mu)$ (represented by a gray shadow) at the home unit cell ($\vec{R} = \vec{0}$) and the periodic images in the lattice of a ϕ_ν orbital (dotted pattern). The matrix element $H_{\mu\nu}(\vec{T}'')$ will vanish if the orbitals $\phi_\mu(\vec{r}' - \vec{\tau}_\mu)$ and $\phi_\nu(\vec{r}' - \vec{\tau}_\nu - \vec{T}'')$ don't overlap. Regarding the sketch, the only $\phi_\nu(\vec{r}' - \vec{\tau}_\nu - \vec{T}'')$ orbitals which intersect with $\phi_\mu(\vec{r}' - \vec{\tau}_\mu)$ are the ones placed in the cells characterized by the lattice vectors $\vec{R} = \vec{0}, \vec{R}_1, \vec{R}_2, \vec{R}_3$. The orbital overlaps between $\phi_\mu(\vec{r}' - \vec{\tau}_\mu)$ and $\phi_\nu(\vec{r}' - \vec{\tau}_\nu - \vec{T}'')$ are shown in green, pink, orange and blue respectively.

With the definition given in Eq. (1.14), then Eq. (1.12) transforms into

$$\langle \phi_\mu(\vec{k})|H|\phi_\nu(\vec{k}')\rangle = N\delta_{\vec{k},\vec{k}'}A^*_{\mu\vec{k}}A_{\nu\vec{k}'}\left(\sum_{\vec{T}''} e^{i\vec{k}'\cdot\vec{T}''}H_{\mu,\nu}(\vec{T}'')\right). \tag{1.15}$$

Finally, assuming that the normalization factors of the Bloch orbitals are $1/\sqrt{N}$, we finally arrive to

$$H_{\mu,\nu}(\vec{k}) = \langle \phi_\mu(\vec{k})|H|\phi_\nu(\vec{k})\rangle = \sum_{\vec{T}} e^{i\vec{k}\cdot\vec{T}}H_{\mu,\nu}(\vec{T}). \tag{1.16}$$

1.3.2 Overlap matrix elements

By analogy, the matrix elements of the overlap matrix between Bloch basis orbitals $\phi_\mu(\vec{k})$ and $\phi_\nu(\vec{k}')$ are also diagonal in $\vec{k}$ and equals

$$S_{\mu,\nu}(\vec{k}) = \langle\phi_\mu(\vec{k})|\phi_\nu(\vec{k})\rangle = \sum_{\vec{T}} e^{i\vec{k}\cdot\vec{T}} S_{\mu,\nu}(\vec{T}). \tag{1.17}$$

1.3.3 The secular equation

Since the hamiltonian conserves $\vec{k}$, an eigenfunction of the Schrödinger equation in a basis always can be written in the form

$$\psi_{i\vec{k}}(\vec{r}) = \sum_{\mu} c_{\mu i}(\vec{k})\phi_{\mu\vec{k}}(\vec{r}). \tag{1.18}$$

Inserting this expansion of the eigenfunctions into the Schrödinger equation,

$$\begin{aligned} &H\psi_{i\vec{k}}(\vec{r}) = \varepsilon_i(\vec{k})\psi_{i\vec{k}}(\vec{r}), \\ &H\left(\sum_{\mu} c_{\mu i}(\vec{k})\phi_{\mu\vec{k}}(\vec{r})\right) = \sum_{\mu} c_{\mu i}(\vec{k})H\phi_{\mu\vec{k}}(\vec{r}) = \varepsilon_i(\vec{k})\sum_{\mu} c_{\mu i}(\vec{k})\phi_{\mu\vec{k}}(\vec{r}), \end{aligned} \tag{1.19}$$

Multiplying both sides of Eq. (1.19) by $\phi^*_{\nu\vec{k}}(\vec{r})$ and integrating over all space,

$$\begin{aligned} &\sum_{\mu} c_{\mu i}(\vec{k}) \int_{\text{all space}} d\vec{r}\phi^*_{\nu\vec{k}}(\vec{r})H\phi_{\mu\vec{k}}(\vec{r}) = \varepsilon_i(\vec{k})\sum_{\mu} c_{\mu i}(\vec{k}) \int_{\text{all space}} \phi^*_{\nu\vec{k}}(\vec{r})\phi_{\mu\vec{k}}(\vec{r}), \\ &\sum_{\mu} c_{\mu i}(\vec{k})H_{\nu,\mu}(\vec{k}) = \varepsilon_i(\vec{k})\sum_{\mu} c_{\mu i}(\vec{k})S_{\nu,\mu}(\vec{k}). \end{aligned} \tag{1.20}$$

Transposing all the terms to the left hand side in Eq. (1.20),

$$\sum_{\mu}\left(H_{\nu,\mu}(\vec{k}) - \varepsilon_i(\vec{k})S_{\nu,\mu}(\vec{k})\right)c_{\mu i}(\vec{k}) = 0, \tag{1.21}$$

which constitutes the secular equation that must be solved for a given wave vector $\vec{k}$ within the first Brillouin zone.

1.3.4 Calculation of the matrix elements

According to Eq. (1.21), the band structure of the system can be computed. For a given $\vec{k}$ point in the first-Brillouin zone, the square Hamiltonian and Overlap matrices, of range N_{basis}, are build following Eq. (1.16) and Eq. (1.17), respectively. Finally, the generalized eigenvalue problem given by Eq. (1.21) is solved to find the eigenvalues, $\varepsilon_i(\vec{k})$, and the coefficients of the expansion of the eigenfunctions in terms of the basis set, $c_{\mu i}(\vec{k})$.

The problem now is how to calculate in an efficient way the Hamiltonian $[H_{\mu\nu}(\vec{T})]$ and the Overlap $[S_{\mu\nu}(\vec{T})]$ matrix elements between the orbitals included in the basis set.

In the *empirical* tight-binding models, the parameters are fitted to reproduced the experimental band structures. Although they can achieve excellent agreement with specific properties, their transferability (the ability to accurately describe systems or conditions beyond those used in the fitting) can be limited, requiring careful validation. For the bands, the parameters are usually designed to fit selected eigenvalues for a particular crystal structure and lattice constants. Extensive tables can be found in the works by Papaconstantopoulos [59] and Harrison [60].

In the mid-1990s, several approaches to ab initio tight-binding were introduced [61]. These methods employed a minimal basis set, consisting only of numerical atomic orbitals for valence electrons. The Hamiltonian in Eq. (1.14) was replaced with the Kohn-Sham Hamiltonian, requiring integrals involving one to four atomic centers when expressed in a basis of atomic orbitals. To improve computational efficiency, many contributions to the Hamiltonian matrix elements were omitted, retaining only the one- and two-center terms. It was assumed that the omitted terms could be accounted for through a parametrization of the remaining integrals—an approximation that, though uncontrolled, could be justified by demonstrating the smallness of these neglected terms. The matrix elements depend solely on interatomic distances and orientations. Consequently, they can be precomputed for each pair of atoms and stored in a table as functions of distance. For any given interatomic distance, intermediate values can be interpolated from these stored data.

The foundational work of Slater and Koster [62] enabled the efficient computation of two-center integrals, where orbitals are located on separate atomic centers. These integrals are categorized based on the relative orientation of orbitals with respect to the interatomic axis: σ (parallel), π (perpendicular), or δ (for d-orbitals). When orbitals are not perfectly aligned or perpendicular, the integrals can be determined as linear combinations of these types by applying a coordinate rotation. Despite the great simplification that leads to an extremely useful approach to understanding electron in materials, the Slater-Koster method considering only two-center matrix elements are not transferable between materials.

The proper description of the electronic band structure usually requires the consideration of many atomic orbitals. For instance, in the case of $SrTiO_3$ we need to include all the valence orbitals of O, Sr and Ti atoms and even the semi-core states for the metallic cations. To achieve accurate results using SIESTA, it is necessary to employ single-ζ for the semicore and double-ζ polarized basis sets for the valence shell. This implies that for each unit cell analyzed in the study, 72 numerical atomic orbitals must be included. The diagonalization of the Kohn-Sham Hamiltonian yields the same number of bands per $\vec{k}$-point in the first Brillouin zone, ranging from the deeply localized semicore Sr and O orbitals to numerous conduction band orbitals corresponding to 72 bands per unit cell. Even more, for $SrTiO_3$, a characteristic distortion is the antiferrodistortive phase. The smallest computational cell required to study this system is a $\sqrt{2} \times \sqrt{2} \times 1$ unit cell, achieved by rotating the x- and y-axes by 45°. However, if the axes are to remain fixed, the required simulation cell corresponds to a $2 \times 2 \times 2$ supercell. This configuration necessitates the diagonalization of matrices with dimensions of 576 (72 orbitals per unit cell × 8 unit cells). If we extend our analysis to simulate defects in this system, such as polarons, the computational cost of diagonalization, which scales as $O(N^3_{\mathrm{orb/cell}})$ we can ascertain how the first-principles treatment of the super cell to simulate defects is unfeasible.

In many cases, it is beneficial to restrict the electronic degrees of freedom to the states most relevant to the problem under study. Specifically, in our case, these states correspond to the top of the valence band and the bottom of the conduction band. This focus is particularly valuable for understanding the mechanisms underlying electron and hole polarons, as well as excitonic effects. Additionally, it is possible to restrict the analysis to the top of the valence band when studying electron polarons. Such a reduction in electronic degrees of freedom allows the investigation to concentrate on the Bloch states directly involved in the behavior of crystal defects.

A viable strategy to restrict the electronic degrees of freedom and lower the computational cost of calculations while preserving accuracy is to replace the original basis set with a minimal orthonormal basis constructed from Wannier functions. These Wannier functions faithfully reproduce the bands of interest and provide an efficient basis for analyzing the electronic degrees of freedom within the framework of second-principles methodology.

1.4 Bloch functions and Wannier functions

The tight-binding approach, introduced in Sec. 1.3, relies on localized atomic orbitals as a basis set for expanding the electronic wavefunctions. An alternative yet closely related approach employs Wannier functions, which provide a maximally localized representation of the electronic states in real space. Wannier functions are particularly advantageous in capturing the underlying physics of electronic systems, especially when dealing with strongly correlated or localized electrons. Indeed, the construction of localized Wannier orbitals from the Kohn-Sham states calculated using density-functional theory represents an elegant and systematic way to obtain realistic (materials-specific) tight-binding models. The Maximally localized Wannier functions (MLWF) approach automatically leads to a tight-binding parameterization of the relevant bands so that the parameters of the model can simply be “read off” and no fitting procedure is required.

1.4.1 Isolated band case

The main property of crystal structure is its periodicity. Consequently, the electronic structure calculations in crystals are performed by applying periodic boundary conditions. In the study of the one-particle Hamiltonian, $H = p^2/2m + V(\vec{r})$ with a real periodic potential $V(\vec{r})$ the Hamiltonian H is invariant to translations and then commutes with the lattice-translation operator $T_{\vec{T}}$. Consequently, as eigenfunctions of the Hamiltonian can be chosen the Bloch orbitals,

$$\psi_{n\vec{k}}(\vec{r}) = u_{n\vec{k}}(\vec{r})e^{i\vec{k}\cdot\vec{r}} \tag{1.22}$$

determined by a plane wave arising from the translational symmetry (characterized by a $\vec{k}$-vector) modulated by a periodic function with the same periodicity of the potential and defined in real space, $u_{n\vec{k}}(\vec{r})$. By construction, Bloch functions are delocalized and span on over all the space. They are also orthogonal and represent a basis of a Hilbert space. The Bloch functions are not univocally defined due to its dependence on an arbitrary phase. One particularly appropriate choice of these phases is the use of a smooth and periodic gauge [63].

In order to find localized functions in the real space storing the same information as the Bloch basis set, the Fourier transform is applied. The Fourier-transform partners of the Bloch functions are known as Wannier functions [64] which represent a basis set of the Hilbert space too.

Attending to a band diagram, a Bloch band is considered *isolated* if never crosses the bands above or below it, in other words, the band does not exhibit degeneracy with another band at any point within the Brillouin zone. However, if a collection of bands are interconnected by degeneracies and simultaneously do not intersect with bands lower or higher in energy, they constitute a composite group. For the sake of simplicity, we are going to consider the case of one isolated band, meaning that the band never crosses the bands above or below it. Taking the Bloch function $\psi_{n\vec{k}}(\vec{r})$ linked to the n index band and wave vector $\vec{k}$, and the Wannier function labeled by a cell index $\vec{R}$ and a band-like index n, $\chi_{n\vec{R}}(\vec{r})$, we can go from reciprocal space to real space by

$$|\chi_{n\vec{R}}\rangle = \frac{V_{\text{cell}}}{(2\pi)^3} \int_{\text{BZ}} \mathrm{d}\vec{k} e^{-i\vec{k}\cdot\vec{R}} |\psi_{n\vec{k}}\rangle . \tag{1.23}$$

The inverse transformation is given by

$$|\psi_{n\vec{k}}\rangle = \sum_{\vec{R}} e^{-i\vec{k}\cdot\vec{R}} |\chi_{n\vec{R}}\rangle . \tag{1.24}$$

This is an special case of an unitary transformation where V_{cell} is the volume of the primitive unit cell and the integral is defined in the Brillouin zone (BZ). Consequently, we can view the Bloch and Wannier functions as providing two different basis sets describing the same manifold of states associated with the electronic band in question.

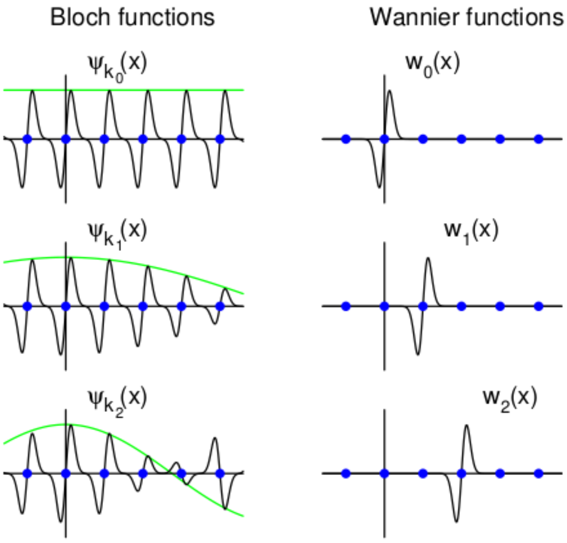


Figure 1.4: Bloch and Wannier functions for a chain of atoms in 1D (blue dots). Left: Bloch functions of a single band for three different values of k. The green curves show the envelopes e^{ikx} of the Bloch functions. Right: Wannier functions obtained for the same band, producing periodic images of one another. Reprinted with permission from Ref. [65].

1.4.2 Properties

Among the properties of the Wannier functions in the single band case, we can stand out the following, as nicely summarized in Ref. [63].

- The Wannier functions are localized in real space. A Wannier function associated with a band n can be understood as a function localized in a specific unit cell, labelled by the vector $\vec{R}$ in the subscript. While it is primarily centered in this unit cell, its tails may extend into neighboring unit cells. As a result, the amplitude of the Wannier function gradually decays with increasing distance from the unit cell where it is localized. That is

$$\left|\chi_{n\vec{R}}(\vec{r})\right| \to 0 \text{ as } |\vec{r}-\vec{R}| \text{ gets larger.} \tag{1.25}$$

  As result, they are perfect to generate range-limited models.
- In real space, a Wannier function can be defined at each unit cell, with the functions in different unit cells being translationally equivalent to one another,

$$\chi_{n\vec{R}}(\vec{r}) = \chi_{n\ \vec{R}+\vec{R}'}(\vec{r}+\vec{R}'). \tag{1.26}$$

  Or more formally

$$|\chi_{n\ \vec{R}+\vec{R}'}\rangle = T_{\vec{R}'}|\chi_{n\vec{R}}\rangle \tag{1.27}$$

  where $T_{\vec{R}'}$ is the operator that translates the system by a lattice vector $\vec{R}'$.

Proof. Applying the definition of a Wannier function, Eq. (1.23),

$$\chi_{n\,\vec{R}+\vec{R}'}(\vec{r}+\vec{R}') = \frac{V_{\text{cell}}}{(2\pi)^3}\int_{\text{BZ}} \mathrm{d}\vec{k}e^{-i\vec{k}\cdot(\vec{R}+\vec{R}')}\psi_{n\vec{k}}(\vec{r}+\vec{R}'), \tag{1.28}$$

The Bloch theorem states $\psi_{n\vec{k}}(\vec{r}+\vec{R}') = e^{i\vec{k}\cdot\vec{R}'}\psi_{n\vec{k}}(\vec{r})$, resulting

$$\begin{aligned}\chi_{n\,\vec{R}+\vec{R}'}(\vec{r}+\vec{R}') &= \frac{V_{\text{cell}}}{(2\pi)^3}\int_{\text{BZ}} \mathrm{d}\vec{k}e^{-i\vec{k}\cdot(\vec{R}+\vec{R}')}e^{i\vec{k}\cdot\vec{R}'}\psi_{n\vec{k}}(\vec{r}) = \\ &= \frac{V_{\text{cell}}}{(2\pi)^3}\int_{\text{BZ}} \mathrm{d}\vec{k}e^{-i\vec{k}\cdot\vec{R}}\psi_{n\vec{k}}(\vec{r}) = \chi_{n\vec{R}}(\vec{r})\end{aligned} \tag{1.29}$$

□

- The Wannier functions form an orthonormal basis set,

$$\langle\chi_{m\vec{R}'}|\chi_{n\vec{R}}\rangle = \delta_{\vec{R}'\vec{R}}\delta_{mn}. \tag{1.30}$$

- The Wannier functions span the same subspace of the Hilbert space as is spanned by the Bloch functions from which they are constructed. To elucidate this property, let us define $\mathscr{P}_n$ as the projector operator onto band n. This operator can then be expressed as,

$$\mathscr{P}_n = \frac{V_{\text{cell}}}{(2\pi)^3}\int_{\text{BZ}} |\psi_{n\vec{k}}\rangle\langle\psi_{n\vec{k}}|\,\mathrm{d}^3k = \sum_{\vec{R}}|\chi_{n\vec{R}}\rangle\langle\chi_{n\vec{R}}|\,. \tag{1.31}$$

From this it follows that the total charge density in real space $\rho_n(\mathbf{r})$ in band n is the same when computed in every representation,

$$\rho_n(\vec{r}) = -e\langle\vec{r}|\mathscr{P}_n|\vec{r}\rangle = -e\frac{V_{\text{cell}}}{(2\pi)^3}\int_{\text{BZ}} |\psi_{n\vec{k}}(\vec{r})|^2\mathrm{d}^3k = -e\sum_{\vec{R}}|\chi_{n\vec{R}}(\vec{r})|^2. \tag{1.32}$$

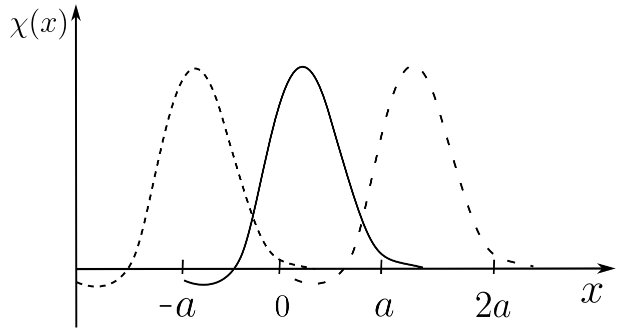


Figure 1.5: Schematic representation of three aligned Wannier functions, $\chi_{nR}(x)$, corresponding to band n in a one-dimensional crystal characterized by a lattice constant a. The solid curve is associated with the Wannier function in the reference unit cell $R = 0$, while the dotted and dashed curves illustrate those located in the unit cells at positions $R = -a$ and $R = a$, respectively. The Wannier functions are translationally invariant of one another and orthonormal. This picture has been adapted from Ref. [63].

All these properties are illustrated in the schematic representation of Fig. 1.5 for a simple one-dimensional case. The lattice constant of the unit cell is denoted by a. In this diagram, it is evident how the Wannier functions are centered within a specific unit cell in real space, are normalized, and exhibit localization. Additionally, neighboring Wannier functions are periodic images of each other. Notably, the functions also include regions of negative amplitude, ensuring that the scalar product between adjacent Wannier functions vanishes. This orthogonality arises from the cancellation of contributions with opposite signs in the integral over x.

All the equations and properties above can be generalized to the case of isolated band manifolds.

1.4.3 Gauge freedom

The Wannier functions are strongly non-unique as a consequence of the phase indeterminacy of Bloch orbitals. There exists a “gauge” freedom in the definition of the Bloch functions. For instance, a Bloch function $\psi_{n\vec{k}}(\vec{r})$ can be replaced by another Bloch function $\widetilde{\psi}_{n\vec{k}}(\vec{r})$ where a phase factor has been introduced,

$$\left|\widetilde{\psi}_{n\vec{k}}\right\rangle = e^{i\varphi_n(\vec{k})}\left|\psi_{n\vec{k}}\right\rangle \tag{1.33}$$

considering the periodic part of the Bloch function the previous definition is equivalent to

$$\left|\widetilde{u}_{n\vec{k}}\right\rangle = e^{i\varphi_n(\vec{k})}\left|u_{n\vec{k}}\right\rangle \tag{1.34}$$

where $\varphi_n(\vec{k})$ is a real function periodic in the reciprocal space. The replacement of the Bloch function $\left|\psi_{n\vec{k}}\right\rangle$ by the Bloch function $\left|\widetilde{\psi}_{n\vec{k}}\right\rangle$ does not change the physical properties of the problem.

However, the gauge freedom in Bloch functions, Eqs. (1.33) and (1.34), has an impact in the Wannier functions definition. In other words, the spread and the shape of the Wannier functions depend on the $\varphi_n(\vec{k})$ phase. An important consideration is that the Schrödinger equation does not specify a particular gauge for the Bloch functions. As a result, the non-uniqueness of the Wannier functions emerges from Eq. (1.23) as a consequence of the gauge freedom of the Bloch functions.

To identify a specific set during the construction of Wannier functions in an insulating material, it is essential to apply a rigorously defined localization criterion. In this work, we present the two most commonly employed methodologies: maximally localized Wannier functions (MLWFs) [66] and Wannier functions obtained through projection, following the framework of Ref. [67]. Both approaches will be discussed in detail in the subsequent Secs. 1.4.5 and 1.4.6.

1.4.4 Multiband case

Until now, we have worked on the isolated band case. However, bands can display nonanalytic behaviours as crossings with other bands and singularities. This complexity is particularly observed in the top of the valence band [in blue attending to Fig. 1.6(a)] and the bottom of the conduction band [in green regarding Fig. 1.6(a)] of the cubic paraelectric phase of SrTiO$_3$.

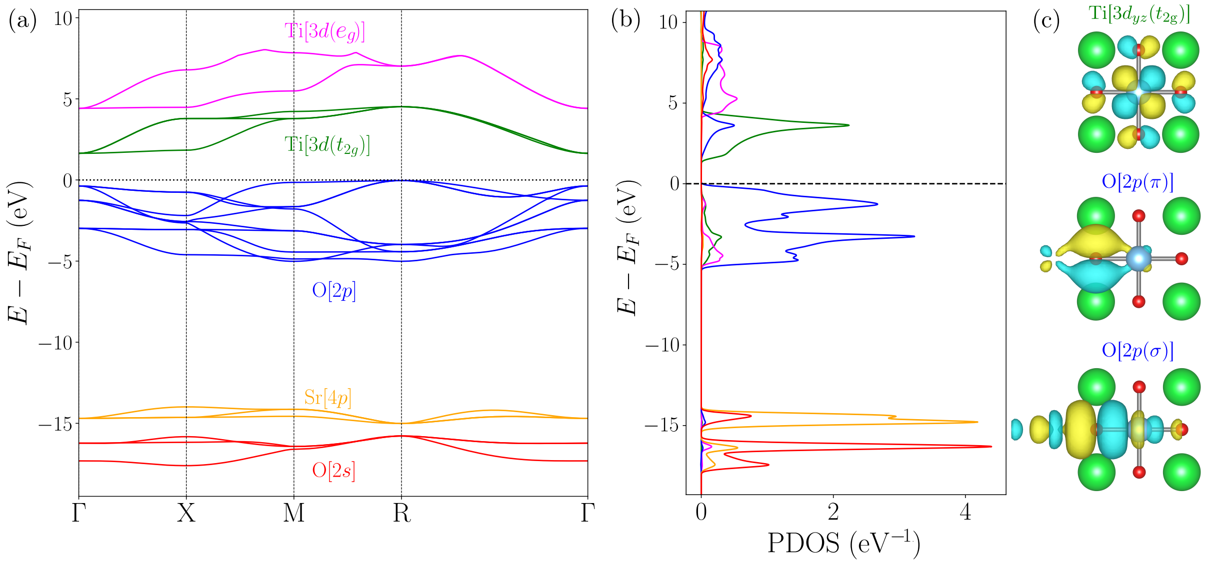


Figure 1.6: (a) Electronic band structure of SrTiO$_3$ in the cubic phase at the equilibrium lattice constant. The diagram displays the bands corresponding to the Sr $4p$ semicore levels (orange), the O $2s$ (red) and $2p$ (blue) valence states, and the Ti $3d$ conduction bands, with t_{2g} and e_g symmetries represented in green and magenta, respectively. The energy zero point is set to the top of the valence band. (b) Projected density of states (PDOS) for cubic SrTiO$_3$ at the equilibrium lattice constant. Contributions from O $2s$ (red), O $2p$ (blue), Ti t_{2g} (green), Ti e_g (magenta), and Sr $4p$ semicore (orange) states are shown. (c) Real-space representation of maximally localized Wannier functions: two orbitals with dominant O p-like character at the top of the valence band, and one Ti-centered orbital with predominant d-character associated with the bottom of the conduction.

To address this issue, we have moved away from the concept of isolated bands and we consider multiband *manifolds* defined as a set of J consecutive energy bands, covering an energy range, with the possibility of exhibiting internal crossing and degeneracies among them, but that do not become degenerate with any lower or higher band outside of the energy range anywhere in the Brillouin zone. The energy degeneracies observed at band crossings leading to non-analytic behavior generate poorly localized functions under the Fourier transform. Consequently, for a band manifold is not possible to directly insert the Bloch functions obtained from the electronic structure calculation, $\{|\psi_{n\vec{R}}\rangle\}$, into Eq. (1.23).

Nevertheless, well localized Wannier functions can still be constructed if we abandon the idea that each Wannier function should be associated with one and only one energy band. The proposed solution consists of defining a unitary transformation, $U(\vec{k})$, which transforms the original Bloch functions $|\psi_{n\vec{k}}\rangle$, typically non-smooth in the Brillouin zone (BZ), into Bloch-like functions that are smooth in $\vec{k}$-space (i.e. they satisfy the smoothness condition: $\nabla_{\vec{k}}|\widetilde{\psi}_{n\vec{k}}\rangle$ is regular at all $\vec{k}$). These Bloch-like functions are represented by $|\widetilde{\psi}_{n\vec{k}}\rangle$ and are computed as

$$|\widetilde{\psi}_{n\vec{k}}\rangle = \sum_{m=1}^{J} U_{mn}(\vec{k})\,|\psi_{m\vec{k}}\rangle\,. \tag{1.35}$$

The resulting $\widetilde{\psi}_{n\vec{k}}$ are no longer, in general, eigenstates of H. According to the unitary rotation matrices $U(\vec{k})$, they are defined at each $\vec{k}$-point within the Brillouin zone (BZ), possess a dimensionality of J, and exhibit periodicity in reciprocal space.

As a result, the Wannier orbitals set $\{|\chi_{n\vec{R}}\rangle\}$ represents each one of the J bands composing a manifold and spans the same subspace as the Bloch functions $\{\,|\psi_{n\vec{k}}\rangle\}$ considered as a group, $\{\,|\widetilde{\psi}_{n\vec{k}}\rangle\}$

$$|\chi_{n\vec{R}}\rangle = \frac{V_{\text{cell}}}{(2\pi)^3}\int_{\text{BZ}} \mathrm{d}\vec{k} e^{-i\vec{k}\cdot\vec{R}}\,|\widetilde{\psi}_{n\vec{k}}\rangle\,. \tag{1.36}$$

The unitary rotation $U(\vec{k})$ represents a gauge transformation. Then, Eq. (1.35) is replaced in Eq. (1.36) resulting

$$|\chi_{n\vec{R}}\rangle = \frac{V_{\text{cell}}}{(2\pi)^3}\int_{\text{BZ}} e^{-i\vec{k}\vec{R}} \sum_{m=1}^{J} U_{mn}(\vec{k})\,|\psi_{m\vec{k}}\rangle\, d\vec{k}. \tag{1.37}$$

For instance, revisiting the case of $\mathrm{SrTiO_3}$, the manifold formed by the top of the valence band is conformed by $J = 9$ bands while the bottom of the conduction band is composed by $J = 3$ bands.

1.4.5 Wannier functions via projection

To construct smooth and well-localized Wannier functions, one effective approach is to use projections over atomic-like orbitals. This method begins by considering N localized trial orbitals of the simulation cell ($\vec{R} = \vec{0}$) $\{\phi_n(\vec{r})\}$ which typically corresponds to the number of resulting Wannier functions. These trial orbitals serve as initial guesses or "seeds" for the Wannier functions, and are chosen to reflect the chemical and physical nature of the bands comprising the manifold of interest. The primary requirement for these orbitals is their localization in real space, justifying their atomic-like character.

In particular, in the context of using WANNIER90 as a post-processing tool with SIESTA software, which is the method employed in this work, the trial orbitals $\phi_n(\vec{r})$ are selected as the atomic-like numerical orbitals from SIESTA basis set. The basis set in SIESTA in real space, and thus, the trial orbital $\phi_n(\vec{r})$ refers to this real-space representation.

In general, to compute Wannier functions through projection, one begins considering the definition provided in Eq. (1.23), where Wannier functions are expressed as the Fourier transform of Bloch states[68]. In order to perform the projection of Bloch functions onto the trial atomic-like orbitals ϕ_n the first step is to express these trial orbitals in reciprocal space using a Fourier transform. This process must account for the translational symmetry of the crystal, which introduces periodic replicas of the trial orbitals $\phi_n(\vec{r}-\vec{R})$. For compactness of notation, the atomic-like orbitals will be expressed as

$$\phi_n(\vec{r}-\vec{R}) \equiv \phi_{n\vec{R}}, \tag{1.38}$$

$$|\phi_{n\vec{k}}\rangle = \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} |\phi_{n\vec{R}}\rangle . \tag{1.39}$$

At this stage, the objective is to determine the Wannier functions that most closely approximate the given basis set. This is achieved by projecting the states $\{|\phi_{n\vec{k}}\rangle\}$ onto the Bloch manifold. With respect to the wavevector $\vec{k}$, this projection yields

$$|\varphi_{n\vec{k}}\rangle = \sum_{m=1}^{J} |\psi_{m\vec{k}}\rangle \langle \psi_{m\vec{k}}|\phi_{n\vec{k}}\rangle. \tag{1.40}$$

Although $|\varphi_{n\vec{k}}\rangle$ functions are smooth in reciprocal space, they would not be orthonormal. To solve this problem, the solution is based on considering the overlap matrix $(S_{\vec{k}})_{mn} = \langle \varphi_{m\vec{k}}|\varphi_{n\vec{k}}\rangle_V$ where V refers to an integral over one cell. As a result

$$\begin{aligned} |\widetilde{\psi}_{n\vec{k}}\rangle &= \sum_{l=1}^{N} |\varphi_{l\vec{k}}\rangle (S_{\vec{k}}^{-1/2})_{ln} = \sum_{l=1}^{N} |\varphi_{l\vec{k}}\rangle \langle \varphi_{l\vec{k}}|\varphi_{n\vec{k}}\rangle^{-1/2} = \sum_{l=1}^{N} \left[\sum_{m=1}^{J} |\psi_{m\vec{k}}\rangle \langle \psi_{m\vec{k}}|\phi_{l\vec{k}}\rangle \right] \langle \varphi_{l\vec{k}}|\varphi_{n\vec{k}}\rangle^{-1/2} = \\ &= \sum_{m=1}^{J} \left[\sum_{l=1}^{N} \langle \varphi_{l\vec{k}}|\varphi_{n\vec{k}}\rangle^{-1/2} \langle \psi_{m\vec{k}}|\phi_{l\vec{k}}\rangle \right] |\psi_{m\vec{k}}\rangle = \sum_{m=1}^{J} U_{mn}(\vec{k}) |\psi_{m\vec{k}}\rangle , \end{aligned} \tag{1.41}$$

so that the Löwdin transformation is expressed as

$$U_{mn}(\vec{k}) = \sum_{l=1}^{N} \langle \phi_{l\vec{k}}|\phi_{n\vec{k}}\rangle^{-1/2} \langle \psi_{m\vec{k}}|\phi_{l\vec{k}}\rangle . \tag{1.42}$$

Wavefunctions $|\widetilde{\psi}_{n\vec{k}}\rangle$ in Eq. (1.41) represent the Löwdin orthonormalized Bloch-like states. As a result, the new states $|\widetilde{\psi}_{n\vec{k}}\rangle$ are related to the original Bloch-states $|\psi_{n\vec{k}}\rangle$ by a unitary transformation. They represent the Bloch-like states the most similar to the atomic orbitals.

Once we have obtained Bloch-like states with a smooth gauge in $\vec{k}$, these states can be substituted into Eq. (1.23).

The Bloch-like states $\widetilde{\psi}_{n\vec{k}}$ are uniquely defined by the selected manifold and the trial orbitals. According to Eq. (1.40), if a phase factor is introduced by employing a Bloch function $|\psi'_{m\vec{k}}\rangle = e^{i\beta}|\psi_{m\vec{k}}\rangle$, it is canceled out by the term $\langle \psi'_{m\vec{k}}|$. As a result, the gauge freedom is eliminated.

As an example, Fig. 1.6(c) illustrates some of the Wannier functions obtained for $\mathrm{SrTiO_3}$, specifically two p-orbitals of oxygen atoms and a t_{2g}-orbitals centred on titanium atoms, obtained through projections onto a basis set comprising trial SIESTA orbitals, denoted as $|\phi_\mu\rangle$. This basis set enables the construction of well-localized Wannier functions that span the top nine occupied valence bands and the bottom t_{2g} empty conduction bands.

1.4.6 Maximally localized Wannier functions

Marzari and Vanderbilt developed a *localization criterion* in order to define a methodology to generate well-defined Wannier functions eliminating the phase indeterminacy [68]. The methodology is based on the minimization of the spread of the Wannier functions. For a Wannier function $\chi_{n\vec{0}}$ the spread of a Wannier function represents the mean square variation of the Wannier electron density away from its mean position (centres of the Wannier functions), $\vec{r}_n = \langle \chi_{n\vec{0}}|\hat{r}|\chi_{n\vec{0}}\rangle$,

$$\Omega_{n\vec{0}} = \langle \chi_{n\vec{0}}|r^2|\chi_{n\vec{0}}\rangle - \langle \chi_{n\vec{0}}|\hat{r}|\chi_{n\vec{0}}\rangle^2 . \tag{1.43}$$

The localization functional is defined as the sum of the quadratic spreads of the Wannier functions of the simulation cell ($\vec{R}_n = 0$) around their centers, resulting

$$\Omega = \sum_n \left[\langle \chi_{n\vec{0}}|r^2|\chi_{n\vec{0}}\rangle - \langle \chi_{n\vec{0}}|\hat{r}|\chi_{n\vec{0}}\rangle^2 \right] . \tag{1.44}$$

In order to minimize the spread, Ω, the first step consist on expressing the functional as a function of Bloch-orbitals by means of the expression Eq. (1.37). This step is perform as a post-processing methodology. The ground state Bloch functions $|\psi_{m\vec{k}}\rangle$ are obtained in a self consistency convergence in a conventional electronic structure calculation. The minimization of the spread is focused on the definition of the unitary rotations $U_{mn}(\vec{k})$ (parameters of the minimization) which minimizes the spread of the Wannier functions.

1.4.7 Mapping hamiltonian matrix elements from SIESTA basis to Wannier basis

To determine the energy bands of a system, it is necessary to solve the $N_{\text{orb}} \times N_{\text{orb}}$ complex eigenvalue problem defined in Eq. (1.21) for every $\vec{k}$-point in the reciprocal space mesh. Here, N_{orb} denotes the total number of atomic orbitals in the unit cell, which directly determines the dimensionality of the problem. Solving this eigenvalue equation yields two key results: the eigenvalues, representing the energy levels of the system, and the eigenvectors, corresponding to the Bloch wavefunctions associated with these energy levels. The Bloch wavefunctions are expressed as linear combinations of the atomic orbitals within the SIESTA basis set, capturing the periodic nature of the crystal and the quantum mechanical behavior of the electrons. According to Eq. (43) of Ref. [38],

$$|\psi_{m\vec{k}}\rangle = \sum_{\mu'} e^{i\vec{k}\cdot\vec{R}_{\mu'}} c_{\mu' m}(\vec{k}) \, |\phi_{\mu'}\rangle , \tag{1.45}$$

where the sum in μ' extends over all the atomic orbitals in space, and $c_{\mu' n} \equiv c_{\mu n}$ if $\mu' \equiv \mu$ (this notation indicates that $\phi_{\mu'}$ and ϕ_μ are equivalent orbitals, related by a lattice vector). Here replacing Eq. (1.45) into Eq. (1.37), and making use of Eq. (21) of Ref. [65],

$$\frac{V}{(2\pi)^3} \int d\vec{k} \rightarrow \frac{1}{N} \sum_{\vec{k}} \tag{1.46}$$

then

$$
\begin{aligned}
|\chi_{n\vec{R}}\rangle &= \frac{1}{N}\sum_{\vec{k}} e^{-i\vec{k}\cdot\vec{R}} \sum_{m=1}^{J} U_{mn}(\vec{k}) \left(\sum_{\mu'} e^{i\vec{k}\cdot\vec{R}_{\mu'}} c_{\mu' n}(\vec{k})\,|\phi_{\mu'}\rangle\right) \\
&= \sum_{\mu'} \left(\sum_{\vec{k}} \sum_{m=1}^{J} \frac{1}{N} U_{mn}(\vec{k}) e^{i\vec{k}\cdot(\vec{R}_{\mu'}-\vec{R})} c_{\mu' n}(\vec{k})\right) |\phi_{\mu'}\rangle \\
&= \sum_{\mu'} T_{\mu' n\vec{R}}\,|\phi_{\mu'}\rangle .
\end{aligned} \tag{1.47}
$$

The preceding expression has been derived by generalizing to a manifold consisting of J bands.

From hamiltonian matrix elements expressed in numerical atomic orbitals it is possible to obtain the hamiltonian matrix elements expanded in Wannier functions basis set.

$$
\begin{aligned}
\langle \chi_{m\vec{R}'}|H|\chi_{n\vec{R}}\rangle &= \left\langle \sum_{\mu'} T_{\mu' m\vec{R}'}\phi_{\mu'} \middle| H \middle| \sum_{\nu'} T_{\nu' n\vec{R}}\phi_{\nu'} \right\rangle \\
&= \sum_{\mu'}\sum_{\nu'} T^{*}_{\mu' m\vec{R}'} T_{\nu' n\vec{R}} \langle \phi_{\mu'}|H|\phi_{\nu'}\rangle \\
&= \sum_{\mu'}\sum_{\nu'} T^{*}_{\mu' m\vec{R}'} T_{\nu' n\vec{R}} H_{\mu'\nu'} .
\end{aligned} \tag{1.48}
$$

1.5 Tight binding in Wannier functions

As previously mentioned, the band structure can be represented within a Wannier function basis. In this section, we address this formulation following the approach presented in Ref. [63]. For the sake of simplicity, we assume again the case of isolated bands, a manifold composed by one band. It can be proved that

$$
\langle \chi_{n\vec{0}}|H|\chi_{n\vec{R}}\rangle = E_{n\vec{R}}, \tag{1.49}
$$

where

$$
E_{n\vec{R}} = \frac{V_{\mathrm{cell}}}{(2\pi)^3} \int_{\mathrm{BZ}} e^{-i\vec{k}\cdot\vec{R}} E_{n\vec{k}} \mathrm{d}^3 k, \tag{1.50}
$$

so the diagonal elements between Wannier functions are the coefficients in the Fourier expansion of the band energy. From a practical point of view, these matrix elements can be obtained from the knowledge of the band structure, the $\vec{k}$-points sampling and the volume of the unit cell, V_{cell}. These operations are done within the WANNIER90 code. Deeming a physical perspective, Eq. (1.49) represents the same matrix elements in real space as shown in Fig. 1.3, but in a basis of Wannier functions and not in atomic orbitals. Then, applying the inverse Fourier transform we can compute the band structure in the reciprocal space knowing the Hamiltonian matrix elements written in a basis of Wannier functions,

$$
E_{n\vec{k}} = \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} E_{n\vec{R}}. \tag{1.51}
$$

Analyzing Eq. (1.51), it is formally the same as the Hamiltonian matrix expressed in a basis of Bloch-like basis functions. In other words, the Wannier functions provide an exact tight-binding representation of the dispersion $E_{n\vec{k}}$ of band n where tight binding on-site energy is obtained by $E_{n\vec{R}=0}$ and the hoppings are given by $E_{n\vec{R}}$ to its neighbor located at $\vec{R}$.

Equation (1.51) is exact and, by construction, accurately reproduces the band dispersion. From a computational point of view, performing an infinite summation is not feasible. However, owing to the localized character of the Wannier functions, the hopping matrix elements decay rapidly with increasing distance. As a result, only a limited number of hopping terms need to be retained in practice,

$$E_{n\vec{k}} \approx \sum_{|\vec{R}|<\delta r_{\mathrm{h}}} e^{i\vec{k}\cdot\vec{R}} E_{n\vec{R}}, \tag{1.52}$$

where δr_{h} is a radius cutoff designed as *cutoff-h*. This approximation is systematically convergent. In order to obtain a better representation of the band diagram, the increasing of the δr_h value is enough being the asymptotically limit of the bands the ones obtained by the Bloch orbitals.

The expression presented in Eq. (1.51) was derived from the Wannierization of an isolated band. However, in many practical scenarios, one must consider the Wannierization of a group of entangled bands, particularly when band crossings occur within a given manifold, as discussed in Sec. 1.4.4. In such cases, the interpretation of Eq. (1.51) can be extended to encompass entire band manifolds, following the formalism outlined in Ref. [63]. In the case of the study of a manifold composed for a J bands, the Eq. 1.49 is also satisfied.

However, as discussed in Sec. 1.4.4, these bands may be associated to Bloch functions which exhibit non-analytic behavior. Consequently, it becomes necessary to construct analytic Bloch-like functions, $\widetilde{\psi}_{m\vec{k}}$ to enable a well-defined Wannier representation. In such basis set, the Hamiltonian matrix is not diagonal,

$$\widetilde{H}_{mn\vec{k}} = \langle \widetilde{\psi}_{m\vec{k}} | H | \widetilde{\psi}_{n\vec{k}} \rangle . \tag{1.53}$$

The terms $\widetilde{H}_{mn\vec{k}}$ represent the matrix elements of a $J \times J$ Hamiltonian matrix. Then, instead of Eq. (1.51), we compute the energies solving the secular equation

$$\det \left[\widetilde{H}_{mn\vec{k}} - E_{n\vec{k}} \right] = 0. \tag{1.54}$$

constructed by considering

$$\widetilde{H}_{mn\vec{k}} = \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} H_{mn\vec{R}}, \tag{1.55}$$

where

$$H_{mn\vec{R}} = \langle \chi_{m\vec{0}} | H | \chi_{n\vec{R}} \rangle . \tag{1.56}$$

At this point, in order to study the relation between the Hamiltonian matrix elements $H_{mn\vec{R}}$ and the energies in Eq. (1.49). We begin by considering the basis set of Bloch functions, denoted as $\{\psi_{n\vec{k}}\}$ where $n = 1, \ldots, J$. This basis is assumed to be orthonormal within the simulation cell, with a volume V_{cell}. As a result, the corresponding periodic parts of the Bloch functions, $\{u_{n\vec{k}}\}$, also satisfy the orthonormality condition within the same domain,

$$\int_{V_{\text{cell}}} \left|\psi_{n\vec{k}}\right|^2 d^3r = \int_{V_{\text{cell}}} \left|u_{n\vec{k}}\right|^2 d^3r = 1. \tag{1.57}$$

In the following reasoning has been taken into account the periodicity of $\{u_{n\vec{k}}\}$ functions,

$$\begin{aligned}
\langle \psi_{m\vec{k}} | \psi_{n\vec{k}'} \rangle &= \int_{\mathbb{R}} e^{-i(\vec{k}-\vec{k}')\cdot\vec{r}} u^*_{m\vec{k}}(\vec{r}) u_{n\vec{k}'}(\vec{r}) d^3r \\
&= \sum_{\vec{R}} \int_{\text{cell}} e^{-i(\vec{k}-\vec{k}')\cdot(\vec{R}+\vec{r})} u^*_{m\vec{k}}(\vec{r}+\vec{R}) u_{n\vec{k}'}(\vec{r}+\vec{R}) d^3r \\
&= \sum_{\vec{R}} \int_{\text{cell}} e^{-i(\vec{k}-\vec{k}')\cdot(\vec{R}+\vec{r})} u^*_{m\vec{k}'}(\vec{r}) u_{n\vec{k}}(\vec{r}) d^3r \\
&= \left(\sum_{\vec{R}} e^{-i(\vec{k}-\vec{k}')\cdot\vec{R}}\right) \int_{\text{cell}} e^{-i(\vec{k}-\vec{k})\cdot\vec{r}} u^*_{m\vec{k}}(\vec{r}') u_{n\vec{k}}(\vec{r}) d^3r \\
&= \frac{(2\pi)^3}{V_{\text{cell}}} \, \delta(\vec{k}-\vec{k}') \int_{\text{cell}} u^*_{m\vec{k}'}(\vec{r}) u_{n\vec{k}}(\vec{r}) \; d^3r = \\
&= \frac{(2\pi)^3}{V_{\text{cell}}} \, \delta(\vec{k}-\vec{k}') \int_{\text{cell}} u^*_{m\vec{k}'}(\vec{r}) u_{n\vec{k}}(\vec{r}) \; d^3r = \frac{(2\pi)^3}{V_{\text{cell}}} \, \delta(\vec{k}-\vec{k}') \, \delta_{mn}.
\end{aligned} \tag{1.58}$$

Once this characteristic of the Bloch function is fixed, we compute the Hamiltonian matrix elements expressed in a Wannier functions basis set

$$H_{mn\vec{R}} = \langle \chi_{m\vec{0}} | H | \chi_{n\vec{R}} \rangle, \tag{1.59}$$

$$|\chi_{n\vec{R}}\rangle = \frac{V_{\text{cell}}}{(2\pi)^3} \int_{\text{BZ}} e^{-i\vec{k}\cdot\vec{R}} |\widetilde{\psi}_{n\vec{k}}\rangle d\vec{k} = \frac{V_{\text{cell}}}{(2\pi)^3} \int_{\text{BZ}} e^{-i\vec{k}\cdot\vec{R}} \sum_{p=1}^{J} U_{pn}(\vec{k}) |\psi_{p\vec{k}}\rangle d\vec{k}, \tag{1.60}$$

$$|\chi_{m\vec{0}}\rangle = \frac{V_{\text{cell}}}{(2\pi)^3} \int_{\text{BZ}} |\widetilde{\psi}_{m\vec{k}}\rangle d\vec{k} = \frac{V_{\text{cell}}}{(2\pi)^3} \int_{\text{BZ}} \sum_{l=1}^{J} U_{lm}(\vec{k}') |\psi_{l\vec{k}}\rangle d\vec{k}', \tag{1.61}$$

$$\begin{aligned}
\langle \chi_{m\vec{0}} | H | \chi_{n\vec{R}} \rangle &= \frac{V^2_{\text{cell}}}{(2\pi)^6} \int_{\text{BZ}} \int_{\text{BZ}} e^{-i\vec{k}\cdot\vec{R}} \sum_{l=1}^{J} \sum_{p=1}^{J} U^{\dagger}_{ml}(\vec{k}') U_{pn}(\vec{k}) \langle \psi_{l\vec{k}'} | H | \psi_{p\vec{k}} \rangle d\vec{k}' \; d\vec{k} = \\
&= \frac{V^2_{\text{cell}}}{(2\pi)^6} \int_{\text{BZ}} \int_{\text{BZ}} e^{-i\vec{k}\cdot\vec{R}} \sum_{l=1}^{J} \sum_{p=1}^{J} U^{\dagger}_{ml}(\vec{k}') U_{pn}(\vec{k}) E_{p\vec{k}} \langle \psi_{l\vec{k}'} | \psi_{p\vec{k}} \rangle d\vec{k}' \; d\vec{k} = \\
&= \frac{V^2_{\text{cell}}}{(2\pi)^6} \int_{\text{BZ}} \int_{\text{BZ}} e^{-i\vec{k}\cdot\vec{R}} \sum_{l=1}^{J} \sum_{p=1}^{J} U^{\dagger}_{ml}(\vec{k}') U_{pn}(\vec{k}) E_{p\vec{k}} \frac{(2\pi)^3}{V_{\text{cell}}} \delta(\vec{k}-\vec{k}') \delta_{lp} \; d\vec{k}' \; d\vec{k} = \\
&= \frac{V_{\text{cell}}}{(2\pi)^3} \int_{\text{BZ}} e^{-i\vec{k}\cdot\vec{R}} \sum_{p=1}^{J} U^{\dagger}_{mp}(\vec{k}) U_{pn}(\vec{k}) E_{p\vec{k}} d\vec{k}.
\end{aligned} \tag{1.62}$$

Finally, we obtain the relation between the Hamiltonian matrix elements expressed in a Wannier functions basis set and the energy bands,

$$H_{mn\vec{R}} = \frac{V_{\text{cell}}}{(2\pi)^3} \int_{\text{BZ}} e^{-i\vec{k}\cdot\vec{R}} \sum_{p}^{J} U^{\dagger}_{mp}(\vec{k}) U_{pn}(\vec{k}) E_{p\vec{k}} d\vec{k}. \tag{1.63}$$

If the Brillouin zone is discretized, applying Eq. (1.46),

$$H_{mn\vec{R}} = \frac{1}{N}\sum_{j=0}^{N-1} e^{-i\vec{k}\cdot\vec{R}} \sum_{p}^{J} U^{\dagger}_{mp}(\vec{k}) U_{pn}(\vec{k}) E_{p\vec{k}} = \frac{1}{N}\sum_{j=0}^{N-1} e^{-i\vec{k}\cdot\vec{R}} \widetilde{H}_{mn\vec{k}}. \tag{1.64}$$

Finally, equivalently with Eq. (1.52) a δr_h cutoff can be applied,

$$H_{mn\vec{k}} \approx \sum_{|\vec{R}|<\delta r_{\mathrm{h}}} e^{i\vec{k}\cdot\vec{R}} H_{mn\vec{R}}. \tag{1.65}$$

An important consideration related to the Wannierization procedure is that we can restrict our bands to the treatment of the set of active electrons in our problem. That means we can wannierize and keep in our calculations only the bands that really play a role in the physical problem under study. For example, in the case of the polaron existence which we are interested in, it will be enough to keep the *top-valence* and the *bottom-conduction* bands related to the O $2p$ and Ti $t_{2\mathrm{g}}$ characters respectively. This leads to a tight binding model integrated by 12 bands per unit cell, in contrast to first-principles calculations, performed with 72 bands per unit cell, arising a high reduction of the computational cost.

Finally, we must hightlight that from the point of view of the band manifold study, the Wannier functions maintain the orthonormality already expressed in the isolated band case, Eq. (1.30). As a result, for a basis set composed by Wannier functions the overlap matrix S is the identity,

$$S_{mn}(\vec{R}) = \langle \chi_{m\vec{R}'} | \chi_{n\vec{R}} \rangle = \delta_{\vec{R}'\vec{R}} \delta_{mn}. \tag{1.66}$$

1.6 Disentanglement

In the preceding sections, we described the procedure for Wannierize isolated bands or groups of bands forming a manifold that is energetically separated from other bands across the entire Brillouin zone by well-defined energy gaps [65]. However, in certain systems, even when considering entangled bands within manifolds, achieving such separation is not always feasible.

One illustrative example is the case of $SrTiO_3$, a material studied in this work. As previously mentioned, the focus is on the valence bands with oxygen p-character and the conduction bands with $t_{2\mathrm{g}}$-character. Analyzing the band structure, as shown in Fig. 1.6(a), it is evident that the valence band manifold does not overlap with any other bands across the Brillouin zone. On the other hand, the conduction bands exhibit a different behavior: although the titanium $t_{2\mathrm{g}}$ bands appear isolated at first glance, they cross in energy with the titanium e_{g} bands. This crossing becomes more pronounced near the Γ-point and its vicinity when examining the band structure of a $2 \times 2 \times 2$ supercell, as depicted in Fig. 1.8. This figure shows the fat bands plot of the conduction band, where the width of the band is modulated based on the contribution of a particular orbital to that band at each $\vec{k}$-point. Due to the crossing between $t_{2\mathrm{g}}$ and e_{g} bands, disentanglement [69] is required during Wannierization to separate the $t_{2\mathrm{g}}$ bands from the e_{g} bands within the energy window of interest.

When bands are entangled, the target bands for Wannierization overlap with other states above or below the desired energy range, leading to hybridization. In $SrTiO_3$, the top eigenvalues of the t_{2g} bands hybridizes with the lower-energy e_g states. Addressing such situations requires a careful selection of the energy window and the associated J-dimensional Bloch manifold, which corresponds to the states we aim to wannierize and from which J Wannier functions will be constructed. Consequently, selecting the energy window for the Wannierization of the bottom of the conduction band requires performing a *disentanglement* procedure to separate the e_g bands from the t_{2g} bands.

The disentanglement procedure [69] involves constructing J smooth Bloch-like states $|\widetilde{\psi}_{n\vec{k}}\rangle$ by combining contributions from all bands within the selected energy window, which is referred to as the "outer window". At each k-point, the $\mathscr{J}_{\vec{k}}$ Bloch states that fall within this energy range are identified, and a unitary transformation is applied to these states to produce the desired number of smooth Bloch-like states, corresponding to the number of Wannier functions being sought. The transformation is expressed as,

$$|\widetilde{\psi}_{n\vec{k}}\rangle = \sum_{m=1}^{\mathscr{J}_{\vec{k}}} U_{mn}(\vec{k})\, |\psi_{m\vec{k}}\rangle, \tag{1.67}$$

where $U(\vec{k})$ is a rectangular $\mathscr{J}_{\vec{k}} \times J$ matrix. This matrix is determined by ensuring that the resulting states are as smooth as possible across $\vec{k}$-space, according to Ref. [65]. By carefully tailoring this process, the disentanglement ensures that the target bands are extracted cleanly, free from unwanted hybridization with nearby states.

In the $SrTiO_3$ system, the outer energy window is defined to encompass both the t_{2g} and e_g bands. The upper and lower bounds of this energy window are chosen as illustrated in Fig. 1.7.

To refine the disentanglement, an additional "inner window" can be defined, within which certain Bloch states are preserved exactly in the projected manifold. This approach prevents unwanted mixing of higher-energy states that possess some t_{2g}-character into the t_{2g} manifold. The energy window is illustrated in Fig. 1.7.

The projection process involves selecting J trial orbitals ϕ_n, which are used as a basis to construct the smooth Bloch-like states. At each k-point, the selected Bloch states are projected onto these trial orbitals,

$$|\varphi_{n\vec{k}}\rangle = \sum_{m=1}^{\mathscr{J}_{\vec{k}}} |\psi_{m\vec{k}}\rangle \langle \psi_{m\vec{k}} \,|\, \phi_n\rangle, \tag{1.68}$$

where ϕ_n are the trial functions.

For $SrTiO_3$, selecting an energy window targeting the t_{2g} bands and projecting onto t_{2g} orbitals can inadvertently include higher-energy, undesired states with significant t_{2g}-character. By enforcing a frozen inner window, this mixing is avoided, and the t_{2g} manifold is accurately preserved. As illustrated in Fig. 1.7 in red dashed lines, this approach reproduces the three t_{2g} bands faithfully, demonstrating the effectiveness of the procedure.

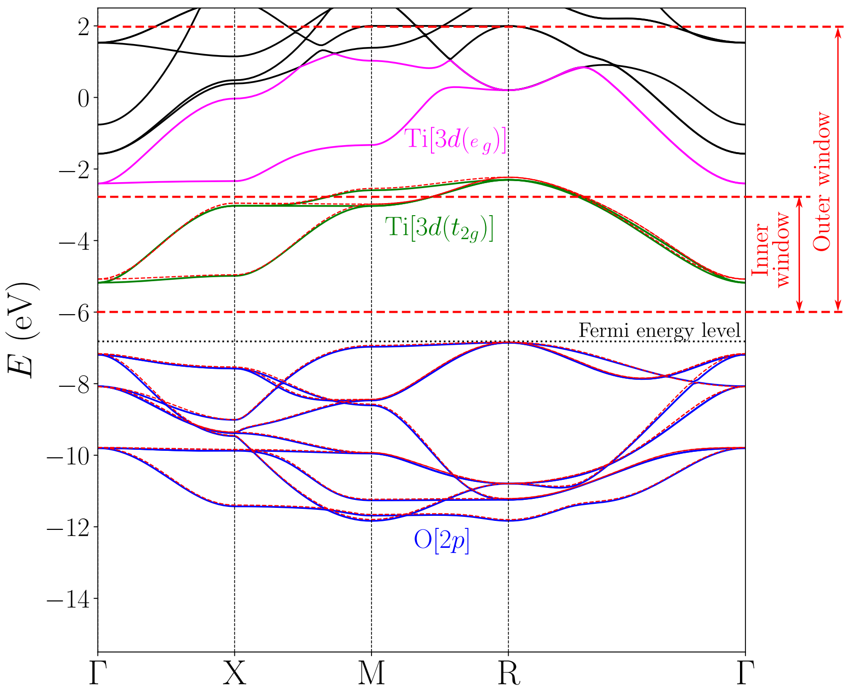


Figure 1.7: Electronic band structure of the cubic paraelectric phase of SrTiO$_3$, with emphasis on the bottom of the conduction band for the purpose of applying the disentanglement procedure. This approach is necessary to address the energy overlap between the t_{2g} bands (shown in green) and the e_g bands. The outer energy window spans the range -6.0 eV $\leq E \leq 2.0$ eV, while the inner energy window corresponds to the narrower interval -6.0 eV $\leq E \leq -3.0$ eV, within which the Bloch states $|\psi_{n\vec{k}}\rangle|$ are kept fixed. Superimposed on the DFT bands computed with SIESTA are the bands employing a Wannier functions basis set (in red), obtained applying the disentanglement procedure.

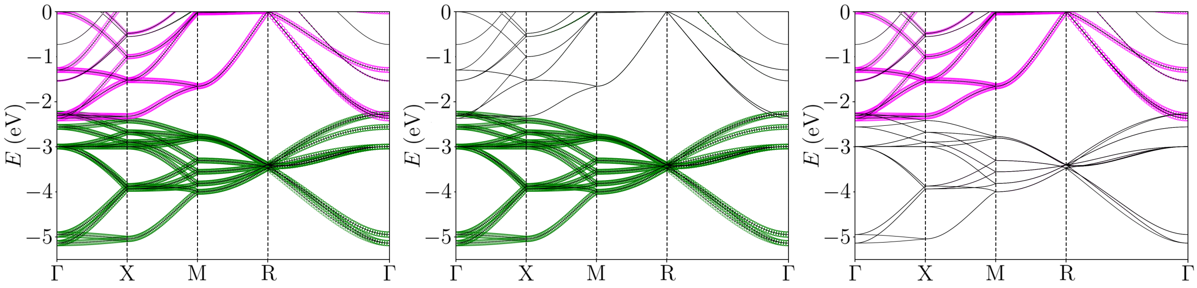


Figure 1.8: (a) Fat band representation of the conduction band for the cubic centrosymmetric phase of the SrTiO$_3$ system, computed using a $2 \times 2 \times 2$ supercell. The fat bands representation for the bottom of the conduction band is provided, highlighting 24 bands with t_{2g} character (corresponding to three d-orbitals per titanium atom across eight titanium atoms) and 16 bands with e_g character (arising from two d-orbitals per titanium atom across eight titanium atoms), shown in magenta. Panels (b) and (c) show the isolated contributions from the t_{2g} and e_g orbitals, respectively.

Chapter 2

Band structure representation using Hamiltonian matrix elements in the Wannier basis

2.1 Practical Study of WANNIER90: Internal Mechanisms of the Code

A byproduct of the Wannierization of Bloch band manifolds using the WANNIER90 code is the computation of the tight-binding matrix elements of the Hamiltonian, as well as the matrix elements of the position operator, all expressed in the Wannier function basis,

$$H_{mn\vec{R}} = \langle \chi_{m\vec{0}} | H | \chi_{n\vec{R}} \rangle. \tag{2.1}$$

In Eq. (2.1), one of the Wannier functions is centered in the home unit cell ($\chi_{m\vec{0}}$), while the other resides in a replica unit cell used to construct a supercell ($\chi_{m\vec{R}}$). The replica's origin is shifted by a Bravais vector $\vec{R}$ relative to the home unit cell. The size of the supercell is determined by the number of $\vec{k}$-points used during the Wannierization process. The final expression used to compute $H_{mn\vec{R}}$ was discussed in Sec. 1.5, more precisely in Eq. (1.64).

In this chapter, we will describe the construction of this supercell and examine the convergence of the tight-binding matrix elements concerning the sampling of reciprocal space.

2.1.1 Simulation box

Let's begin our discussion by defining the *simulation box* as the unit cell used to perform calculations in the first-principles code. This box can be the primitive cell, the conventional cell, or any other configuration deemed suitable for a specific problem by the user. Obviously, smaller simulation boxes lead to faster first-principles calculations. However, due to the application of periodic boundary conditions, a small simulation box may lack the flexibility needed to accommodate significant lattice distortions. For example, oxygen octahedra rotations in perovskites cannot be accurately modeled using a simulation box that contains only five atoms per unit cell.

To clarify these concepts and the notation used in this thesis, Fig. 2.1 presents a schematic representation of a square lattice (it can be trivially generalized to a three-dimensional lattice).

In Fig. 2.1(a), the simulation box (outlined by the red square) corresponds to the *primitive unit cell*, containing a single atom in the basis. The Bravais lattice vectors are labeled as $\vec{a}_1$ and $\vec{a}_2$. Assuming each atom hosts a single Wannier function, the simulation box contains only one Wannier function, denoted as $\chi_{m_1,\vec{0}}(\vec{r})$. According with the properties of the Wannier functions described in Sec. 1.4.2, the Wannier functions are translational images of one another. That means that for any lattice vector $\vec{R} = \ell_1\vec{a}_1 + \ell_2\vec{a}_2$, where ℓ_1 and ℓ_2 are a couple of integers, then there is an image of $\chi_{m_1,\vec{0}}(\vec{r})$ centered on $\vec{R}$, that will be denoted as $\chi_{m_1,\vec{R}}(\vec{r})$.

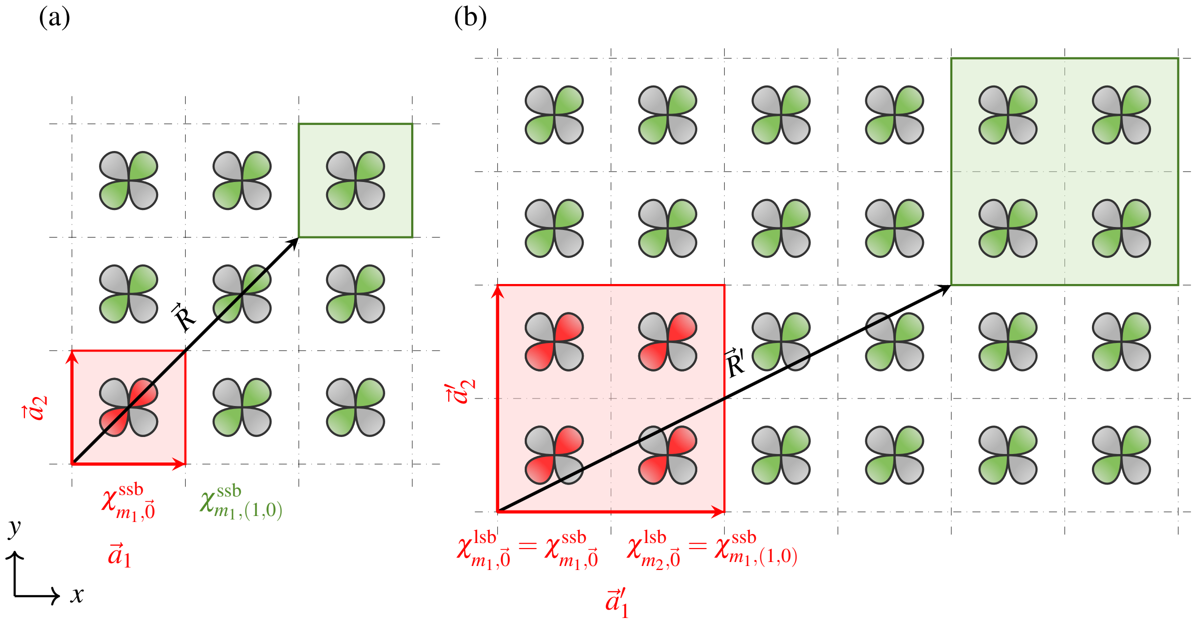


Figure 2.1: Bidimensional system where the primitive cell is a square containing a d_{xy} Wannier function. The primitive cells in the diagram are delineated by dashed lines. Two possible simulation boxes are illustrated (highlighted in red). In panel (a), the selected simulation box corresponds to the primitive cell (denoted as the *small simulation box*, *ssb*), which is characterized by lattice vectors $\vec{a}_1$ and $\vec{a}_2$ and contains a single Wannier function, denoted $\chi^{\text{ssb}}_{m_1,\vec{0}}$. In panel (b), the selected simulation box encompasses four primitive cells (denoted as the *large simulation box*, *lsb*), defined by lattice vectors $\vec{a}_1{}'$ and $\vec{a}_2{}'$. This box contains four Wannier functions: $\chi^{\text{lsb}}_{m_1,\vec{0}}$, $\chi^{\text{lsb}}_{m_2,\vec{0}}$, $\chi^{\text{lsb}}_{m_3,\vec{0}}$, and $\chi^{\text{lsb}}_{m_4,\vec{0}}$. Using explicit symmetry constraint on the unitary transformation matrices, the unit cell consistency can be rigorously preserved, so $\chi^{\text{lsb}}_{m_2,\vec{0}} = \chi^{\text{ssb}}_{m_1,(1,0)}$. For both cases, a translational image of the simulation box, located at a lattice vector $\vec{R}$ and $\vec{R}'$, is shaded in green.

In some cases, it may be advantageous to enlarge the simulation box. This is illustrated in Fig. 2.1(b), where a larger square, now containing four atoms in the basis, is selected (again outlined in red). Consequently, the simulation box now accommodates four Wannier functions, labeled as m_i, with i ranging from 1 to 4. Using explicit symmetry constraint on the unitary transformation matrices, the unit cell consistency can be rigorously preserved [70]. That means that the physical properties or quantities of solids resulted from electronic structure calculations should not depend on the specific choice of the unit cell. That means that in Fig. 2.1(b), $\chi^{\mathrm{lsb}}_{m_2,\vec{0}} = \chi^{\mathrm{ssb}}_{m_1,\vec{R}=(1,0)}$, where "lsb" and "ssb" refer to the large simulation box [Fig. 2.1(b)] and the small simulation box [Fig. 2.1(a)] respectively. In the large simulation box, the translational images are determined by lattice vectors $\vec{R}' = \ell_1\vec{a}'_1 + \ell_2\vec{a}'_2$, where $\vec{a}'_1$ and $\vec{a}'_2$ are the Bravais lattice vectors that expand the larger simulation box. In our sketch, these lattice vectors are twice as large as those of the primitive unit cell.

2.1.2 Relationship between Brillouin zone sampling and supercells in real space

The simulation boxes defined above are periodically replicated along all directions in real space, following the methodology prescribed by the corresponding Bravais lattice vectors. Therefore, the potential satisfies the imposed periodic boundary conditions,

$$V(\vec{r}) = V(\vec{r}+\vec{R}), \tag{2.2}$$

where $\vec{R}$ is a linear combination of the corresponding Bravais lattice vectors, as defined earlier. It follows directly that the system remains invariant under the action of the translation operator $T_{\vec{R}}$, which shifts the entire system by the displacement $\vec{R}$ [63]. Consequently, the eigenvectors of the Hamiltonian must obey Bloch's theorem, as discussed in Sec. 1.3. These eigenvectors can be uniquely identified by a continuous wave vector $\vec{k}$ within the first Brillouin zone and a discrete index n, referred to as the band index. According to Bloch's theorem, the eigenfunctions are given by the product of a function that exhibits the same periodicity as the simulation cell and a phase factor determined by the wave vector $\vec{k}$,

$$\psi_{n\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{n\vec{k}}(\vec{r}), \tag{2.3}$$

where $u_{n\vec{k}}(\vec{r}) = u_{n\vec{k}}(\vec{r}+\vec{R})$. From this condition, it immediately follows that

$$\psi_{n\vec{k}}(\vec{r}+\vec{a}_i) = e^{i\vec{k}\cdot\vec{a}_i} \psi_{n\vec{k}}(\vec{r}). \tag{2.4}$$

Since the system exhibits periodicity, the corresponding electronic density remains unchanged across all replicas of the simulation box. This follows from the fact that the phase factors cancel out when computing the modulus squared of the eigenfunction by multiplying it with its complex conjugate. However, the eigenfunction itself is *not* inherently periodic within the simulation box, except at the special point $\vec{k} = 0$. Nevertheless, the wave function retains periodicity within a supercell in real space. Indeed, after Eq. (2.4)

$$\psi_{n\vec{k}}(\vec{r}+p\vec{a}_i) = e^{ip\vec{k}\cdot\vec{a}_i} \psi_{n\vec{k}}(\vec{r}), \tag{2.5}$$

and therefore the wave function exhibits periodicity if $\vec{k}\cdot\vec{a}_i = 2\pi q/p$, where p and q are integers. Considering a simple one-dimensional system with a lattice constant a, this condition simplifies to $k = 2\pi q/(pa)$. For this case:

- If $p = 2$, the values of k for which ψ_{nk} is periodic are $k = 0$ ($q = 0$) and $k = \pi/a$ ($q = 1$). The next value, $k = 2\pi/a$ ($q = 2$), is equivalent to $k = 0$, as they are related by a reciprocal lattice vector. Similarly, $k = 3\pi/a$ ($q = 3$) is equivalent to $k = \pi/a$. Consequently, a k-sampling of two points within the simulation box is equivalent to sampling a supercell of the simulations box of size 2 at the Γ point.
- For a larger integer $p = 4$, the wave function remains periodic for $k = 0$, $k = \pm\pi/2a$, and $k = \pi/a$. Consequently, performing a k-sampling of four points within the original simulation box is equivalent to sampling exclusively at the Γ point in a supercell of size 4.

Based on this analysis, we conclude that investigating the simulation box with a $\vec{k}$-point sampling of N_1, N_2, N_3 points within its Brillouin zone is mathematically equivalent to employing a supercell with lattice vectors defined as $\vec{A}_i = N_i\vec{a}_i$, sampled solely at the Γ point. Consequently, the $\vec{k}$-point sampling in reciprocal space corresponds to an equivalent Born–von Kármán supercell in real space.

In the thermodynamic limit, where the sampling of the first Brillouin zone within the reciprocal unit cell becomes continuous ($N_i \to \infty$, $i = 1,2,3$), this is equivalent to performing a calculation at a single $\vec{k}$-point within a supercell in real space that extends infinitely.

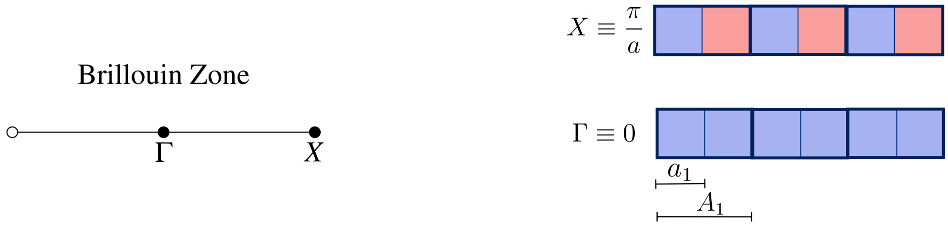


Figure 2.2: Equivalence between k-space sampling in the first Brillouin zone of the primitive unit cell and real-space simulations in a supercell. The simulation box, represented here as a periodic repetition of small squares with lattice constant a, corresponds to the primitive unit cell in this simplified illustration. The associated first Brillouin zone, with a size of $2\pi/a$, extends from $[-\pi/a, \pi/a]$. Two distinct k-points are selected to sample the one-dimensional Brillouin zone: $X \equiv \pi/a$ (upper panel) and $\Gamma = 0$ (lower panel). In the first case (X point), the wave function alternates in sign between adjacent unit cell replicas, as indicated by the color variation in the upper panel. A phase recovery occurs after shifting by two unit cells. At the Brillouin zone center (Γ point), the wave function remains unchanged in phase. Ultimately, both wave functions exhibit periodicity within a supercell consisting of two repetitions of the primitive unit cell along the x-direction, effectively forming a Born-von Kármán supercell, highlighted with thick lines.

To illustrate the transition from reciprocal space to real space, we are going to consider a one-dimensional system, taken along the x cartesian direction, and we perform different k-samplings for a squared unit cell. Firstly, we select a k-sampling of two points within the first Brillouin zone. The first one is the Γ

point, characterized by the same phase in all unit cells (modulus 2π) colored in blue in picture Fig. 2.2. The second one is the $X = \pi/a$ point, where two consecutive cells are out of phase (blue positive phase and red negative phase). Consequently, for the X point the wavefunction ψ_k is not periodic in the unit cell but the wavefunction keeps the periodicity of the lattice in a supercell composed by two unit cells. By this way, considering two k-points in the unit cell, it is equivalent to employ a Γ point in the supercell composed by two unit cells. The equivalent representation for a k-sampling of four points within the first Brillouin zone is shown in Fig. 2.3 where the equivalent Born-von Kármán supercell in real space contains four unit cells.

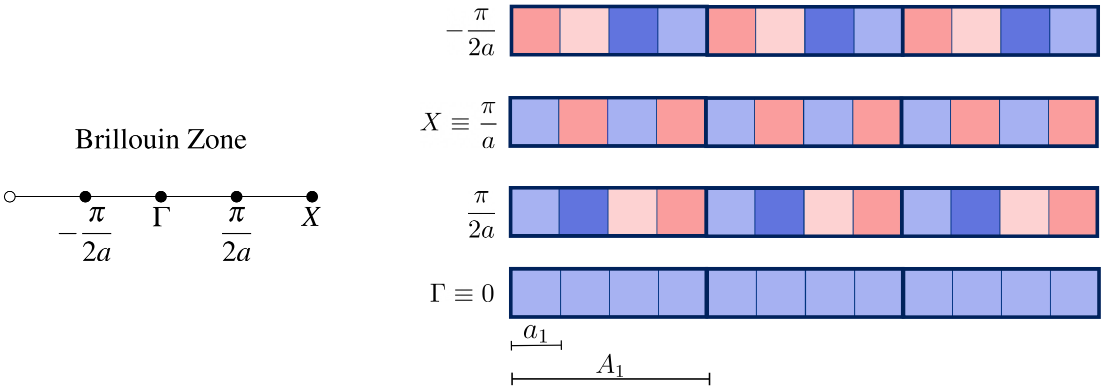


Figure 2.3: Same as in Fig. 2.2 but using four k-points to sample the first Brillouin zone: $\Gamma = 0$, $X = \pi/a$, and $k = \pm\pi/2a$. Now, the four corresponding wavefunctions exhibit periodicity within a supercell consisting of four repetitions of the primitive unit cell along the x-direction.

2.1.3 $\vec{k}$-mesh sampling in the computation of Wannier functions

In the formal definition of a Wannier function as the Fourier transform of Bloch wave functions, the integral over the entire Brillouin zone is inherently assumed, as already shown in Eq. (1.23), reproduced here for the shake of clarity

$$|\chi_{n\vec{R}}\rangle = \frac{V_{\text{cell}}}{(2\pi)^3} \int_{\text{BZ}} \mathrm{d}\vec{k}\, e^{-i\vec{k}\cdot\vec{R}} |\psi_{n\vec{k}}\rangle \,. \tag{2.6}$$

Nevertheless, from a computational perspective, directly evaluating this Brillouin zone integral is infeasible. Instead, it is commonly approximated by a summation over a finite set of selected $\vec{k}$-points. The more general strategy for constructing these $\vec{k}$-point grids was introduced by Monkhorst and Pack [71]. The accuracy of this discretization scheme is intrinsically linked to the density of the $\vec{k}$-point sampling, with finer grids systematically reducing the errors arising from the approximation of the continuous integral.

From a mathematical standpoint, this methodology effectively substitutes the integral in Eq. (2.6) with a discrete summation over the N $\vec{k}$-points composing the grid,

$$\frac{V_{\text{cell}}}{(2\pi)^3}\int_{\text{BZ}} d\vec{k} \rightarrow \frac{1}{N}\sum_{\vec{k}} w(\vec{k}), \tag{2.7}$$

ensuring computational tractability while maintaining convergence towards the continuum limit. In this substitution, V_{cell} denotes the volume of the simulation box, and w represents the weights assigned to the $\vec{k}$-points.

From a practical perspective, this implies that the first-principles code computes the wave eigenfunctions, $\psi_{n\vec{k}}$, and eigenvalues, $E_{n\vec{k}}$, only for the selected set of special $\vec{k}$-points. This set of special $\vec{k}$-points forming a grid within the first Brillouin zone is determined using the following procedure.

Considering the simulation box defined by the set of lattice vectors $\{\vec{a}_i\}$ where $i \in \{1,2,3\}$, the reciprocal lattice vectors are computed as

$$\vec{a}_1^* = \frac{2\pi}{V_{\text{cell}}}\vec{a}_2 \times \vec{a}_3, \tag{2.8}$$

$$\vec{a}_2^* = \frac{2\pi}{V_{\text{cell}}}\vec{a}_3 \times \vec{a}_1, \tag{2.9}$$

$$\vec{a}_3^* = \frac{2\pi}{V_{\text{cell}}}\vec{a}_1 \times \vec{a}_2, \tag{2.10}$$

where V_{cell} represents the volume of the simulation unit cell in real space

$$V_{\text{cell}} = \vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3) = \vec{a}_2 \cdot (\vec{a}_3 \times \vec{a}_1) = \vec{a}_3 \cdot (\vec{a}_1 \times \vec{a}_2). \tag{2.11}$$

Now, taking the $\vec{k}$-grid as (N_1, N_2, N_3), the $\vec{k}_j-$points are computed as

$$\vec{k}_j = \frac{q_1}{N_1}\vec{a}_1^* + \frac{q_2}{N_2}\vec{a}_2^* + \frac{q_3}{N_3}\vec{a}_3^* \tag{2.12}$$

where q_i are integers such that $0 \leq q_i \leq N_i - 1$.

Based on these calculations, as described in Sec. 1.5, the matrix elements of the Hamiltonian in real space can be constructed within a Wannier function basis, following the expression $H_{mn\vec{R}} = \langle \chi_{m\vec{0}} | H | \chi_{n\vec{R}} \rangle$ [Eq. (2.1)]. Here, the vector $\vec{R}$ spans the replicas of the simulation box within a supercell compatible with the chosen $\vec{k}$-point sampling. For a given $\vec{R}$, the Hamiltonian matrix in real space is a square matrix of dimension J, where J corresponds to the number of Wannier functions within the simulation box. In the asymptotic limit, where the number of $\vec{k}$-points in the Brillouin zone approaches infinity—leading to an infinitely extended supercell in real space—the Hamiltonian matrix elements $H_{mn\vec{R}}$ provide an exact tight-binding-like representation of the system's band structure. Due to the localized nature of Wannier functions, the tight-binding Hamiltonian matrix elements decay rapidly with increasing distance, meaning that only a limited number of hopping terms typically need to be retained. The interaction range considered in the simulations is determined by the $\vec{k}$-point sampling used in the Wannierization process, which in turn defines the size of the corresponding supercell in real space. These tight-binding matrix elements are directly outputted by the WANNIER90 code in a file with the suffix `_tb.dat`. This file constitutes one of the most critical components for extracting parameters to be employed in second-principles simulations.

2.1.4 Selection of $\vec{\mathbf{R}}$-vectors for the computation of $\mathbf{H}_{\mathbf{mn}\vec{\mathbf{R}}}$

To clarify the meaning of the different parameters that enter in the expression of the tight-binding matrix elements, and how the lattice vectors $\vec{R}$ are chosen, let us take an example of a two-dimensional hexagonal simulation box that coincides with the primitive unit cell, of lattice vectors $\vec{a}_1$ and $\vec{a}_2$. Let as assume that there are two Wanniers per simulation box [Fig. 2.4(a)], labeled as m and n. Let us assume a $\vec{k}$-point sampling with $N_i = 2$ for $i = 1,2$, that means a (2×2) grid in the first Brillouin zone. According to the discussion in the previous subsection, this corresponds to perform a simulation in a Born-von Kármán supercell made of two repetitions of the primitive unit cell. The Bravais lattice vectors used to displace the simulation box to build the supercell are denoted by $\vec{R}$. In the cartoon of Fig. 2.4(b), the supercell is made after replicating the home unit cell $\vec{R}_{\vec{0}}$ by the lattice vectors $\vec{R}' = 0\vec{a}_1 - 1\vec{a}_2$, $\vec{R}'' = -1\vec{a}_1 - 1\vec{a}_2$, and $\vec{R}''' = -1\vec{a}_1 + 0\vec{a}_2$. The Wanniers are periodic with the periodicity of the primitive unit cell. As explained before, a computation in the primitive cell with unit cell vectors $\vec{a}_i$ with a sampling in $\vec{k}$ is equivalent to perform the calculation in the Born-von Kármán supercell with unit cell vectors $\vec{A}_i = N_i\vec{a}_i$ at the Γ point [Fig. 2.4(b)].

In order to select the $\vec{R}$ lattice vectors for which the $H_{mn\vec{R}}$ elements will be calculated, the Wigner-Seitz cell of the Born-von Kármán supercell built before is found [Fig. 2.4(c)]. The selection of the Bravais lattice vectors $\vec{R}$ is carried out by centering the Wigner-Seitz supercell at the origin of the Born-von Kármán supercell. The chosen Bravais lattice vectors $\vec{R}$ are those that fall within the boundaries of the Wigner-Seitz supercell. In our example, the home unit cell $\vec{R}_0$ is positioned at the center of the Wigner-Seitz supercell. Additionally, the vectors $\vec{R}'$ and $\vec{R}'''$, located at the centers of the faces, are included, and are relabeled in Fig. 2.4(d) as $\vec{R}_1$ and $\vec{R}_2$, respectively.

However, the vector $\vec{R}''$ does not lie within the Wigner-Seitz supercell, and therefore, the matrix elements between the Wannier functions at the home unit cell and this replica will not be computed. As depicted in Fig. 2.4(d), additional vectors pointing toward the centers of the remaining faces of the hexagonal Wigner-Seitz supercell are also considered. These vectors are labeled $\vec{R}_j$, where $j = 3,\dots,6$. Each vector $\vec{R}_j$ has assigned a weight factor of $1/N_{\vec{R}_j}$, where $N_{\vec{R}_j}$ represents the degeneracy of the vector. The degeneracy factor $N_{\vec{R}_j}$ corresponds to the number of periodic images of the Wigner-Seitz supercell (centered at the origin of the Born-von Kármán supercell) that share the given lattice vector $\vec{R}_j$. It is particularly relevant for Bravais lattice vectors that lie on the boundary of the Wigner-Seitz supercell, including its faces, edges, or vertices. In the specific example considered, the lattice vectors $\vec{R}_j$ with $j = 1,\dots,6$ fall on the faces of the Wigner-Seitz supercell. Consequently, these vectors are shared between two Wigner-Seitz supercells, leading to a degeneracy of $N_{\vec{R}_j} = 2$ for $j = 1,\dots,6$. Conversely, the lattice vector $\vec{R}_0$ is located within the interior of the supercell, resulting in a degeneracy of $N_{\vec{R}_0} = 1$. Following the previous reasoning, if a lattice vector were to fall on a vertex in this example, the corresponding degeneracy $N_{\vec{R}}$ would be three. In a WANNIER90 calculation, the degeneracy numbers $N_{\vec{R}}$ are listed at the top of the `tb.dat` file, as depicted in Fig. 2.5.

The total sum of the weights associated with the selected Bravais lattice vectors $\vec{R}_j$ must be equal to the total number of $\vec{k}$-points. In particular, considering this example, the number of $\vec{k}$-points is $N = 4$. Regarding the Bravais lattice vectors $\vec{R}$, the weight of $\vec{R}_0$ is given by $1/N_{\vec{R}_0} = 1$, while for $\vec{R}_j$ with $j = 1, \ldots, 6$, the corresponding weights are $1/N_{\vec{R}_j} = 1/2$. Consequently, the total sum amounts to four, as expected.

Once the Bravais lattice vectors $\vec{R}$ are selected, WANNIER90 code computes the Hamiltonian matrix elements $H_{mn\vec{R}_j}$ and they are stored in the `_tb.dat` file in units of eV (see Fig. 2.5). The index m runs over the Wannier functions in the home unit cell, $\vec{R}_0$, while the index n corresponds to the Wannier functions placed in $\vec{R}_j$ for $j = 0, \ldots, 6$. In Fig. 2.4(e), the cells associated with these lattice vectors are shaded in red ($\vec{R}_0$) and yellow ($\vec{R}_j$ for $j = 1, \ldots, 6$).

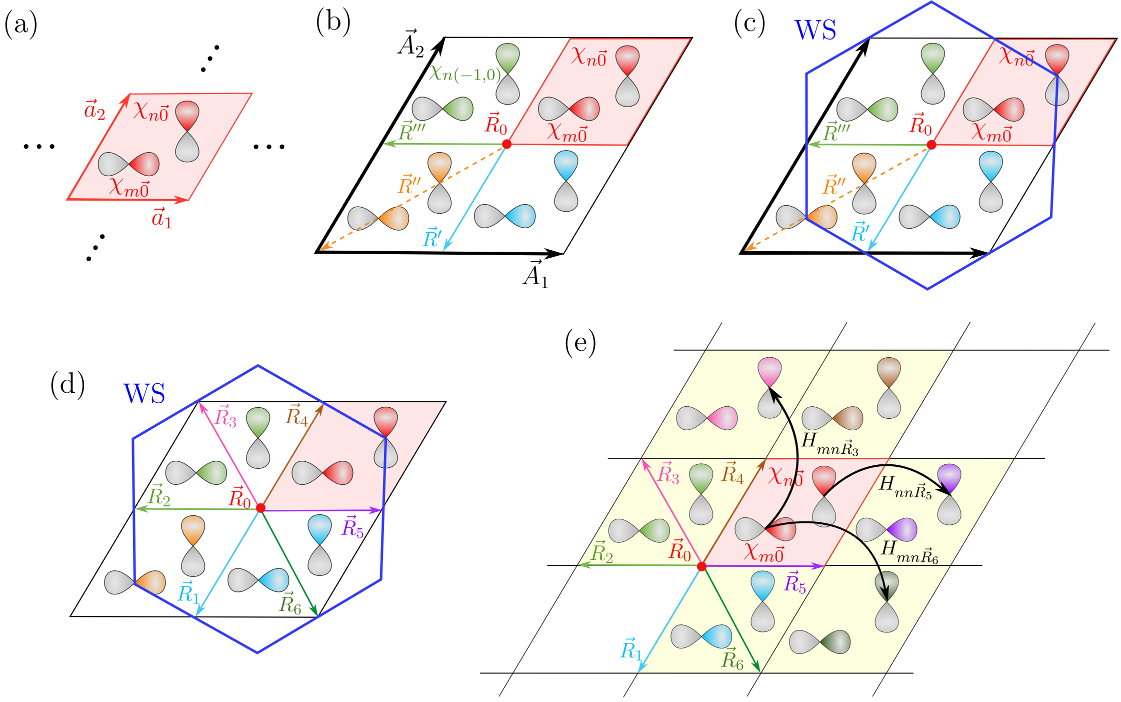


Figure 2.4: The figure illustrates the methodology employed by WANNIER90 to determine the Bravais lattice vectors $\vec{R}$, associated with periodic replicas of the simulation box, for the computation of the Hamiltonian matrix elements $H_{mn\vec{R}}$ in a two dimensional hexagonal system. In panel (a), the simulation box, which corresponds to the primitive unit cell of a two-dimensional hexagonal lattice, is depicted. The simulation box is defined by the lattice vectors $\vec{a}_1$ and $\vec{a}_2$ and contains two Wannier functions, $\chi_{m\vec{0}}$ and $\chi_{n\vec{0}}$. Panel (b) introduces the Born-von Kármán supercell constructed for a 2×2 $\vec{k}$-point sampling, with lattice vectors $\vec{A}_1$ and $\vec{A}_2$ defining its extension. The supercell is constructed by replicating the home unit cell $\vec{R}_0$ (highlighted in red) through the application of the lattice translation vectors $\vec{R}'$, $\vec{R}''$, and $\vec{R}'''$. In panel (c), the Wigner-Seitz supercell of the Born–von Kármán supercell is depicted in blue and is centered at the origin. This supercell provides a natural framework for selecting the relevant Bravais lattice vectors for the WANNIER90 calculation. The selected $\vec{R}$ vectors are those that fall within its boundaries, as illustrated in panel (d), and are denoted as $\vec{R}_j$ with $j = 0, \ldots, 6$. Finally, in panel (e), the cells associated to the $\vec{R}_j$ selected Bravais lattice vectors are highlighted in yellow. In the WANNIER90 code, they are going to be employed in the calculation of $H_{mn\vec{R}_j}$ Hamiltonian matrix elements. The WANNIER90 code computes and stores these matrix elements in the `_tb.dat` file, explicitly considering interactions between a Wannier function m centered in the home unit cell and a Wannier function n centered at $\vec{R}_j$.

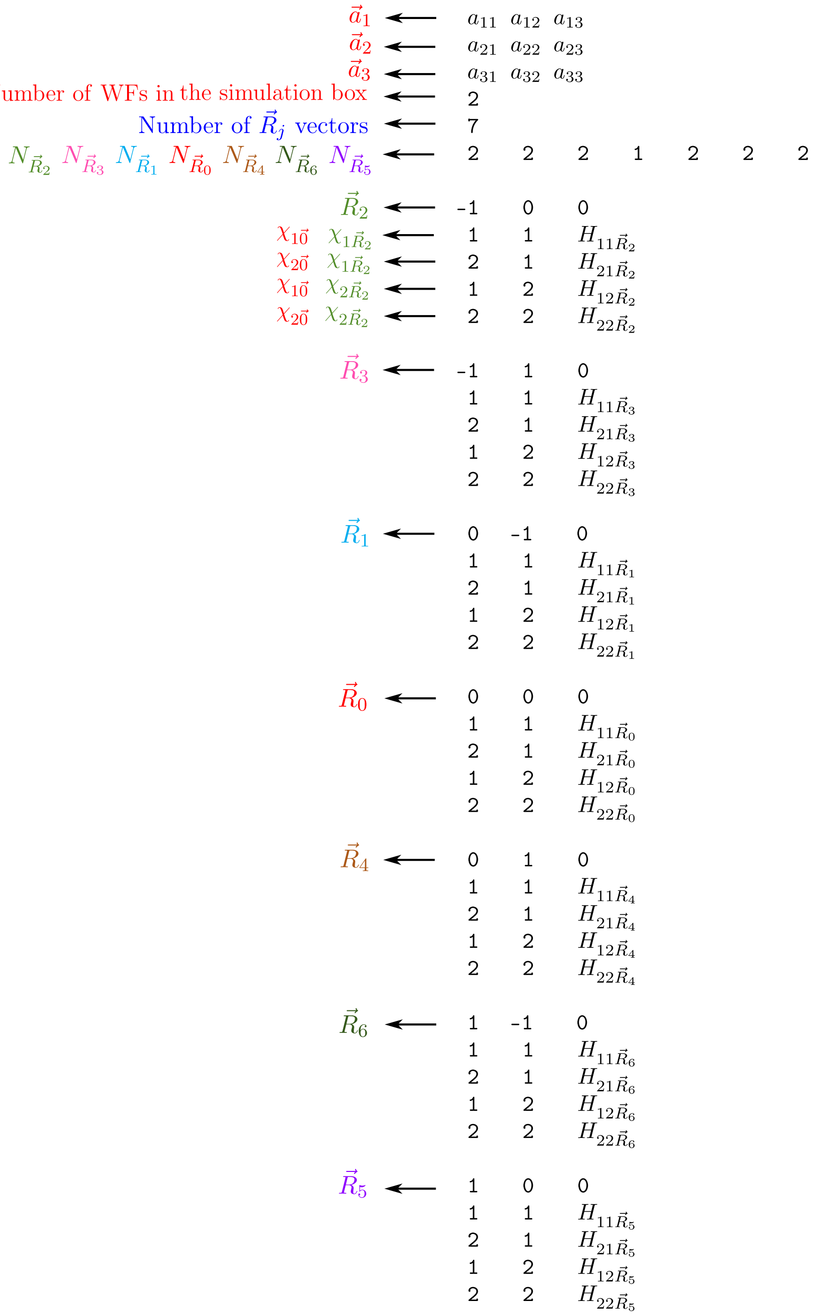


Figure 2.5: Illustration of the `_tb.dat` file corresponding to a two-dimensional hexagonal lattice system characterized by two Wannier functions in the simulation box and a $\vec{k}$-point sampling of 2×2, as depicted in Fig. 2.4. The numerical data stored in the file is shown in black, while additional annotations in different colors, consistent with Fig. 2.4, have been included to emphasize key information. According to the structure of the file, the file first lists the lattice vectors $\{\vec{a}_i\}$ with $i = 1, 2, 3$ defining the simulation box, followed by the number of Wannier functions. Next, the number of reciprocal lattice vectors $\vec{R}_j$ and their corresponding degeneracies, $N_{\vec{R}_j}$ with $j = 0, \ldots, 6$. These are presented in the same order as the Bravais lattice vectors $\vec{R}_j$ appear throughout the file. Subsequently, the coordinates of each Bravais lattice vector $\vec{R}_j$ are expressed in terms of the lattice basis $\{a_i\}$. Below each Bravais lattice vector, the Hamiltonian matrix elements $H_{mn\vec{R}_j}$ are presented, where m refers to a Wannier function centered in the home unit cell, $\chi_{m\vec{0}}$, and n corresponds to the periodic replica of Wannier function $\chi_{n\vec{0}}$ located in unit cell $\vec{R}_j$, denoted as $\chi_{n\vec{R}_j}$.

2.1.5 Invariance of the Hamiltonian matrix elements $H_{mn\vec{R}}$ under a translational vector $\vec{T}$ in the Born-von Kármán supercell

The real space Hamiltonian matrix elements $H_{mn\vec{R}}$ remain invariant under translations of the form $\vec{R} \to \vec{R}+\vec{T}$, where $\vec{T}$ is a translational vector of the Born–von Kármán superlattice. As an example, Fig. 2.6 illustrates the Hamiltonian matrix element $H_{mn\vec{R}_2}$, which represents the interaction between the Wannier function $\chi_{m\vec{0}}$ (centered at the lattice site $\vec{R}_0$, shown in red) and $\chi_{n\vec{R}_2}$ (centered at $\vec{R}_2$, shown in light green). The value of this interaction is equivalent to that of the Hamiltonian matrix elements $H_{mn\vec{R}_2+\vec{T}}$, which represent the interaction between $\chi_{m\vec{0}}$ and any Wannier function $\chi_{n\vec{R}_2+\vec{T}}$ (also shown in light green), where $\vec{T}$ is a translational vector of the Born-von Kárman superlattice.

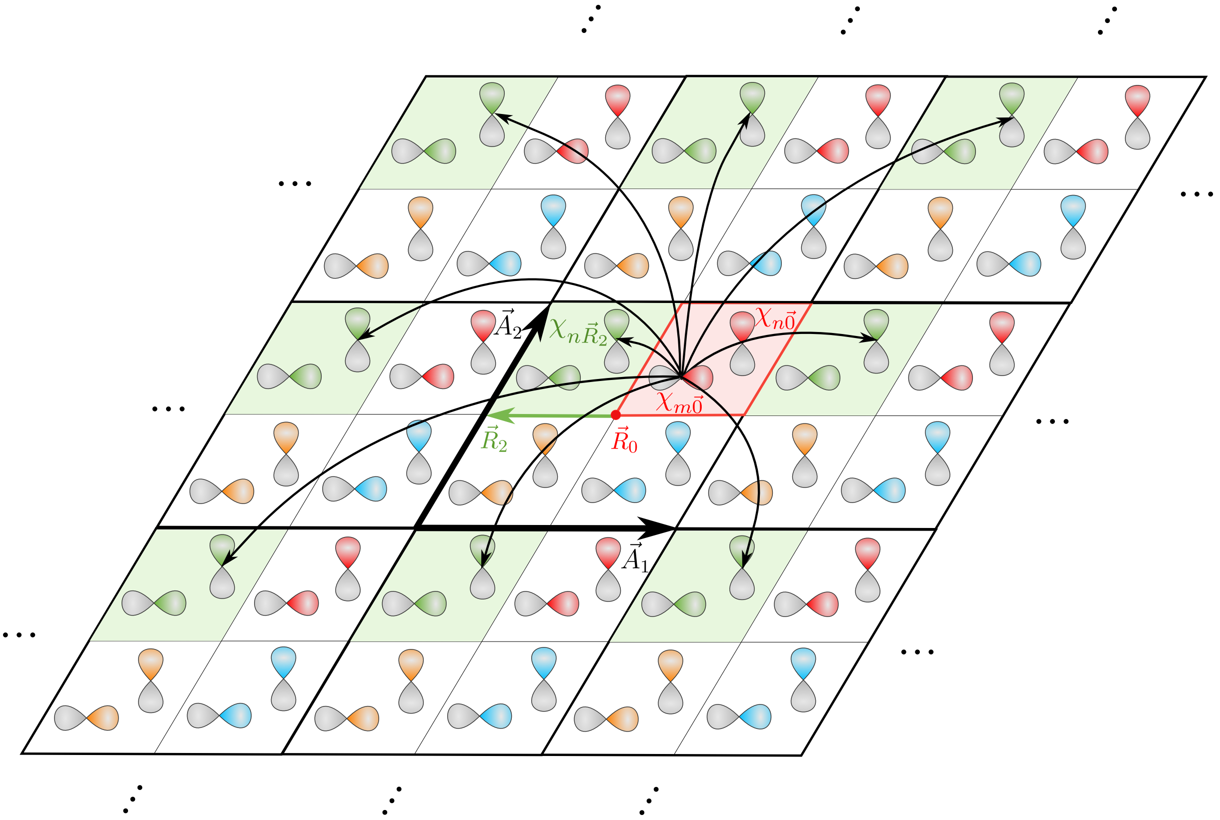


Figure 2.6: Schematic representation of the interactions between the Wannier function $\chi_{m\vec{0}}$, located in the simulation box (shaded in red), and all the periodic replicas of the Wannier function $\chi_{n\vec{R}_2}$ located at $\vec{R}' = \vec{R}_2 + \vec{T}$, generated through translations $\vec{T}$ of the Born–von Kármán supercell (defined by the lattice vectors $\vec{A}_1$ and $\vec{A}_2$). All periodic images of the cell $\vec{R}_2$ are shaded in green. As a consequence of translational symmetry, the Hamiltonian matrix elements satisfy $H_{mn\vec{R}_2} = H_{mn(\vec{R}_2+\vec{T})}$.

Now, the equality $H_{mn\vec{R}} = H_{mn\vec{R}+\vec{T}}$, as shown in Fig. 2.6, is demonstrated analytically. According to the previous notation, the translational vectors of the Born-von Kármán superlattice verify $\vec{T} \in \langle\{\vec{A}_i = N_i\vec{a}_i\}\rangle$. Consequently, a general expression for a translational vector is $\vec{T} = p_1N_1\vec{a}_1 + p_2N_2\vec{a}_2 + p_3N_3\vec{a}_3$ where p_1, p_2, p_3 are integer numbers. At this point, we compute $H_{mn\vec{R}}$ and $H_{mn\vec{R}+\vec{T}}$ employing the relation in

Eq. (1.64). Firstly, we obtain $H_{mn\vec{R}}$,

$$H_{mn\vec{R}} = \langle \chi_{m\vec{0}} | H | \chi_{n\vec{R}} \rangle = \frac{1}{N} \sum_{j=0}^{N-1} e^{-i\vec{k}_j \cdot \vec{R}} \widetilde{H}_{mn\vec{k}_j}. \tag{2.13}$$

Secondly, we compute $H_{mn\vec{R}+\vec{T}}$,

$$\begin{aligned} H_{mn\vec{R}+\vec{T}} &= \frac{1}{N} \sum_{j=0}^{N-1} e^{-i\vec{k}_j \cdot (\vec{R}+\vec{T})} \widetilde{H}_{mn\vec{k}_j} = \frac{1}{N} \sum_{j=0}^{N-1} e^{-i(\vec{k}_j \cdot \vec{R} + \vec{k}_j \cdot \vec{T})} \widetilde{H}_{mn\vec{k}_j} = \\ &= \frac{1}{N} \sum_{j=0}^{N-1} e^{-i\left[\vec{k}_j \cdot \vec{R} + \left(\frac{q_1}{N_1}\vec{a}_1^* + \frac{q_2}{N_2}\vec{a}_2^* + \frac{q_3}{N_3}\vec{a}_3^*\right) \cdot (p_1 N_1 \vec{a}_1 + p_2 N_2 \vec{a}_2 + p_3 N_3 \vec{a}_3)\right]} \widetilde{H}_{mn\vec{k}_j} = \\ &= \frac{1}{N} \sum_{j=0}^{N-1} e^{-i(\vec{k}_j \cdot \vec{R} + q_1 p_1 \vec{a}_1^* \cdot \vec{a}_1 + q_2 p_2 \vec{a}_2^* \cdot \vec{a}_2 + q_3 p_3 \vec{a}_3^* \cdot \vec{a}_3)} \widetilde{H}_{mn\vec{k}_j} = \frac{1}{N} \sum_{j=0}^{N-1} e^{-i[\vec{k}_j \cdot \vec{R} + 2\pi(q_1 p_1 + q_2 p_2 + q_3 p_3)]} \widetilde{H}_{mn\vec{k}_j} = \\ &= \frac{1}{N} \sum_{j=0}^{N-1} e^{-i\vec{k}_j \cdot \vec{R}} e^{-2\pi i(q_1 p_1 + q_2 p_2 + q_3 p_3)} \widetilde{H}_{mn\vec{k}_j} = \frac{1}{N} \sum_{j=0}^{N-1} e^{-i\vec{k}_j \cdot \vec{R}} \widetilde{H}_{mn\vec{k}_j} = H_{mn\vec{R}}. \end{aligned} \tag{2.14}$$

Here, from the first to the second line the expression for the sampled $\vec{k}_j$-points, given by Eq. (2.12), is applied alongside the expression for the translational vector $\vec{T}$. Moving from the second to the third line, the orthogonality condition $\vec{a}_i^* \cdot \vec{a}_j = 0$ for $i \neq j$ is used. In third line, the expressions for the reciprocal lattice vectors and the definition of the volume, as given in Eq. (2.11), are considered.

2.1.6 Band energy diagrams from Wannier interpolation

For a given vector of the Bravais lattice $\vec{R}$ the real-space Hamiltonian matrices $H_{\vec{R}}$, whose entries $H_{mn\vec{R}}$ were introduced in Sec. 2.1.4, are square and have dimensions equal to the number of Wannier functions within the simulation box. This corresponds to the number of wannierized states (bands), denoted as J. Examples of these square $H_{\vec{R}}$ matrices for different values of $\vec{R}$ have been given in Fig. 2.5.

The next question is whether from these real-space Hamiltonian matrix elements, $H_{mn\vec{R}}$, is possible to compute the collection of J energy bands $E_{n\vec{k}}$ at an arbitrary $\vec{k}$ point of the first Brillouin zone, where n indexes the different energy bands. The answer is positive, and the procedure is called the *Wannier interpolation*. In Fig. 2.8 we show a scheme of this methodology. The determination of these energy bands in reciprocal space requires a two-step procedure.

The first step involves the construction of the Hamiltonian matrices in reciprocal space for each $\vec{k}$, denoted as $\widetilde{H}_{\vec{k}}$. In principle, the Hamiltonian matrix elements in reciprocal space, $\widetilde{H}_{mn\vec{k}}$, are obtained by applying the inverse transformation to the real-space Hamiltonian matrix elements, $H_{mn\vec{R}}$- stored in the `_tb.dat` file [Fig. 2.5],

$$\widetilde{H}_{mn\vec{k}} = \sum_{\vec{R}} \frac{1}{N_{\vec{R}}} e^{i\vec{k} \cdot \vec{R}} H_{mn\vec{R}}, \tag{2.15}$$

where the summation runs over the set of lattice vectors $\vec{R}$ selected in Sec. 2.1.4, and the factor $1/N_{\vec{R}}$ represents the corresponding weight assigned to each vector. Remember that $\sum_{\vec{R}} 1/N_{\vec{R}} = N$, where N is the number of $\vec{k}_j$-points selected in the sampling of the first Brillouin zone.

However, it is important to note that in the construction of the Hamiltonian matrix elements $\widetilde{H}_{mn\vec{k}}$, it is essential to preserve the symmetry of the system. To ensure this symmetry, an intermediate step is introduced in the application of Eq. (2.15) during the computation of the matrix elements $\widetilde{H}_{mn\vec{k}}$. This step accounts for the fact that, as established in Sec. 2.1.5, the Hamiltonian matrix elements $H_{mn\vec{R}}$ remain invariant under a translation $\vec{T} \in \{N_i\vec{a}_i : i = 1,2,3\}$, such that $H_{mn\vec{R}} = H_{mn\vec{R}+\vec{T}}$. This can be illustrated in our simplified example, as shown in Fig. 2.7. For the sake of clarity, we shall focus on a particular matrix element between the Wannier m in the home unit cell (whose center is given by $\vec{r}_m$), and the Wannier n centered at the unit cell displaced by $\vec{R}_3$, $H_{mn\vec{R}_3}$ [represented in Fig. 2.7(a) by a black arrow]. The center of the n Wannier function in the home unit cell is denoted by $\vec{r}_n$. Here, we can clearly distinguish the matrix elements that are equivalent to $H_{mn\vec{R}_3}$ by a translational vector of the Born-von Kármán supercell. Four of them are shown and represented in Fig. 2.7 in dark blue:

- The translation vector $\vec{T}_{mn\vec{R}_3} = (0,0)$ corresponds to the original position of $\chi_{n\vec{R}_3}$, marked with a blue dot in Fig. 2.7.
- The translation vector $\vec{T}'_{mn\vec{R}_3} = 0\vec{a}_1 - 2\vec{a}_2$ shifts $\chi_{n\vec{R}_3}$ from $\vec{r}_n + \vec{R}_3$ to $\chi_{n\vec{R}''}$, which is now located at $\vec{r}_n + \vec{R}_3 + \vec{T}'_{mn\vec{R}_3} = \vec{r}_n + \vec{R}''$ [orange in Fig. 2.7(b)].
- The translation vector $\vec{T}''_{mn\vec{R}_3} = 2\vec{a}_1 - 2\vec{a}_2$ shifts $\chi_{n\vec{R}_3}$ from $\vec{r}_n + \vec{R}_3$ to $\chi_{n\vec{R}_6}$, which is now located at $\vec{r}_n + \vec{R}_3 + \vec{T}''_{mn\vec{R}_3} = \vec{r}_n + \vec{R}_6$ [dark green in Fig. 2.7(c)].
- The translation vector $\vec{T}'''_{mn\vec{R}_3} = 2\vec{a}_1 + 0\vec{a}_2$ shifts $\chi_{n\vec{R}_3}$ from $\vec{r}_n + \vec{R}_3$ to $\chi_{n\vec{R}''''}$, which is now located at $\vec{r}_n + \vec{R}_3 + \vec{T}'''_{mn\vec{R}_3} = \vec{r}_n + \vec{R}''''$ [yellow in Fig. 2.7(c)].

Then, the next step is to identify between those interactions the ones that minimizes the distances. In order to do this,

a) For each pair of Wannier functions m,n and $\vec{R}$, the optimal choice of translation vectors $\vec{T}$ is the one obtained by the minimization of the distance between the two Wannier functions m and n,

$$|\vec{r}_m - (\vec{r}_n + \vec{R} + \vec{T})|. \tag{2.16}$$

b) If there exist $\mathcal{N}_{mn\vec{R}}$ vectors $\vec{T}$ which minimize the Wannier functions distance, all of them are included in Eq. (2.15) multiplied by a weight factor $1/\mathcal{N}_{mn\vec{R}}$.

Finally, the general employed expression to compute the Hamiltonian matrix Eq. (2.15) is replaced by

$$\widetilde{H}_{mn\vec{k}} = \sum_{\vec{R}} \frac{1}{N_{\vec{R}}} \frac{1}{\mathcal{N}_{mn\vec{R}}} \sum_{l=1}^{\mathcal{N}_{mn\vec{R}}} e^{i\vec{k}\cdot\left(\vec{R}+\vec{T}^{(l)}_{mn\vec{R}}\right)} H_{mn\vec{R}}, \tag{2.17}$$

where the summation runs over the set of lattice vectors $\vec{R}$ selected in Sec. 2.1.4, and the factor $1/N_{\vec{R}}$ represents the corresponding weight assigned to each vector.

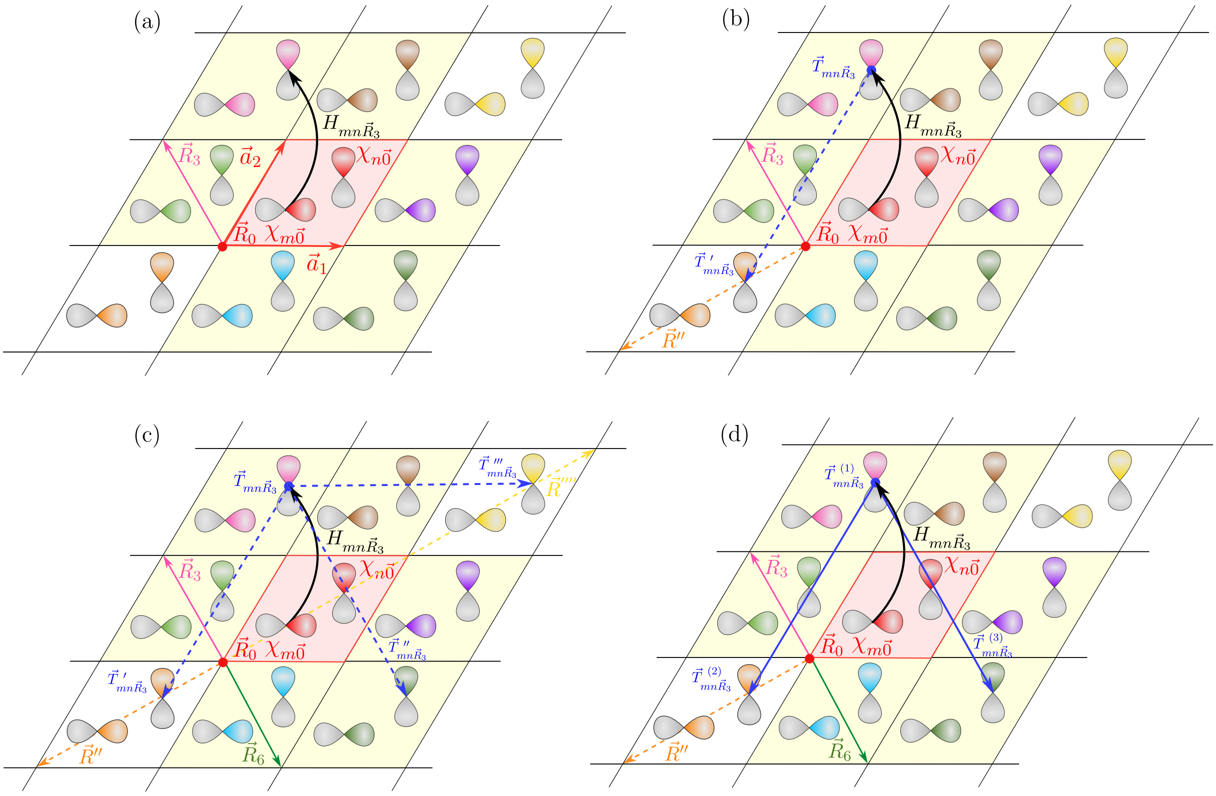


Figure 2.7: Hamiltonian matrix element $H_{mn\vec{R}_3}$ between Wannier functions $\chi_{m\vec{0}}$ and $\chi_{n\vec{R}_3}$, as well as all symmetry-equivalent functions generated by translation vectors of the Born–von Kármán supercell. (a) Illustration of the Hamiltonian matrix element between the Wannier function $\chi_{m\vec{0}}$, centered at the home unit cell (depicted in red), and the Wannier function $\chi_{n\vec{R}_3}$, centered at the lattice site $\vec{R}_3$ (shown in magenta). (b) Two translational vectors of the Born–von Kármán supercell (in blue) that generate periodic replicas of the Wannier function $\chi_{n\vec{R}_3}$. The first vector, $\vec{T}_{mn\vec{R}_3} = (0,0)$, corresponds to the original Wannier function $\chi_{n\vec{R}_3}$, while the second translational vector, $\vec{T}'_{mn\vec{R}_3} = 0\vec{a}_1 - 2\vec{a}_2$, produces the periodic image $\chi_{n\vec{R}''}$ (displayed in orange). (c) Four translational vectors within the Born–von Kármán supercell that give rise to four periodic replicas of $\chi_{n\vec{R}_3}$. These include the two vectors described in (b), along with $\vec{T}''_{mn\vec{R}_3} = 2\vec{a}_1 - 2\vec{a}_2$, which generates the replica $\chi_{n\vec{R}_6}$ (in green), and $\vec{T}'''_{mn\vec{R}_3} = 2\vec{a}_1 + 0\vec{a}_2$, leading to the replica $\chi_{n\vec{R}''''}$ (in yellow). (d) Translational vectors (represented as solid blue vectors) that define periodic images $\chi_{n\vec{R}_3+\vec{T}}$ of $\chi_{n\vec{R}_3}$ in the Born-von Kármán supercell, such that each image lies at the same spatial distance from $\chi_{m\vec{0}}$ as the original function $\chi_{n\vec{R}_3}$.

Considering our example, among the four translational vectors $\vec{T}$ which lead to the four replicas of $\chi_{n\vec{R}_3}$, we select those which minimize the distance between the Wannier functions $\chi_{m\vec{0}}$ and $\chi_{n\vec{R}_3+\vec{T}}$. Attending to Fig. 2.7(c) we can see that there are three periodic replicas of $\chi_{n\vec{R}_3}$ which are at the same distance to $\chi_{m\vec{0}}$. These Wannier functions are associated to the translation vectors $\vec{T}_{mn\vec{R}_3}$, $\vec{T}'_{mn\vec{R}_3}$ and $\vec{T}''_{mn\vec{R}_3}$, which are now relabeled as $\vec{T}^{(1)}_{mn\vec{R}_3}$, $\vec{T}^{(2)}_{mn\vec{R}_3}$ and $\vec{T}^{(3)}_{mn\vec{R}_3}$, respectively. These translational vectors are represented by solid blue vectors in Fig. 2.7(d). Finally, because we have selected three traslational vectors $\vec{T}$ which lead to a three periodic replicas of $\chi_{n\vec{R}_3}$ at the same distance of $\chi_{m\vec{0}}$, we incorporate $\mathcal{N}_{mn\vec{R}_3} = 3$. Consequently, the contribution of the $\vec{R}_3$ cell to the Hamiltonian matrix element $H_{mn\vec{k}}$ is given by

$$\frac{1}{N_{\vec{R}_3}}\frac{1}{\mathcal{N}_{mn\vec{R}_3}}\left[e^{i\vec{k}\cdot\left(\vec{R}_3+\vec{T}^{(1)}_{mn\vec{R}_3}\right)}+e^{i\vec{k}\cdot\left(\vec{R}_3+\vec{T}^{(2)}_{mn\vec{R}_3}\right)}+e^{i\vec{k}\cdot\left(\vec{R}_3+\vec{T}^{(3)}_{mn\vec{R}_3}\right)}\right]H_{mn\vec{R}_3}=\frac{1}{2}\cdot\frac{1}{3}\left[e^{i\vec{k}\cdot\vec{R}_3}+e^{i\vec{k}\cdot\vec{R}''}+e^{i\vec{k}\cdot\vec{R}_6}\right]H_{mm\vec{R}_3}. \tag{2.18}$$

This procedure is applied to all the elements in the sum over $\vec{R}$ in Eq. (2.17). Similarly, the remaining entries constituting the $\widetilde{H}_{\vec{k}}$ matrices can be obtained in an analogous manner.

Once the Hamiltonian matrices $\widetilde{H}_{\vec{k}}$ have been constructed in reciprocal space, the next step involves diagonalizing these matrices to determine the energy eigenvalues $E_{n\vec{k}}$ associated with the $\vec{k}$-points belonging to the first Brillouin zone.

At this stage, the energy bands $E_{n\vec{k}}$ can be evaluated at arbitrary $\vec{k}$-points within the first Brillouin zone. We can wonder to which extend the eigenvalues $E_{n\vec{k}}$ obtained in the tight-binding reproduce the ones obtained from a first-principles simulations. We prove below that they are exactly the same, by construction, for the $\vec{k}$-points used in the sampling to build the Wanniers, denoted as $\vec{k}_j$ in Sec. 2.1.3.

To prove the previous statement, we first show that the phase factor $e^{i\vec{k}_j\cdot\vec{R}}$ remains invariant under the transformation $\vec{R}\to\vec{R}+\vec{T}$, employed in Eq. (2.17). It is equivalent to prove the equality $\vec{k}_j\cdot\vec{T}=2\pi n$ where $n\in\mathbb{Z}$, since this condition ensures $e^{i\vec{k}_j\cdot\vec{T}}=1$. This follows immediately from the definition of the sampled $\vec{k}_j$-points as defined in Eq. (2.12). Indeed, if we consider the lattice translation vector defined as $\vec{T}=p_1N_1\vec{a}_1+p_2N_2\vec{a}_2+p_3N_3\vec{a}_3$, where $p_1,p_2,p_3\in\mathbb{Z}$, and N_1,N_2,N_3 are the numbers of $\vec{k}_j$-points sampled along each reciprocal lattice direction, then the scalar product $\vec{k}_j\cdot\vec{T}$ can be evaluated explicitly,

$$\begin{aligned}\vec{k}_j\cdot\vec{T}&=\left(\frac{q_1}{N_1}\vec{a}^*_1+\frac{q_2}{N_2}\vec{a}^*_2+\frac{q_3}{N_3}\vec{a}^*_3\right)\cdot\left(p_1N_1\vec{a}_1+p_2N_2\vec{a}_2+p_3N_3\vec{a}_3\right)=\\&=q_1p_1\vec{a}^*_1\cdot\vec{a}_1+q_2p_2\vec{a}^*_2\cdot\vec{a}_2+q_3p_3\vec{a}^*_3\cdot\vec{a}_3=2\pi(q_1p_1+q_2p_2+q_3p_3),\end{aligned} \tag{2.19}$$

where we have used the identity $\vec{a}^*_i\cdot\vec{a}_j=2\pi\delta_{ij}$. As a result, for $\vec{k}\in\{\vec{k}_j\}$, and applying the property $H_{mn\vec{R}}=H_{mn\vec{R}+\vec{T}}$, which is satisfied for any translational vector $\vec{T}$ of the Born–von Kármán supercell, the following equality holds,

$$e^{i\vec{k}_j\cdot(\vec{R}+\vec{T})}H_{mn\vec{R}+\vec{T}}=e^{i\vec{k}_j\cdot\vec{R}}H_{mn\vec{R}}. \tag{2.20}$$

Applying this result in Eq. (2.17),

$$\begin{aligned}\widetilde{H}_{mn\vec{k}_j}&=\sum_{\vec{R}}\frac{1}{N_{\vec{R}}}\frac{1}{\mathcal{N}_{mn\vec{R}}}\sum_{l=1}^{\mathcal{N}_{mn\vec{R}}}e^{i\vec{k}_j\cdot\left(\vec{R}+\vec{T}^{(l)}_{mn\vec{R}}\right)}H_{mn\vec{R}}=\sum_{\vec{R}}\frac{1}{N_{\vec{R}}}\frac{1}{\mathcal{N}_{mn\vec{R}}}\sum_{l=1}^{\mathcal{N}_{mn\vec{R}}}e^{i\vec{k}_j\cdot\vec{R}}H_{mn\vec{R}}=\\&=\sum_{\vec{R}}\frac{1}{N_{\vec{R}}}\frac{1}{\mathcal{N}_{mn\vec{R}}}\mathcal{N}_{mn\vec{R}}e^{i\vec{k}_j\cdot\vec{R}}H_{mn\vec{R}}=\sum_{\vec{R}}\frac{1}{N_{\vec{R}}}e^{i\vec{k}_j\cdot\vec{R}}H_{mn\vec{R}}.\end{aligned} \tag{2.21}$$

In particular, considering the example illustrated in Fig. 2.4, the real-space Hamiltonian matrices $H_{\vec{R}}$ are square matrices of dimension two, corresponding to the number of Wannier functions in the simulation cell (equal to the number of bands). From these matrices, it is possible to exactly reconstruct the original band energies $E_{n\vec{k}_j}$ (with $n = 1,2$) at each sampled $\vec{k}_j$-point,

- Band 1: $E_{1\vec{k}_0}$, $E_{1\vec{k}_1}$, $E_{1\vec{k}_2}$ $E_{1\vec{k}_3}$,
- Band 2: $E_{2\vec{k}_0}$, $E_{2\vec{k}_1}$, $E_{2\vec{k}_2}$ $E_{2\vec{k}_3}$.

The $\vec{k}_j$-points are computed by Eq. (2.12) leading to $\vec{k}_j \in \{(0,0),(\pi/a,0),(0,\pi/a),(\pi/a,\pi/a)\}$. For the rest of the $\vec{k}$-points of the first Brillouin zone, the energy values will be interpolated following the previous methodology.

In the case of an infinitely dense grid of $\vec{k}_j$-points, the $E_{n\vec{k}}$ values become exact for any $\vec{k}$, and the sum in Eq. (2.15) transforms into an infinite series. Due to the real-space localization of the Wannier functions, the matrix elements $H_{mn\vec{R}}$ become negligibly small when the distance between the Wannier centers exceeds a critical length L, which represents the "bandwidth" of the Wannier Hamiltonian. As a result, only a finite number of terms significantly contribute to the sum in Eq. (2.15) . This means that even with a finite $N_1 \times N_2 \times N_3$ grid, employed to reduce the computational cost, the interpolation remains accurate, as long as the "sampling rate" N_i along each cell vector $\vec{a}_i$ is large enough to satisfy the condition $N_i\,|\vec{a}_i| > 2L$, analogous to the Nyquist-Shannon sampling criterion [72].

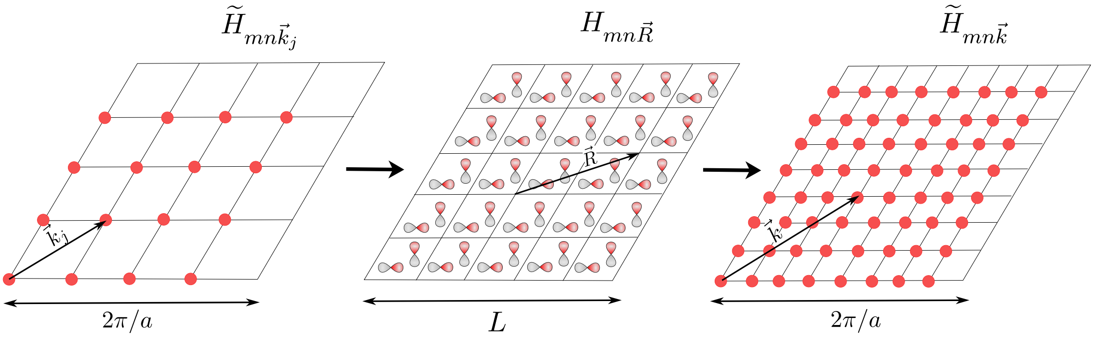


Figure 2.8: Schematic representation of the Wannier interpolation methodology. The left panel shows the grid defined in the Brillouin zone to perform the first-principles simulation. There, the Hamiltonian matrix elements $\widetilde{H}_{mn\vec{k}_j}$ are directly computed. Later, a Fourier transform is performed to express the Hamiltonian matrix elements in real space, obtaining Hamiltonian matrix elements $H_{mn\vec{R}}$. The right panel illustrates a dense mesh of $\vec{k}$-points of the Brillouin zone employed to compute by the inverse Fourier transform the $\widetilde{H}_{mn\vec{k}}$ values by means of an interpolation procedure, employing $H_{mn\vec{R}}$ elements. Adapted from Ref. [65].

2.2 Practical examples

2.2.1 Linear chain of H-atoms. The primitive unit cell.

The preceding Section detailed the methodology for obtaining interpolated energy bands $E_{n\vec{k}}$ at arbitrary $\vec{k}$-points within the first Brillouin zone, employing the real-space Hamiltonian matrix elements $H_{mn\vec{R}}$.

In this section, we apply the aforementioned method to a prototypical one-dimensional system composed of a linear chain of hydrogen atoms, each separated by a lattice constant a, and associated with one Wannier function per atom. The purpose of this analysis is to elucidate: (i) the number of hopping parameters included in the `tb.dat` file; (ii) the spatial decay behavior of these parameters as the distance between Wannier function centers increases; and (iii) the convergence of the interpolated tight-binding band structure toward the reference first-principles band structure as a function of $\vec{k}_j$-point sampling density.

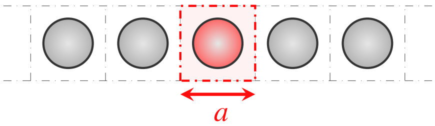


Figure 2.9: In SIESTA, the simulation box, shaded in red, contains one s-like Wannier function. The lattice constant of the simulation box is a.

The simulation cell, as implemented in the SIESTA framework (presented in Sec. 2.1.1 and highlighted in red in Fig. 2.9), consists of a single hydrogen atom and it is characterized by a lattice constant a. Bloch eigenstates are computed for various $\vec{k}_j$-point grids, after which a wannierization procedure is performed to construct the corresponding Wannier functions. The resulting Wannier function, $\chi_{m\vec{0}}$, resembles an s-like orbital localized on a hydrogen atom (represented by the yellow isosurface in Fig. 2.10). To maintain orthogonality with neighboring functions, the Wannier function includes two lobes of opposite phase on either side of the central peak (depicted as blue regions in Fig. 2.10). It is worth emphasizing that these Wannier functions are obtained by maximizing their projection onto the localized hydrogen $1s$ atomic orbital defined within the SIESTA basis set, rather than by optimizing for maximal real-space localization. Finally, the periodic Wannier function replicas $\chi_{m\vec{R}}$ corresponding to the different Bravais lattice vectors, as given by Eq. (1.27), are represented schematically (avoiding the detailed structure) by gray spheres in Fig. 2.9.

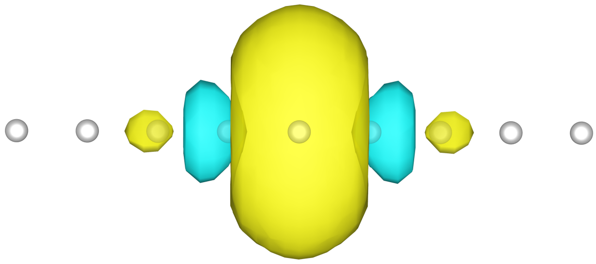


Figure 2.10: Wannier function with an s-like character, centered on a hydrogen atom within a linear chain. The different colors indicate the sign of the wavefunction phase. The presence of nodes is necessary to satisfy the orthonormality condition of the basis set.

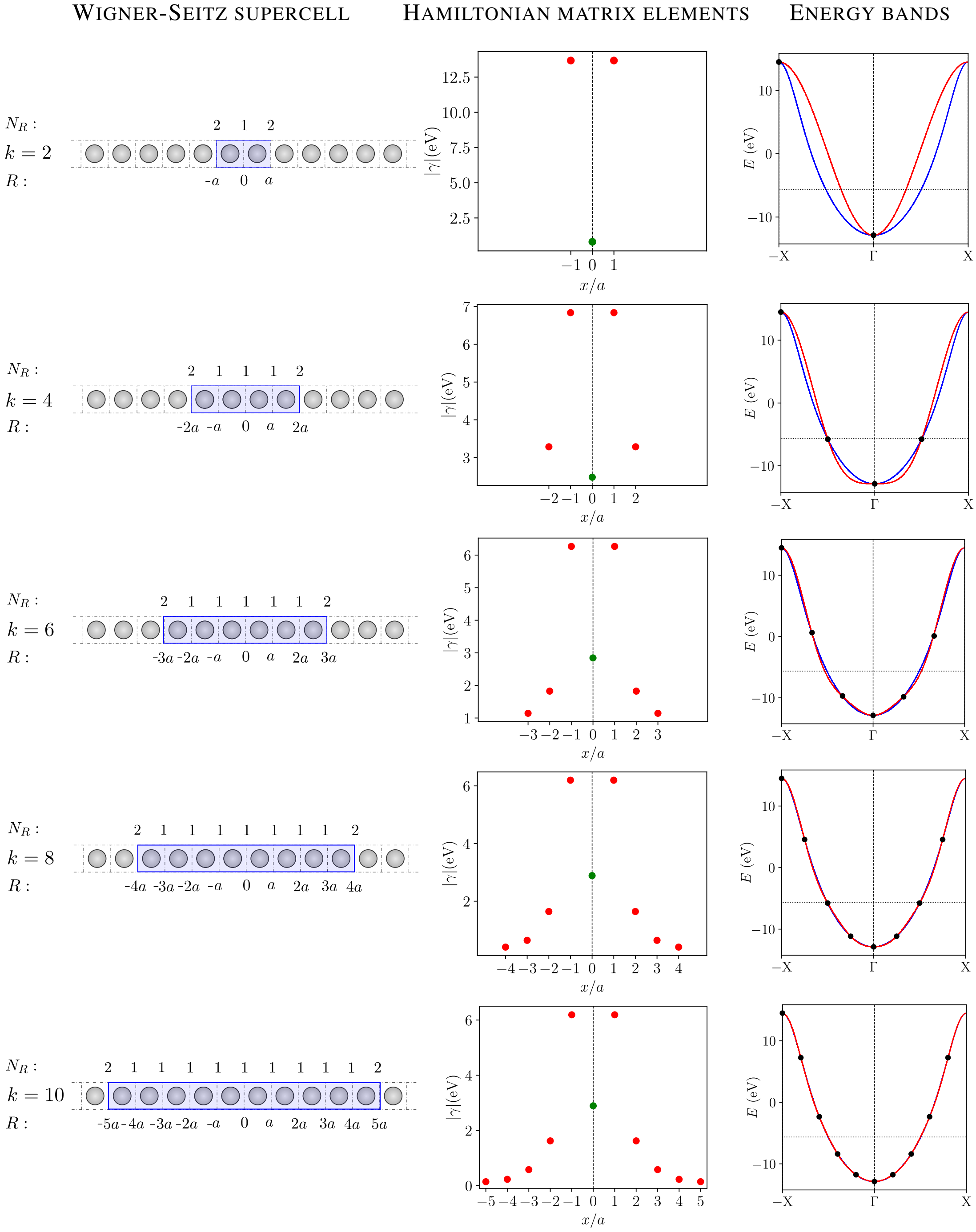


Figure 2.11: Schematic representation of various k_j-point sampling schemes for a linear hydrogen chain. For each sampling, the Wigner-Seitz supercell of the corresponding Born–von Kármán supercell is shaded in blue. Below the chain, the selected R lattice vectors are shown, while their corresponding degeneracies N_R are displayed above. The middle section of the figure presents the magnitude of the Hamiltonian matrix elements between the Wannier function in the simulation cell and its periodic images located at each R lattice vector, as a function of distance. On the right, the interpolated electronic band structures obtained using different k_j-point samplings (red) are compared with first-principles results from SIESTA (blue). Sampled k_j-points are marked in black; at these points, Wannier-interpolated and first-principles energies coincide. The horizontal line denotes the Fermi level.

To illustrate the impact of the k_j-point sampling (see Sec. 2.1.3), we analyze in Fig. 2.11 the behavior of several key quantities as a function of the increasing number of sampled k_j-points. As a first step, we compare the number of Bravais lattice vectors $\vec{R}$, defined according to Sec. 2.1.4. These vectors are related to the real-space supercell employed in the WANNIER90 formalism. We begin with a k_j-point sampling in reciprocal space consisting of two points, $N = 2$. Attending to Sec. 2.1.3 these k_j-points are $k_1 = 0$ and $k_2 = \pi/a$. This k_j-grid in reciprocal space defines an equivalent Born–von Kármán supercell in real space, following the discussion in Sec. 2.1.2. Thus, a k_j-sampling of two corresponds to constructing a supercell in real space composed of two replicas of the original simulation box. Its corresponding Wigner-Seitz supercell is shaded in blue in Fig. 2.11. Subsequently, the Wigner–Seitz cell of the Born-von Kármán supercell is constructed to define the set of Bravais lattice vectors $\vec{R}$. According to the procedure described in Sec. 2.1.4, the selected $\vec{R}$ vectors are $R_0 = 0$, with $N_{R_0} = 1$, and $R_1 = a$, $R_2 = -a$, each with a degeneracy of $N_{R_1} = N_{R_2} = 2$, as both lie on the boundary of the Wigner–Seitz supercell. It can be checked that the sum of the weights, $1/N_{\vec{R}}$, gives the total number of sampled points.

As the number of sampled k_j points increases, the corresponding Born–von Kármán supercell in real space expands, and with it, the associated Wigner–Seitz supercell also enlarges. Consequently, the number of lattice vectors R required to evaluate the Hamiltonian matrix elements H_{mnR} increases, progressively incorporating interactions between more distant Wannier functions. Figure 2.11 displays the evolution of the Wigner–Seitz supercell for the one-dimensional chain as the k-point sampling becomes denser. The selected R vectors according to Sec. 2.1.4 and their associated degeneracies N_R are identified. Notably, vectors lying strictly inside the Wigner-Seitz supercell have $N_R = 1$, while those on the boundaries exhibit $N_R = 2$. It can be verified that the sum of the inverse degeneracies, $\sum 1/N_R$, recovers the total number of sampled k_j points, consistent with the normalization imposed by the Wannier interpolation scheme.

Once the set of lattice vectors R and their corresponding degeneracies have been determined, the next step is to analyze the numerical values of the Hamiltonian matrix elements H_{mmR}. Given that the simulation box contains a single Wannier function, these matrix elements involve only one type of Wannier function and are denoted by $\gamma_R \equiv H_{mmR}$. For each k-point sampling, we represent the Hamiltonian matrix elements γ (measured in eV) as a function of the distance between the Wannier function centered in the home unit cell and a second Wannier function located at a cell associated to the lattice vector R; see Fig. 2.11, "Hamiltonian matrix elements" panel. This figure reveals two key features.

First, within each representation, the matrix elements γ exhibit a rapid decay with increasing distance between the Wannier centers R. This behavior reflects the localized nature of the Wannier functions. From a computational perspective, this property enables a truncation of the sum of the Hamiltonian matrix elements in real space in Eq. (1.65) when evaluating the Hamiltonian in reciprocal space. Specifically, matrix elements γ_R can be neglected beyond a predefined cutoff distance δr_{h}, thereby reducing the computational cost without significantly affecting the accuracy of the results.

Second, by analyzing both the panels in Fig. 2.11 and the data presented in Table 2.1, which reports the values of the matrix elements γ_R for various k-point samplings, clear convergence trends can be identified. The table displays the forward interaction terms only, as the backward interactions are equivalent due to the symmetry of the system. Attending to the table the Hamiltonian matrix elements corresponding to interactions with neighboring sites converge with increasing k-point sampling. This is evident from the consistent γ_R values for equivalent lattice vectors across different k-grids. Additionally, Table 2.1 includes the spatial spread of the Wannier function computed for the simulation box and represented in Fig. 2.10. The results indicate that the Wannier spread also converges as the density of the sampling increases.

Number of k-points for Wannierization	Number of γ	Spread of WF (Å^2)	$\vec{R}/a$	γ (eV)
$(1,1,2)$	3	0.445	$(0,0,0)$	0.806
			$(0,0,1)$	-13.678
$(1,1,4)$	5	0.553	$(0,0,0)$	-2.479
			$(0,0,1)$	- 6.839
			$(0,0,2)$	3.285
$(1,1,6)$	7	0.582	$(0,0,0)$	2.844
			$(0,0,1)$	- 6.266
			$(0,0,2)$	1.825
			$(0,0,3)$	- 1.1449
$(1,1,10)$	11	0.600	$(0,0,0)$	-2.891
			$(0,0,1)$	- 6.189
			$(0,0,2)$	1.621
			$(0,0,3)$	- 0.579
			$(0,0,4)$	0.228
			$(0,0,5)$	- 0.141

Table 2.1: The Hamiltonian matrix elements between the Wannier function centered in the home unit cell and the Wannier functions located in the cell defined by $\vec{R}$ in the forward direction are shown as a function of the number of the k-point sampling. The number of interactions reflects the expansion of the Born–von Kármán supercell as the density of the k-grid increases. Additionally, the spreading of the Wannier function centered in the home unit cell is represented, providing insight into its degree of localization.

Finally, we study the bands obtained from the real space Hamiltonian matrix elements γ_R. The system under consideration exhibits a single electronic band, consistent with the presence of only one Wannier function per unit cell. For each k-point grid, we present the interpolated band structure computed following Eq. (2.17) using the Hamiltonian matrix elements in real space shown in Fig. 2.11. The interpolated bands, obtained via the WANNIER90 formalism, are shown in red. For comparison, the reference band structure derived from first-principles calculations using the SIESTA code is displayed in blue.

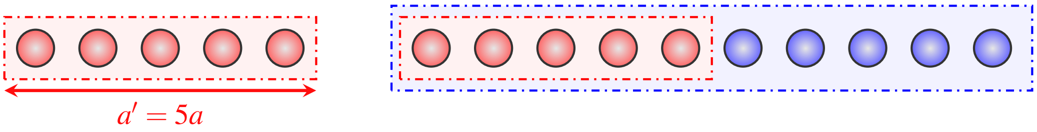
In the evaluation of the band energies E_k, the k_j-points used in the construction of the Wannier functions determine the specific points at which the interpolated energies exactly match those computed from first-principles, as discussed in Sec. 2.1.6. The coordinates associated to the sampled points, (k_j, E_{k_j}), are indicated with black dots in the figure.

As the number of the k_j-points, N, increases, the interpolated band structure shows clear convergence toward the reference band obtained from SIESTA. In particular, for $N = 8$, the agreement between the bands from WANNIER90 and those from the first-principles calculation becomes essentially exact across the entire Brillouin zone.

2.2.2 Linear chain of H-atoms. Larger simulation boxes.

In the previous Section, we analyzed the rapid decay of the real-space Hamiltonian matrix elements as a function of the distance between Wannier functions, considering the case of a single Wannier function per simulation box.

In this Section, we extend the analysis to systems containing multiple Wannier functions within a single simulation box. Specifically, we examine a one-dimensional linear chain of hydrogen atoms where the simulation box now comprises five atoms, leading to a lattice constant $a' = 5a$, shaded in red in Fig. 2.12. As a result, five s-like Wannier functions (schematically represented as red spheres in Fig. 2.12) are included, each centered on a different hydrogen atom.



(a) SIESTA simulation box.

(b) Born von Karman cell.

Figure 2.12: (a) Simulation box employed in SIESTA, consisting of five s-like Wannier functions, each centered on a hydrogen atom. (b) Born–von Kármán supercell (shaded in blue) constructed from the simulation box shown in (a), corresponding to a k-point sampling with $N = 2$. The resulting supercell comprises two repetitions of the original simulation box, leading to a ten Wannier functions in the Born-von Kármán supercell.

For studying this problem we adopt a k_j-point sampling with $N = 2$ points. Consequently, the Born–von Kármán supercell constructed within the WANNIER90 framework consists of two simulation boxes, encompassing a total of ten Wannier functions. This supercell is represented in blue in Fig. 2.12(b). In this figure, the combined set of red and blue spheres corresponds to the Wannier functions in the Born-von Kármán supercell. The corresponding Wigner–Seitz supercell of the Born-von Kármán supercell is depicted in blue in Fig. 2.13. Additionally, Fig. 2.13 includes the set of selected lattice vectors $R \in \{-a', 0, a'\}$ which have been determined following the methodology outlined in Sec. 2.1.4.

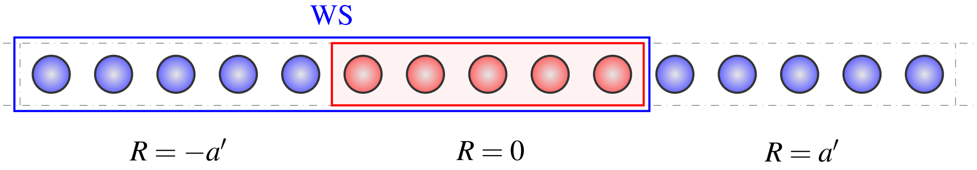


Figure 2.13: Schematic representation of a one-dimensional linear chain of hydrogen atoms, where each sphere corresponds to an s-like Wannier function centered on an individual hydrogen atom. The simulation box, shaded in red, contains five Wannier functions. The Wigner–Seitz cell corresponding to the Born–von Kármán supercell, constructed from a k-point sampling of two points, is outlined in blue. The selected lattice vectors R, determined according to the procedure described in Sec. 2.1.4 and used in the evaluation of the Hamiltonian matrix elements H_{mnR}, are indicated below their associated unit cells.

At this stage, we begin the analysis of the Hamiltonian matrix elements H_{mnR} which are expressed in real space. These matrix elements describe the interaction between the mth Wannier function χ_{m0}, centered at the home unit cell $R = 0$, and the nth Wannier function χ_{nR}, located in the unit cell defined by the lattice vector R. In this system, the indices m, n run over the five Wannier functions in the simulation box, $m, n \in \{0, 1, 2, 3, 4\}$.

Figure 2.14 presents the obtained values of the Hamiltonian matrix elements H_{0nR} measuring the interaction between the Wannier function $\chi_{0,0}$ and a second Wannier function χ_{nR}. The horizontal axis indicates the position of the second Wannier in units of the lattice constant a.

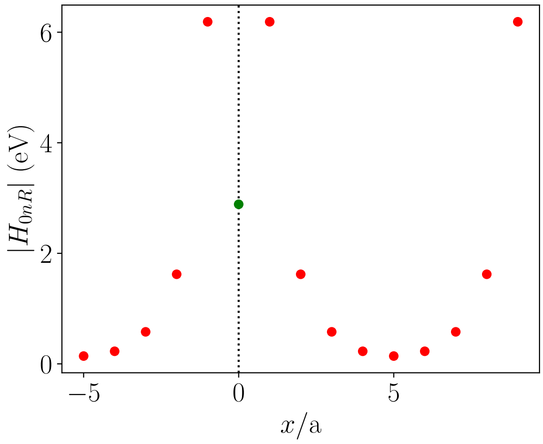


Figure 2.14: Representation of the absolute values of the real-space Hamiltonian matrix elements between the Wannier function $\chi_{1,0}$, located in $R = 0$, and the remaining Wannier functions χ_{nR} placed in the cells $R \in \{-a', 0, a'\}$. The horizontal axis indicates the spatial distance between $\chi_{1,0}$, at the origin, and the corresponding Wannier function χ_{nR}.

The plot illustrates a rapid decay in the magnitude of the Hamiltonian matrix elements up to the fifth nearest-neighbor shell. Although a monotonic decrease with increasing distance between Wannier functions might be expected, deviations from this trend appear beyond the sixth shell. Specifically, the interaction between $\chi_{0,0}$ and $\chi_{1,a'}$ is found to be equivalent to that between $\chi_{0,0}$ and $\chi_{1,-a'}$, both shown in green in Fig. 2.15. Similarly, the Hamiltonian matrix element between $\chi_{0,0}$ and $\chi_{2,a'}$ matches that of $\chi_{0,0}$ with $\chi_{2,-a'}$, represented in purple in Fig. 2.15. Analogous correspondences continue for subsequent pairs.

This behavior arises from the inclusion of interactions with the periodic images of the Born–von Kármán supercell. Wannier functions located beyond the sixth nearest-neighbor cell correspond to those at positions $x < 0$ translated by a vector of the Born-von Kármán supercell, $T = -2a'$. As discussed in Sec. 2.1.5, the Hamiltonian matrix elements $H_{mn\vec{R}}$ are invariant under a translation $\vec{T}$ of the Born-von Kármán supercell, i.e., $H_{mn\vec{R}} = H_{mn\vec{R}'}$, where $\vec{R}' = \vec{R} + \vec{T}$ is a periodic image of $\vec{R}$. Accordingly, the Wannier function $\chi_{1,a'}$ is a periodic replica of $\chi_{1,-a'}$, which lies closer to the reference Wannier function $\chi_{0,0}$, thereby explaining the observed equivalence in their corresponding Hamiltonian matrix elements on Fig. 2.14.

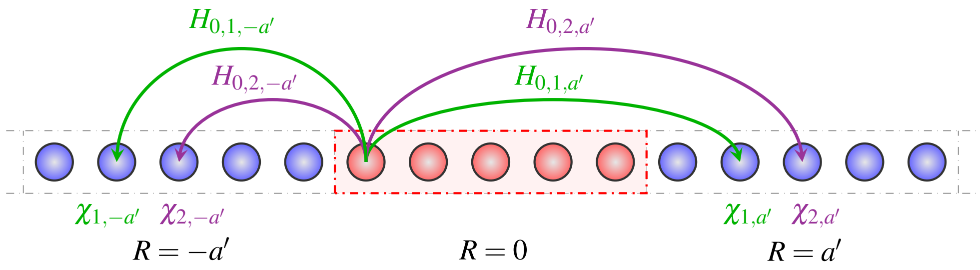


Figure 2.15: Representation of interactions between Wannier functions related by a translational vector of the Born–von Kármán supercell. The interaction between the Wannier functions $\chi_{0,0}$ and $\chi_{1,a'}$, indicated by a green arrow pointing to the right, corresponds to the Hamiltonian matrix element $H_{0,1,a'}$. This interaction is equivalent to that between $\chi_{0,0}$ and $\chi_{1,-a'}$, indicated by a green arrow pointing to the left, as the functions $\chi_{1,a'}$ and $\chi_{1,-a'}$ are related via a translational vector $T = -2a'$ applied to $\chi_{1,a'}$. Similarly, the Wannier functions $\chi_{2,a'}$ and $\chi_{2,-a'}$ are also related by the translation vector $T = -2a'$, making the interactions $H_{0,2,a'}$ and $H_{0,2,-a'}$, represented in purple, equivalent.

Finally, we verify that the simulation presented in this section—performed using an enlarged simulation cell with a lattice constant $a' = 5a$, comprising five Wannier functions and a k-point sampling of two—is equivalent to a simulation using the original, smaller unit cell with lattice constant a, containing a single Wannier function and a k-point sampling of ten. This equivalence is evidenced by comparing Fig. 2.14 with the last row of Fig. 2.11.

2.3 Comparison of the tight-binding formalism in WANNIER90 and the standard text-book formalism:

In Sec. 2.1.6 we have summarized how WANNIER90 interpolates the band structure of a given material starting from a few Hamiltonian matrix elements defined in real space, $H_{mn\vec{R}}$. A priori, that formalism is exactly the same as the one described in Sec. 1.3. However, some care must be taken if we want to use directly the matrix elements in real space stored in the `tb.dat` file. Here, we shall explain the subtleties focusing on the previously analyzed system- a linear chain of hydrogen atoms. We consider a simulation cell with a lattice parameter $a' = 2a$, containing two hydrogen atoms and, consequently, two s-like Wannier functions. The simulation box is shaded in red in Fig. 2.16. These Wannier functions possess the same characteristics as those discussed in the preceding sections.

The energy bands are first computed by diagonalizing the Hamiltonian in reciprocal space, where the corresponding matrix elements have been constructed using the interpolation procedure outlined in Sec. 2.1.6.

2.3.1 WANNIER90 tight binding

For the selected system, we consider a k-point sampling consisting of two points, which, according to the procedure outlined in Sec. 2.1.4, corresponds to the selection of three Bravais lattice vectors $R \in \{-a', 0, a'\}$.

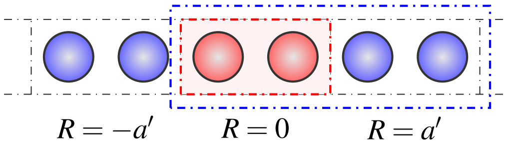


Figure 2.16: In SIESTA, the simulation box, shaded in red, contains two s-line Wannier functions. The corresponding Born-von-Kármán supercell in WANNIER90, for a k-sampling of $N = 2$, is framed in blue. This represents a supercell of two simulation boxes. Periodic replicas of the simulation cell are labeled by the R-vector.

To perform the band interpolation as described by Eq. (2.17), the Hamiltonian matrix elements are systematically constructed by summing over the selected R-lattice vectors.

We begin by analyzing the diagonal terms $H_{1,1,k}$ which relates the interaction between the Wannier function $\chi_{1,0}$ and a second Wannier function $\chi_{1,R}$. Consequently, the interactions involving Wannier orbitals of label “1” include the on-site interaction, the interaction with the second-nearest neighbor at $R = a'$, and the interaction with the second-nearest neighbor at $R = -a'$. These last Wannier functions are related through a translational vector $T = 2a'$, which defines the periodicity of the Born–von Kármán supercell.

Once considered the on-site interaction, renamed as α^{W90} we study the interaction associated to $R = a'$.

For $R = a'$, which has a degeneracy $N_R = 2$, we consider the interaction between $\chi_{1,0}$ and the Wannier function $\chi_{1,a'}$, located in the unit cell defined by this lattice vector. According to the procedure described in Sec. 2.1.6, the interpolation algorithm determines whether the Wannier function $\chi_{1,0}$ interacts more strongly with $\chi_{1,a'}$ or one of its periodic images, specifically $\chi_{1,-a'}$, obtained via the translation $T = 2a'$. In this scenario, $\chi_{1,a'}$ and $\chi_{1,-a'}$ are equidistant from $\chi_{1,0}$, and therefore both contribute equally to the interaction.

Consequently, for $R = a'$, the code averages the contributions of both interactions, assigning a weight of $1/\mathscr{N}_{1,a'}$, where $\mathscr{N}_{1,a'} = 2$. An analogous argument applies for the lattice vector $R = -a'$, as the spatial arrangement and symmetry of the system yield equivalent contributions. These interactions are hereafter denoted as γ_2^{W90} and represented in Fig. 2.17.

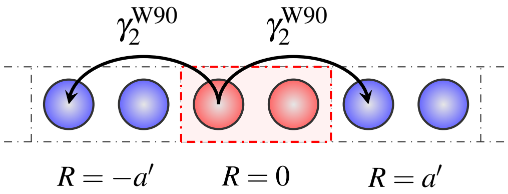


Figure 2.17: The interactions between $\chi_{1,0}$ and the periodic replicas $\chi_{1,a'}$ and $\chi_{1,-a'}$ are denoted as γ_2. Both orbitals are at the same distance of $\chi_{1,0}$..

$$\begin{aligned} H_{1,1,k} = & \frac{1}{4}e^{-ika'}H_{1,1,-a'} + \frac{1}{4}e^{ika'}H_{1,1,-a'} + \\ & + H_{1,1,0} + \\ & + \frac{1}{4}e^{ika'}H_{1,1,a'} + \frac{1}{4}e^{-ika'}H_{1,1,a'} = \\ = & \alpha^{\mathrm{W90}} + \gamma_2^{W90}\cos\left(ka'\right). \end{aligned} \tag{2.22}$$

Similarly, considering the interactions between $\chi_{2,0}$ and $\chi_{2,R}$,

$$\begin{aligned} H_{2,2,k} = & \frac{1}{4}e^{-ika'}H_{2,2,-a'} + \frac{1}{4}e^{ika'}H_{2,2,-a'} + \\ & + H_{2,2,0} + \\ & + \frac{1}{4}e^{ika'}H_{2,2,a'} + \frac{1}{4}e^{-ika'}H_{2,2,a'} = \\ = & \alpha^{\mathrm{W90}} + \gamma_2^{W90}\cos\left(ka'\right). \end{aligned} \tag{2.23}$$

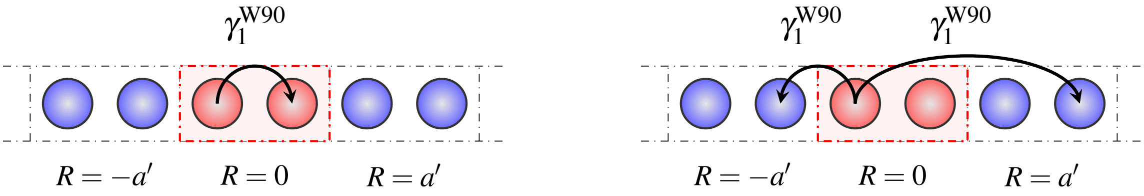
We now analyze the off-diagonal Hamiltonian matrix elements, specifically $H_{12\vec{k}}$, which describe the interaction between the functions $\chi_{1,0}$, centered in the home unit cell, and the type-2 Wannier functions.

For the lattice vector $R = 0$, where the degeneracy factor is $N_{R=0} = 1$, no periodic image of $\chi_{2,0}$ lies closer to $\chi_{1,0}$ than the function itself. Therefore, the degeneracy factor is $\mathscr{N}_{1,2,0} = 1$, and the interaction is denoted by γ_1^{W90}, represented in Fig. 2.18(a).

In the case of the lattice vector $R = a'$, which has a degeneracy $N_R = 2$, the relevant orbital is $\chi_{2,a'}$. However, its periodic image located at $R' = R + T_{1,2,0} = -a'$, with $T_{1,2,0} = -2a'$, is closer to the reference orbital $\chi_{1,0}$. Hence, the translation vector $T_{1,2,0}$ is used to account for the interaction in the Hamiltonian matrix construction. Analogously, for $R = -a'$, there exists no periodic image of $\chi_{2,-a'}$ that is closer to $\chi_{1,0}$ than the orbital itself, and the original interaction is retained. These interaction are reprsented in Fig. 2.18(b).

$$H_{1,2,k} = H_{1,2,0} + \frac{1}{2}e^{-ika'}H_{1,2,a'} + \frac{1}{2}e^{-ika'}H_{1,2,-a'} = \gamma_1^{\mathrm{W90}}(1 + e^{-ika'}) \tag{2.24}$$

$$H_{2,1,k} = H_{1,2,0} + \frac{1}{2}e^{ika'}H_{1,2,a'} + \frac{1}{2}e^{ika'}H_{1,2,-a'} = \gamma_1^{\mathrm{W90}}(1 + e^{ika'}) \tag{2.25}$$



(a) Interactions between $\chi_{1,0}$ and $\chi_{2,0}$. (b) Interactions between $\chi_{1,0}$ and the periodic replicas of $\chi_{2,a'}$.

Figure 2.18: Interactions between the Wannier function $\chi_{1,0}$ and the functions $\chi_{2,R}$. Panel (a) illustrates the interaction between $\chi_{1,0}$ and $\chi_{2,0}$. Panel (b) shows the interaction between $\chi_{1,0}$ and $\chi_{2,a'}$, as well as the interaction with its periodic replica $\chi_{2,-a'}$, associated with a translation vector of the Born–von Kármán supercell, $T = -2a'$, from $\chi_{2,a'}$. The Wannier function $\chi_{2,-a'}$ is closer to $\chi_{1,0}$ than $\chi_{1,a'}$.

Finally, the Hamiltonian matrix as a function of the k vector is obtained,

$$H_k = \begin{pmatrix} \alpha^{\mathrm{W90}} + \gamma_2^{W90}\cos(ka') & \gamma_1^{\mathrm{W90}}(1 + e^{-ika'}) \\ \gamma_1^{\mathrm{W90}}(1 + e^{ika'}) & \alpha^{\mathrm{W90}} + \gamma_2^{W90}\cos(ka') \end{pmatrix} \tag{2.26}$$

The eigenvalues of a two-dimensional hermitian matrix,

$$\begin{pmatrix} A & B \\ B^* & A \end{pmatrix} \tag{2.27}$$

are $\lambda = A \pm \sqrt{B \cdot B^*}$. Consequently, the eigenvalues of Eq. (2.26) are the energy bands

$$E_1(k) = \alpha^{\mathrm{W90}} + 2\gamma_1^{\mathrm{W90}}\cos\left(\frac{ka'}{2}\right) + \gamma_2^{W90}\cos(ka'), \tag{2.28}$$

$$E_2(k) = \alpha^{\mathrm{W90}} - 2\gamma_1^{\mathrm{W90}}\cos\left(\frac{ka'}{2}\right) + \gamma_2^{W90}\cos(ka'). \tag{2.29}$$

2.3.2 Standard tight binding formalism

We now turn our attention to the standard tight-binding formalism. Following this approach, the Hamiltonian matrix elements for an arbitrary $\vec{k}$-point are constructed according to Eq. (1.16).

As a preliminary step, we must identify the number of neighboring shells to be considered, consistent with the methodology employed in the previous analyses with WANNIER90 code. To this end, we center the Wigner–Seitz cell on each Wannier function within the simulation cell. From this inspection, it is evident that we should include the interactions up to the second-nearest neighbors. This is illustrated in Fig. 2.19.

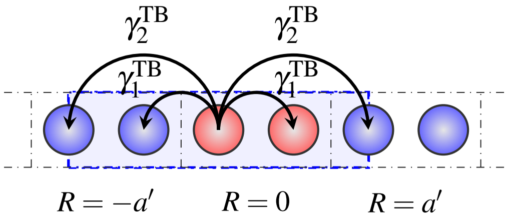


Figure 2.19: Linear chain characterized by two orbitals per simulation cell (red orbitals) and a k-sampling (1,1,2). The Wigner-Seitz supercell, associated with the Born von Karman cell, is shaded in blue and centered on the first orbital to identify the Wannier functions interacting with it. The possible interactions between $\chi_{1,0}$ and the orbitals of the system are included (except the on-side interaction).

Considering the Hamiltonian matrix elements (also referred to as *hoppings*) presented in Fig. 2.19, where α^{TB} represents the on-site matrix element of a Wannier function with itself, and γ_1^{TB} and γ_2^{TB} denote the hoppings to the first- and second-nearest neighbors, respectively, we compute the corresponding Hamiltonian matrix elements in reciprocal space using Eq. (1.16),

$$H_{1,1,k} = \alpha^{\mathrm{TB}} + e^{ika'}\gamma_2^{\mathrm{TB}} + e^{-ika'}\gamma_2^{\mathrm{TB}} = \alpha^{\mathrm{TB}} + 2\gamma_2^{\mathrm{TB}}\cos\left(ka'\right), \tag{2.30}$$

$$H_{2,2,k} = \alpha^{\mathrm{TB}} + e^{ika'}\gamma_2^{\mathrm{TB}} + e^{-ika'}\gamma_2^{\mathrm{TB}} = \alpha^{\mathrm{TB}} + 2\gamma_2^{\mathrm{TB}}\cos\left(ka'\right), \tag{2.31}$$

$$H_{1,2,k} = \gamma_1^{\mathrm{TB}}(1+e^{-ika'}), \tag{2.32}$$

$$H_{2,1,k} = \gamma_1^{\mathrm{TB}}(1+e^{ika'}). \tag{2.33}$$

Since all the Wannier functions are identical and equidistance, it is clear that the on-site Hamiltonian matrix elements, the first-neighbor interactions and second-neighbor interactions will be the same regardless of the orbital under study.

The resulting Hamiltonian matrix follows the shape

$$H_k = \begin{pmatrix} \alpha^{\mathrm{TB}} + 2\gamma_2^{\mathrm{TB}} \cos(ka') & \gamma_1^{\mathrm{TB}}(1+e^{-ika'}) \\ \gamma_1^{\mathrm{TB}}(1+e^{ika'}) & \alpha^{\mathrm{TB}} + 2\gamma_2^{\mathrm{TB}} \cos(ka') \end{pmatrix}, \tag{2.34}$$

where the eigenvalues are the following energies,

$$\begin{aligned} E_1(k) &= \alpha^{\mathrm{TB}} + 2\gamma_2^{\mathrm{TB}} \cos(ka') + 2\gamma_1^{\mathrm{TB}} \cos\left(\frac{ka'}{2}\right) \\ E_2(k) &= \alpha^{\mathrm{TB}} + 2\gamma_2^{\mathrm{TB}} \cos(ka') - 2\gamma_1^{\mathrm{TB}} \cos\left(\frac{ka'}{2}\right). \end{aligned} \tag{2.35}$$

Comparing the eigenvalues obtained from both the WANNIER90 and tight-binding methodologies, we observe that WANNIER90 operates akin to a tight-binding method, satisfying the relations, $\alpha^{\mathrm{TB}} = \alpha^{\mathrm{W90}}$, $\gamma_1^{\mathrm{TB}} = \gamma_1^{\mathrm{W90}}$, $\gamma_2^{\mathrm{TB}} = \gamma_2^{\mathrm{W90}}/2$ in this scenario. If this calculation were repeated for a case using a k-sampling $N = 4$ for WANNIER90, the equivalent tight-binding approximation would consider interactions up to fourth neighbors. Again, the relationship between the interactions obtained from WANNIER90 and those from the tight-binding approximation would be one, except for the interactions at the boundaries, where a factor of 2 would appear again. In these instances, it becomes apparent that the interactions obtained with WANNIER90 at the boundaries are twice those of the tight-binding interactions.

It is noteworthy that in these examples, for a clear theoretical treatment of the WANNIER90 approach, we have worked with a small simulation cell and a minimal number of k-points. Nevertheless, in practical calculations, the interactions at the boundaries tend to diminish to an extent where this effect becomes negligible. This is either due to a large number of k-points or a large simulation cell, or a combination of both factors.

Hence, it follows that when the number of $k-$points is sufficiently large, WANNIER90 becomes equivalent to a tight-binding treatment with a specific number of neighboring shells.

Chapter 3

State of art: Electron-vibration interaction

The interaction between electrons and atomic vibrations is a central topic in condensed matter and molecular physics, essential to understand phenomena such as electrical resistance, charge transport, and superconductivity in solids, as well as electron transfer and non-radiative relaxation processes in molecules [73; 74; 75]. Given its fundamental role across a wide range of physical and chemical processes, as will be shown in detail below, it is very important taking them into account during the construction of second-principles models. And so, this Chapter focuses on the study of the basis of electron–vibration interactions. In the bibliography and along this Chapter, this interaction is commonly referred to as *vibronic coupling* (vibrational–electronic coupling) or known as *electron-phonon* or *electron-lattice* interaction in solid state physics. We will use both terms as equivalent.

In order to include the interaction between electronic and nuclear degrees of freedom, the concept of vibronic coupling arises as a framework to address two fundamental questions: How do nuclear vibrations influence the electronic structure? And conversely, how do electronic states affect nuclear dynamics? To describe geometric distortions we can use three complementary representations: individual atomic displacements $\{\vec{u}_\lambda\}$, atomic positions $\{\vec{\tau}_\lambda\}$, and vibrational normal modes $\{Q_\Gamma\}$, which correspond to collective atomic motions. Each vibrational mode can be expressed as a linear combination of atomic displacements. In the case of a specific mode Q_Γ, where the subscript Γ does not denote a mode at the high-symmetry Γ point but rather refers to any mode of the system, and considering a Cartesian coordinate α,

$$Q_{\Gamma,\alpha} = \sum_{\lambda} c_{\lambda\Gamma,\alpha} u_{\lambda,\alpha} \tag{3.1}$$

with the inverse relation given by,

$$u_{\lambda,\alpha} = \sum_{\Gamma} \widetilde{c}_{\lambda\Gamma,\alpha} Q_{\Gamma,\alpha}. \tag{3.2}$$

In this Chapter, we first illustrate the importance of this interaction across various physical and chemical phenomena and highlight its relevance in numerous applications. After establishing its physical significance, we present various theoretical formalisms that have been developed to describe the electron-lattice interaction.

Furthermore, we explore how vibronic effects can be incorporated from a computational perspective. This Chapter attempts to synthesize the state-of-the-art theoretical developments and practical strategies for treating electron–phonon interactions within first-principles calculations, particularly within DFT. In this context, we address the computational challenges and the corresponding solutions they entail.

In order to study the state of art of the vibronic coupling, a comparative analysis of different approaches from DFT to incorporate the electron–lattice interaction is provided. Among the significant recent contributions discussed in this Chapter are the finite-difference approach to electron–phonon coupling [76], a reciprocal-space formulation to deal with solid state systems [74] and its treatment based on Wannier function representations [68; 69; 74]. Finally, we also present the incorporation of vibronic coupling through an alternative approach: second-principles method [37], which constitutes the key methodology in this work.

3.1 Historical development

Bloch introduced the concept of *electron-phonon interaction* while studying the temperature dependence of electrical resistivity in metals [77]. Thermal vibrations of the crystal lattice—later identified as *phonons*—play a fundamental role in understanding electrical resistivity. Although Bloch did not explicitly refer to phonons, he identified elastic waves as the primary mechanism for electron scattering at low temperatures [77]. In his work, Bloch studied the interaction between two coupled systems—the electron gas, simulating the electrons in a metal, and the lattice, capable of elastic oscillations—using a perturbative approach.

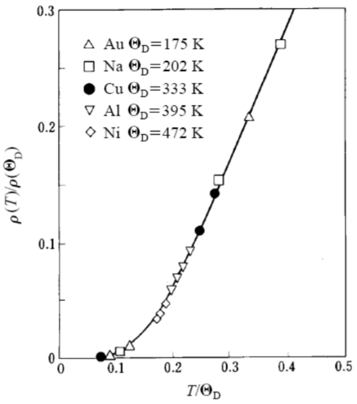


Figure 3.1: Temperature dependence of the electrical resistivity for various metals. This figure shows the T^5 dependence for low temperatures regime. This figure has been extracted from Ref. [78] with permissions.

At sufficiently low temperatures, electron scattering is dominated by acoustic phonons, characterized by large wavelength and low energy. In this regime, it can be assumed that the wavelengths of the elastic waves are much larger than the interatomic distances, allowing the crystal to be treated as a continuum. Furthermore, the amplitudes of these waves are small enough to apply the harmonic approximation. Under these assumptions, Bloch derived the temperature dependence of thermal resistivity demonstrating that it follows a T^5 dependence, leading to the so-called "Bloch T^5 law". In 1940, Bloch predictions were experimentally confirmed by Bardeen [78], Fig. 3.1.

Although Bloch employed a classical description of lattice vibrations by means of long wavelength elastic waves, the derivation of the "Bloch T^5 law" is only possible considering some main fundamental quantum mechanical principles, making his approach effectively semi-classical. In particular, Bloch's theory is based on two key quantum principles: (i) the quantization of the vibrational energies, where phonon occupancy is governed by the Bose-Einstein statistics; (ii) electron scattering by lattice vibrations must obey selection rules leading to quantum energy and momentum conservation laws.

The general quantization of lattice vibrations for studying lattice dynamics was first introduced by Frenkel [79]. While Bloch described lattice vibrations in terms of elastic waves, Frenkel replaced this concept with that of *phonons* [79]. His work formally addressed the mechanics of phonons, which represent the quantized vibrational modes in crystalline solids. This treatment established a foundational framework for understanding how lattice vibrations influence the thermal, acoustic, and electronic properties of materials.

3.2 Applications

The electron–phonon interaction plays a crucial role in the understanding of different solid-state properties. In particular, the vibronic coupling interaction is fundamental in the study of (i) charge and energy transport of several materials including the development of efficient solar cells and other electronic devices [80; 81; 82; 83; 84], semiconductor physics [85], superconductivity [86; 87; 88] and the formation of polarons [88; 89; 90; 91] and localized defects in insulators [92]; (ii) vibronic effects in spectroscopy like the Stokes shift [93; 94; 95] and (iii) structural changes as the symmetry breaking and Jahn-Teller effect [73; 96; 97].

3.2.1 Symmetry breaking and Jahn-Teller effect

From an energetic perspective, symmetry breaking is inherently connected to the emergence of multiple minima on the potential energy surface of a system. This phenomenon holds profound implications across a wide range of physical processes. One of the most interesting examples is the *maser* [98], where the symmetry breaking plays a central role in its functionality, as we explain below. In this case, the underlying process is crucial factor driving the device's operation, demonstrating the tangible impact of symmetry breaking in real-world applications.

A *maser* (Microwave Amplification by Stimulated Emission of Radiation) is a device designed to generate and amplify electromagnetic waves through stimulated emission. It represents one of the pioneering developments that preceded the invention of the laser. While both devices rely on similar amplification principles, the maser operates specifically in the microwave region of the electromagnetic spectrum. The first maser was constructed in 1953 by Charles H. Townes [98] in collaboration with J.P. Gordon and H.J. Zeiger using ammonia (NH_3) as the active medium. This groundbreaking work was awarded by the Nobel Prize in Physics in 1964.

Ammonia is a non-planar molecule that exhibits two equivalent equilibrium configurations—both $C_{3\text{v}}$ pyramidal structures—in which the nitrogen atom resides either above or below the plane defined by the three hydrogen atoms. As a result, the potential energy surface features two degenerate minima corresponding to these geometries, leading to a symmetric double-well potential represented in Fig. 3.2. These wells are separated by an energy barrier of 25 kJ/mol [99] (ten times greater than thermal energy at room temperature) associated with a transition state of higher symmetry, specifically a planar $D_{3\text{h}}$ configuration.

The functionality of the ammonia maser is fundamentally based on the *umbrella inversion* of the ammonia molecule [100], a quantum mechanical phenomenon closely linked to symmetry breaking. The umbrella inversion refers to the process by which the nitrogen atom transitions between the two equivalent equilibrium positions via the tunneling effect inverting the pyramid. This inversion flip-flops repeatedly with a tunneling rate [99]. To understand the underlying physics, it is important to recognize that, due to the equivalence of the two pyramidal geometries, the associated potential wells are also equivalent, as are their vibrational energy levels. Thus, one can initially model the system as two independent harmonic oscillators: oscillator a corresponding to the configuration where the nitrogen atom is "above" the hydrogen plane, and oscillator b corresponding to the configuration where it is "below" the plane. Each oscillator supports quantized vibrational states, and we focus on the fundamental vibrational states, called here φ_a and φ_b, of each configuration [101].

Due to the tunneling effect, these two oscillators begin to interact. Analogous to the molecular orbital approach used in the treatment of the hydrogen molecule, we construct linear combinations of the vibrational wavefunctions φ_a and φ_b. These combinations yield symmetric, Ψ^+ and antisymmetric Ψ^- molecular vibrational wavefunctions [101],

$$\Psi^+ = \frac{1}{\sqrt{2}}(\varphi_a + \varphi_b), \quad \Psi^- = \frac{1}{\sqrt{2}}(\varphi_a - \varphi_b). \tag{3.3}$$

As a result, the interaction of the two wells via tunneling lifts the degeneracy of the vibrational levels, opening an energy splitting between the symmetric, Ψ^+ (lower in energy) and antisymmetric, Ψ^- (higher in energy) states. These energies levels are depicted in Fig. 3.2 by horizontal lines. In this way, the emission of the maser is a result of the desexcitation of the NH_3 molecules from the antisymmetric state to the symmetric state. The energetic separation between these states is 0.8118 cm^{-1} [99], transition associated to the microwave region of the electromagnetic spectrum.

In the case of the NH_3 maser, although the *umbrella inversion* arises from a quantum tunneling process, the origin of the characteristic double-well potential lies in the instability of the high-symmetry planar geometry (D_h). How, then, can we understand the symmetry breaking observed in the ammonia molecule? This instability is a direct consequence of a Pseudo Jahn–Teller (PJT) effect [102].

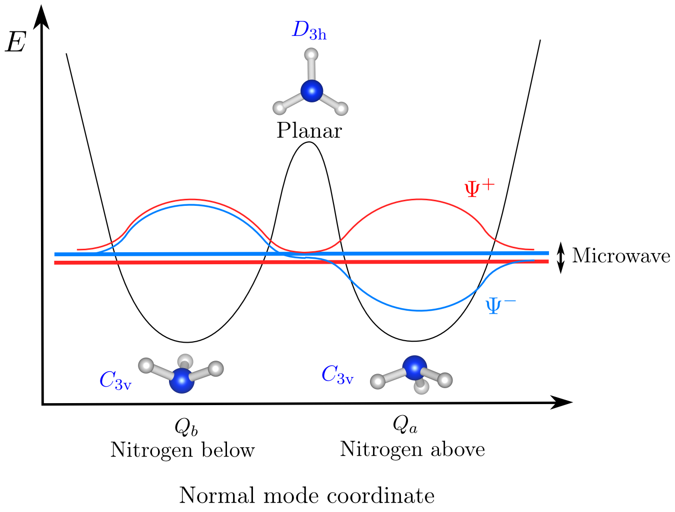


Figure 3.2: Potential energy surface of the ammonia (NH_3) molecule along the normal mode coordinate associated with the oscillatory mode of the nitrogen atom along the axis perpendicular to the plane of the hydrogen atoms. This vibrational mode gives rise to a double-well potential with two minima located at normal mode coordinates Q_a and Q_b, corresponding to the nitrogen atom positioned "above" and "below" the hydrogen plane, respectively. Each minimum exhibits $C_{3\mathrm{v}}$ symmetry. The two wells are separated by an energy barrier associated with a transition state, which corresponds to the planar geometry of the molecule characterized by $D_{3\mathrm{h}}$ symmetry. Quantum tunneling between the two equivalent minima leads to a splitting of the degenerate vibrational energy level into two distinct levels, represented by horizontal lines, corresponding to a symmetric state Ψ^+ (shown in red) and an antisymmetric state Ψ^- (shown in blue). This energy splitting occurs in the microwave region of the electromagnetic spectrum. This figure has been adapted from Ref. [101].

This phenomenon is of considerable importance, as it reflects a broader theoretical framework for understanding symmetry breaking in condensed matter systems. As discussed in Ref. [73], the instability of a high-symmetry configuration in a system can generally be attributed to one of three mechanisms: the Jahn–Teller effect (in the case of degenerate electronic states), the Pseudo Jahn–Teller effect (for non-degenerate states), or the Renner–Teller effect (specific to linear molecular systems).

In order to introduce the Jahn-Teller theorem, two fundamental concepts must first be introduced: the spatial degeneracy of electronic states and the susceptibility of a system leading to symmetry breaking. Firstly, regarding the spatial degeneracy of electronic states: two electronic states are said to be *degenerate*, for fixed nuclear positions, if they have the same electronic energy upon solving the electronic Schrödinger equation. Secondly, in this context, a system susceptible of suffering a symmetry breaking in a degenerate state refers to a configuration associated with a high-symmetry geometry, where the degeneracy arises from the symmetry of the system. Upon the application of a small distortion—lowering the molecular symmetry—the system undergoes a reduction of the total electronic energy, indicating that the high-symmetry degenerate state is energetically unfavorable.

As illustrated in Fig. 3.3, we consider an octahedral complex of the form MX_6, characterized by an open-shell electronic configuration. In this case, the high-symmetry O_h configuration corresponds to an unstable point. Due to the presence of electronic degeneracy in the open-shell state, the system is susceptible to a symmetry-lowering distortion—typically toward a tetragonal geometry (with symmetry $D_{4\mathrm{h}}$)—that lifts the degeneracy of the electronic levels. This distortion leads to an energy splitting and results in a more stable configuration of the complex. Later on, we will examine a concrete example.

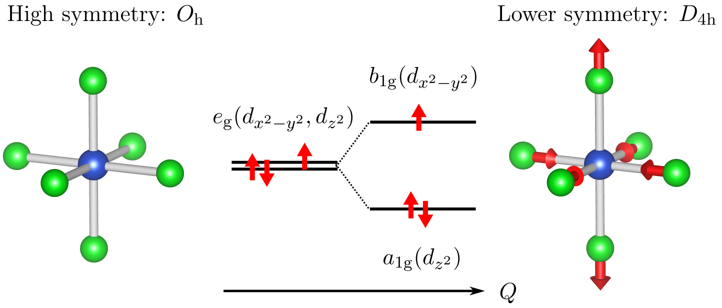


Figure 3.3: Highest occupied molecular orbital (HOMO) of the $CuCl_6^{4-}$ complex in its high-symmetry O_h geometry. This electronic state is unstable, resulting in a lowering of the total energy through a distortion of the geometry—specifically, an elongation along the z-axis—which reduces the molecular symmetry and lifts the orbital degeneracy.

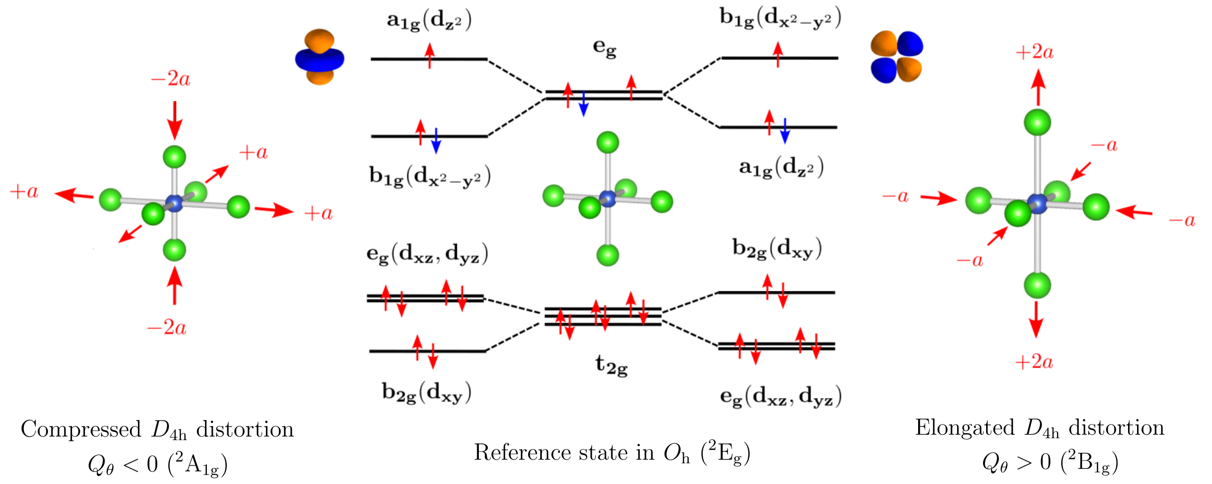
This example illustrates the fundamental interplay between degeneracy and structural distortion. It reveals how a system with a spatially degenerate electronic ground state is unstable and can lower its energy through a symmetry-breaking distortion that lifts the degeneracy. This leads us to the general statement of the Jahn-Teller theorem, which asserts that the nuclear configuration of any nonlinear molecular system in a spatially degenerate electronic state is inherently unstable with respect to nuclear displacements that reduce the system's symmetry, thereby removing the degeneracy [73].

Although originally the Jahn-Teller effect was proposed in molecular systems, Van Vleck (1939) [103] extended the concept to solid-state systems with localized electrons.

The Jahn-Teller effect is particularly relevant in solid-state systems containing open-shell complexes, such as octahedral centers with O_h symmetry, characterized by electronic configurations d^7[104], d^9 [96; 105], p^5 [106], ... One example are insulator crystals composed by charged MX_n complexes where the M cation represents a transition metal characterized by an open d shell, and the nearest-neighbor ligands (n) are anions (F^-, Cl^-, O^{2-}, ...). These complexes exhibit significant electron localization, rendering them particularly susceptible to Jahn-Teller distortions.

As the Jahn-Teller effect arises in the presence of a degenerate electronic ground state, electron localization is essential. It enables stronger interactions between the localized electrons and the surrounding ligands. If the charge density is localized it exerts a strong force over the ligands distorting the lattice. This interaction will be examined in further detail later considering the case of the CuCl_6^{4-} complex.

Among the different examples in which the Jahn-Teller effect can be observed, we find pure solids such as $\mathrm{KCuF_3}$ [107; 108; 109], composed by CuF_6^{4-} complexes with d^9 configuration; CuO crystal [110; 109] characterized by CuO_6^{10-} complexes or systems with defects such as $\mathrm{NaCl:Cu^{2+}}$ involving the CuCl_6^{4-} complex [94]. In particular, to illustrate the vibronic coupling we study the Jahn-Teller effect in the $\mathrm{NaCl:Cu^{2+}}$ system [96]. In its pure form, NaCl consists of NaCl_6^{5-} octahedral complexes, exhibiting O_h symmetry group. Since $\mathrm{Na^+}$ cation is a closed shell ion, its electronic density is spherically symmetric, exerting a uniform force on the surrounding $\mathrm{Cl^-}$ ligands.



Compressed $D_{4\mathrm{h}}$ distortion
$Q_\theta < 0$ ($^2\mathrm{A_{1g}}$)

Reference state in O_h ($^2\mathrm{E_g}$)

Elongated $D_{4\mathrm{h}}$ distortion
$Q_\theta > 0$ ($^2\mathrm{B_{1g}}$)

Figure 3.4: Orbital splitting in the CuCl_6^{4-} octahedral complex induced by the Jahn–Teller effect, resulting in a tetragonal distortion with $D_{4\mathrm{h}}$ symmetry, either elongated or compressed along the z-axis. Figure extracted from Ref. [111].

At this point, if one of the $\mathrm{Na^+}$ cations is substituted by the impurity $\mathrm{Cu^{2+}}$, characterized by an open shell $3d^9$, it leads to a CuCl_6^{4-} complex. According to the molecular orbital diagram of the complex, shown in the center of Fig. 3.4, in the high-symmetry octahedral configuration the triplet level $t_{2\mathrm{g}}$ is complete while the highest occupied level corresponds to a doubly degenerate e_g state—associated with the atomic orbitals $d_{x^2-y^2}$ and d_{z^2}— where one orbital is fully occupied while the other is only partially filled. This leads to a twofold spatial degeneracy in its fundamental ground state. Then, the system undergoes a Jahn-Teller distortion lowering the symmetry to a tetragonal geometry. Since in the undistorted structure the e_g orbital is twofold degenerate and partially occupied, this uneven occupation results in an asymmetric electronic density. In the case where the completely occupied e_g orbital corresponds to the d_{z^2} orbital while $d_{x^2-y^2}$ is partially occupied, the localized nature of the electrons gives rise to anisotropic electronic density, which exerts directional forces on the surrounding ligands [see Fig. 3.5(c)]. This results in a distortion of the local coordination environment, with the ligands along the z-axis being pushed outward (because an excess of electron density in an antibonding orbital oriented in this direction), while those in the xy-plane are drawn closer to the central ion. Consequently, vibronic coupling through the Jahn-Teller effect induces a distortion along the normal mode Q_θ. The configurations corresponding to negative and positive values of Q_θ are illustrated on the left and right sides of Fig. 3.4, respectively. Specifically, for $Q_\theta > 0$, the mode results in an elongation of the two axial bond distances by an amount $2a$, while simultaneously compressing the four equatorial bond distances by an amount a. Conversely, for $Q_\theta < 0$, the complex suffers compression along the axial direction and an expansion within the equatorial plane.

As a result, the system undergoes a symmetry-lowering distortion from its original cubic configuration, becoming elongated along the z-axis and compressed within the xy-plane. This leads to a final geometry consistent with $D_{4\mathrm{h}}$ symmetry, as illustrated on the right in Fig. 3.4.

This symmetry breaking lifts the degeneracy of the e_g states: on the one hand, the d_{z^2} orbital, which now forms an $a_{1\mathrm{g}}$ state, is stabilized (its energy is lowered relative to the e_g doublet); on the other hand, the $d_{x^2-y^2}$ orbital, which corresponds to a $b_{1\mathrm{g}}$ state, is destabilized by an equivalent amount (in the linear approximation). Consequently, the system lowers its total energy compared to the undistorted O_h geometry, being the system in the high symmetry configuration unstable.

Returning to the high-symmetry configuration, when the occupied orbital corresponds to the $d_{x^2-y^2}$, vibronic coupling occurs with equatorial ligands instead of the z-axis. This asymmetric charge distribution exerts a repulsive force on the ligands within the xy-plane, leading to an expansion in that plane, accompanied by a compression along the z-axis [see Fig. 3.5(a)]. In this case, the splitting of the electronic energy levels follows the order illustrated on the left side of Fig. 3.4.

In the electronic band structure, this effect appears as a splitting of degenerate states at the top of the valence band or bottom of the conduction band, altering the bandgap and influencing both the spectroscopic properties and charge transport mechanisms of the material.

Furthermore, the Jahn-Teller effect plays a crucial role in metal-insulator transitions, as observed in manganites and perovskite oxides, where Jahn-Teller distortions induce phase transitions between conducting and insulating states [112].

Beyond the Jahn-Teller effect, certain systems exhibit the pseudo-Jahn-Teller (PJT) effect, where vibronic coupling induces geometric distortions even in the absence of degenerate electronic states. In such cases, the vibrational modes couple the ground electronic state with excited states, leading to a distortion of the molecular or crystalline geometry. An important difference between both effects is that the JT effect exerts a force and not necessarily involves a change in the density while the PJT effect necessarily involves a change in the density (covalency).

An illustrative example of a system exhibiting the pseudo-Jahn–Teller (PJT) effect is Cs_2AgF_4, a layered compound whose high-symmetry configuration is intrinsically tetragonal. Although the ground state in this configuration is non-degenerate, it is unstable due to vibronic coupling with low-lying excited

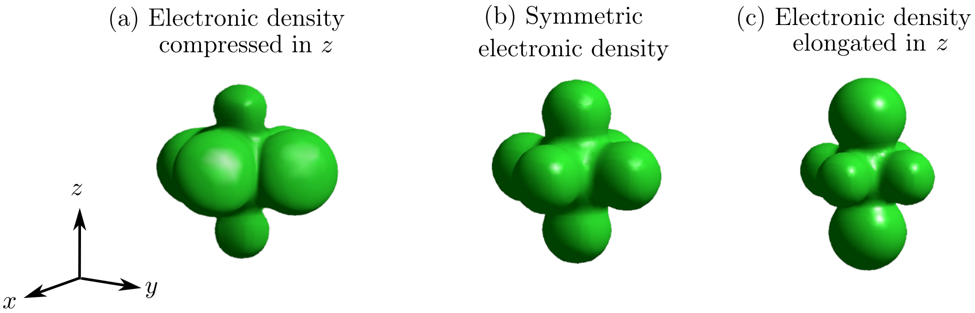


Figure 3.5: (a) Asymmetric electronic density, compressed along the z-axis and expanded in the xy-plane, for a $\mathrm{CuCl_6}^{4-}$ complex with an open-shell d^9 electronic configuration. This density arises from a configuration in which both electrons in the e_g level occupy the $d_{x^2-y^2}$ orbital, while the d_{z^2} orbital is partially occupied by a single electron. As a result, the increased electron density in the $d_{x^2-y^2}$ orbital leads to an expansion of the complex in the xy-plane (due to its antibonding character) and a compression along the z-axis, in contrast to the full-shell structure shown in (b). (b) Symmetric electronic density of an octahedral $\mathrm{NaCl_6}^{5-}$ complex with a closed-shell d^{10} configuration. (c) Asymmetric electronic density, elongated along the z-axis and compressed in the xy-plane, for a $\mathrm{CuCl_6}^{4-}$ complex with an open-shell d^9 configuration. This density corresponds to a configuration in which both electrons in the e_g level occupy the d_{z^2} orbital, while the $d_{x^2-y^2}$ orbital is partially occupied. Consequently, the dominant electron density in the d_{z^2} orbital induces an elongation of the complex along the z-axis, due to its antibonding character.

electronic states. This instability arises from the PJT effect, whereby the ground state interacts with an excited state through vibrational modes, leading to a distortion of the crystal structure [73]. Although the system may initially appear analogous to a Jahn–Teller system, the phonon modes coupled to the electronic states are fundamentally different from those involved in the JT effect. In the JT case, the relevant vibrational modes are Q_θ (representing a tetragonal distortion) and Q_ε (associated to an orthorombic distortion), which transform according to the e_g representation and mediate the electron–phonon interaction. In layered crystals, however, the system couples to a mode resembling the Q_ε mode.

It is important to note that in solids, the manifestation of the Jahn-Teller effect is inherently more intricate than in molecular systems, due to the fact that single-electron states form continuous energy bands in reciprocal space, as opposed to the discrete energy levels characteristic of isolated molecules. In order to perform a rigorous treatment of the JT effect in solids within reciprocal space, two competing theoretical frameworks have been developed: *orbital ordering theory* stablished by Kugel and Khomskii in 1982 [113; 114; 115; 116] and the *cooperative JT effect* proposed by Kanamori [117] and later extended by other researchers [118; 119; 120].

3.2.2 Vibronic coupling in spectroscopy

Vibronic coupling plays a fundamental role in spectroscopy, manifesting in various optical phenomena governed by vibronic coupling. This is particularly evident in processes such as optical absorption, observed in systems like organic dyes [121] and transition metal complexes [122]; fluorescence, manifested in quantum dots [123] and rare-earth-doped oxides [124] and Raman spectroscopy, measured in crystalline silicon [125] and graphene [126], where the coupling between electronic states and lattice vibrations influences spectral features and transition dynamics.

One significant manifestation of electron-phonon coupling is observed in thermally activated optical absorption transitions [127]. At low temperatures, certain optical transitions between electronic states are forbidden by selection rules. However, in some systems, as the temperature increases, the larger vibrations enhance the coupling between electronic states and lattice vibrations, enabling the observation of transitions which were initially symmetry-forbidden.

As an illustrative example, we consider a system initially in its electronic ground state, described by the wavefunction $\Psi^{(0)}$. The potential energy surface (PES) corresponding to this electronic state is depicted in Fig. 3.6 as a function of the vibrational normal mode coordinate Q. This coordinate represents the distortion pathway along a specific vibrational mode that becomes active upon thermal excitation of the system. In particular, we consider a structure in which the vibrational mode is non-totally symmetric, meaning it transforms according to a non-totally symmetric irrep [1], ($Q \neq A_1$). In other words, this mode does not correspond to an isotropic expansion or compression of the system.

At low temperatures ($T \approx 0$), the system remains in the equilibrium configuration of the ground state, residing at the minimum of the PES where the value of the wavefunction is $\Psi^{(0)}(Q_0)$. Under these conditions, we assume that an electric dipole transition between the ground state and the first electronic excited state $\Psi^{(1)}(Q_0)$ is forbidden by selection rules, i.e., the transition dipole moment is null

$$\langle \Psi^{(1)}(Q_0)|\vec{r}|\Psi^{(0)}(Q_0)\rangle = 0. \tag{3.4}$$

This is equivalent to prove that the direct product of the irreducible representations corresponding to $\Psi^{(0)}(Q_0)$, $\vec{r}$, and $\Psi^{(1)}(Q_0)$ does not contain the totally symmetric representation (A_1). In Fig. 3.6 this forbidden transition is represented by a vertical red dashed arrow.

Now, we apply on the structure a finite temperature ($T \neq 0$). Thus, the system experiences thermal excitations causing a distortion of the atomic geometry leading to a new geometry Q' (represented in Fig. 3.6 by the orange color). The resulting ground-state wavefunction for $T \neq 0$ is perturbed (later, we will see the added perturbation represents *first-order vibronic coupling*),

$$\Psi^{(0)}(\vec{r},Q') = \Psi^{(0)}(\vec{r},Q_0) + \left(\partial\Psi^{(0)}(\vec{r})/\partial Q\right)_{Q_0}(Q'-Q_0). \tag{3.5}$$

The thermally distorted wavefunction exhibits different symmetry properties compared to $\Psi^{(0)}(Q_0)$, as it involves a vibrational mode that does not possess the totally symmetric character.

To analyze the transition probability at $T \neq 0$, we evaluate the transition dipole moment considering as

---

[1] For a molecular system with symmetry described by a point group G, the *irreducible representations* (irreps) classify the ways in which different quantum states or operators transform under the symmetry operations of the group.

the initical state the perturbed wavefuntion,

$$\begin{aligned}\langle\Psi^{(1)}(Q')|\vec{r}|\Psi^{(0)}(Q')\rangle &= \langle\Psi^{(1)}(Q')|\vec{r}|\Psi^{(0)}(Q_0)\rangle + \left\langle\Psi^{(1)}(Q')|\vec{r}|\left(\partial\Psi^{(0)}/\partial Q\right)_{Q_0}(Q'-Q_0)\right\rangle = \\ &= \left\langle\Psi^{(1)}(Q')|\vec{r}|\left(\partial\Psi^{(0)}/\partial Q\right)_{Q_0}(Q'-Q_0)\right\rangle \neq 0\end{aligned} \tag{3.6}$$

In this expression, the first term vanishes due to selection rules, expressed in Eq. (3.4). However, the second term may be nonzero if the distortion transforms in such a way as to render the integrand totally symmetric, allowing the transition (represented by a green arrow in Fig. 3.6). As a result, the key to the transition becoming allowed lies in the change of the electronic wavefunction with respect to the nuclear positions.

In this way, the mixing between electronic states through the electron-lattice interaction can transform forbidden transitions at low temperatures into allowed transitions at $T \neq 0$, a phenomenon referred to as *vibronic transitions*. Such transitions, allowed by electron-phonon coupling, are a hallmark of the *Herzberg-Teller effect* [127], which describes how vibrational motion can relax electronic selection rules, enabling otherwise forbidden optical transitions.

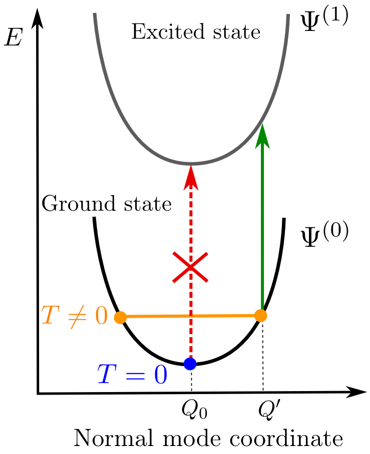


Figure 3.6: Schematic representation of the potential energy surfaces (PES) corresponding to the ground state, $\Psi^{(0)}$, and the excited state, $\Psi^{(1)}$, of a system. The PES are plotted as functions of a normal mode coordinate Q associated with a thermally activated vibrational mode (different from the totally symmetric mode). At zero temperature, the system resides at the equilibrium geometry Q_0, and the electronic transition from the ground to the excited state is symmetry-forbidden, as indicated by the red dashed vertical arrow. Upon thermal excitation, the system becomes vibrationally distorted, spending time in geometries displaced from Q_0, such as Q'. This distortion modifies the symmetry properties of the wavefunction through a first-order perturbative term $(\partial\Psi^{(0)}/\partial Q)_{Q_0}(Q'-Q_0)$, which may relax the selection rules. As a result, the transition dipole moment becomes non-zero, rendering the transition optically allowed. This thermally enabled transition is indicated by the solid green arrow.

Another phenomenon in spectroscopy related to vibronic coupling is the Stokes shift, which arises from the coupling between electronic states and lattice vibrations. This shift, observed in fluorescence and Raman spectroscopy, results from the energy relaxation of an excited electronic state due to vibrational interactions before photon emission, leading to an emission spectrum redshifted relative to the absorption spectrum [93; 94; 95], as shown on the right side of Fig. 3.7.

To explain this effect, we consider a system in the ground electronic state where the potential energy surface (PES) along a vibrational mode coordinate Q is given by the expression

$$E^{(0)}(Q) = \frac{1}{2}K(Q-Q_0)^2. \tag{3.7}$$

Here, the value Q_0 denotes the equilibrium position along the normal mode coordinate. The potential energy surface (PES) of the ground state is depicted in yellow in Fig. 3.7.

Upon the absorption of a photon, the system transitions to the first excited state. It should be noted that, according to the Franck-Condon principle [128], the transition occurs without an immediate change in the system's geometry, as nuclear motion is much slower than electronic transitions [128]. In the excited state, the electronic distribution of the system differs from that in the ground state, leading to a modification of the forces acting on the nuclei. This change induces vibronic coupling since these new forces lead to atomic displacements, causing the system to relax into a different equilibrium geometry (non-radiative relaxation process). This minimum of the excited state corresponds to the normal mode coordinate Q'. Thus, a shift is manifested in the potential energy surface (PES) of the excited state, $E^{(1)}$, which is displaced with respect to that of the ground state (see Fig. 3.7) due to the action of the force $F(Q-Q_0)$, where F is a constant with units of force per unit displacement,

$$E^{(1)}(Q) = E_{\text{exc}} + \frac{1}{2}K(Q-Q_0)^2 - F(Q-Q_0). \tag{3.8}$$

In the previous expression, the energy E_{exc} represents the energy of the vertical transition from the ground state to the excited state evaluated at Q_0. In subsequent sections we will see F represents the *linear vibronic coupling*.

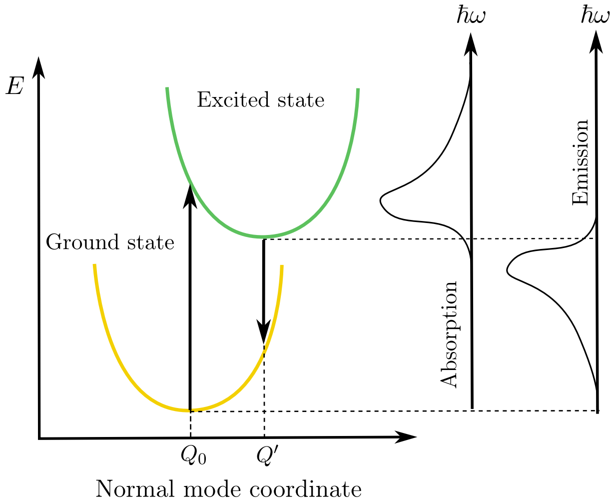


Figure 3.7: Diagram depicting the adiabatic potential energy surfaces of two electronic states in a molecule. Vertical arrows indicate vibronic transitions, accompanied by a schematic representation of the corresponding absorption and emission spectra. This figure has been adapted from Ref. [128].

After excitation, the system eventually relaxes from the excited state to the ground state. Due to non-radiative relaxation via vibronic coupling, the emitted photon has a lower energy than the absorbed photon. This energy difference leads to a shift between the maxima of the absorption and emission spectra, resulting the Stokes shift. Figure 3.7 illustrates the excitation (previously discussed) and decay processes and the corresponding absorption and emission spectra.

As an example of a system that exhibits a Stokes shift, we examine this phenomenon in a candidate material for the development of emissive layers in organic light-emitting diodes (OLEDs) [129; 130; 131]: the molecule known as 9-methoxypyrroloisoquinolinetrione. This compound exhibits a large Stokes shift in the transition $\pi \to \pi^*$. The excitation, associated with an absorption energy of 354 nm [132], leads to a relocalization of the electronic density. This redistribution of the electronic density is depicted in Fig. 3.8. In the first panel, the electronic density for the ground state, π, is shown in yellow. In the second panel, the electronic density of the excited state, π^*, is represented in green. In the third panel, the transition density is illustrated indicating in yellow regions of electron density depletion and in green isosurfaces associated to regions of electron density gain.

As previously discussed, the redistribution of electronic density in the excited state generates forces on the atoms, resulting in bond length changes: bonds with antibonding character lengthen, while others shorten. This constitutes the vibronic coupling.

Finally, the system undergoes desexcitation at 514 nm [132], producing the observed Stokes shift in the comparison of the absorption and emission spectra.

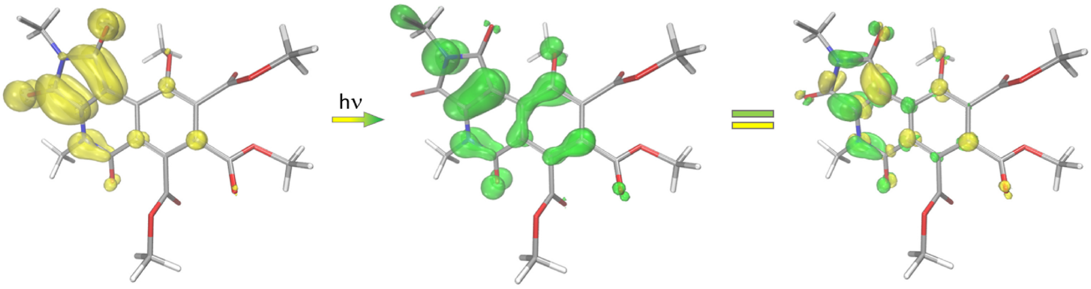


Figure 3.8: Electron density redistribution associated with the $\pi \to \pi^*$ electronic excitation in the 9-methoxypyrroloisoquinolinetrione molecule. The first panel illustrates the electron density corresponding to the initially occupied orbitals involved in the excitation, represented by yellow isosurfaces. The second panel depicts the spatial distribution of the unoccupied orbitals that are populated upon excitation, shown as green isosurfaces. The third panel presents the transition density, where yellow and green isosurfaces indicate regions of electron depletion and accumulation, respectively, capturing the net shift in electron density induced by the excitation. Figure extracted from Ref. [132] with permissions.

3.2.3 Localized states: polarons

A *polaron* is a quasiparticle that emerges from the interaction between a charge carrier (electron or hole) and the lattice of a material. When the system is doped with the charge carrier, it exerts forces on the surrounding atoms, inducing a local distortion of the lattice, as can be observed in Fig. 3.9. This distortion creates a potential well that energetically stabilizes the carrier leading to the localization of the charge carrier. The polaron thus represents the combined entity of the charge and its accompanying lattice deformation [128; 133].

Several physical phenomena are mediated by the presence of polarons. This quasiparticle significantly influences a wide range of processes, including charge transport [134; 135], colossal magnetoresistance [136], ferroic ordering [137], photoemission [138], and thermoelectricity [139].

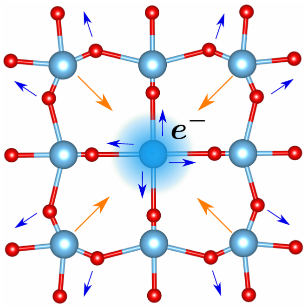


Figure 3.9: Schematic representation of a polaron. Blue and red spheres depict the atoms forming the lattice of the system. The blue cloud illustrates the electronic density associated with an excess electron that contributes to the formation of the polaron. This electron is localized on the central blue atom. In its vicinity, the electronic density induces a force that distorts the surrounding lattice, reflecting the electron–lattice coupling. Blue and orange arrows indicate the direction of atomic displacements resulting from this coupling, which characterizes the polaronic distortion.

Previously, we provided an overview of the definition of a polaron. Here, we further elaborate on the physical mechanism underlying its formation by presenting a schematic model that illustrates the process. In this framework, we consider a system in its ground state, whose potential energy surface (PES) is schematically depicted in green in Fig. 3.10. The associated potential energy is approximated by a harmonic potential of the form

$$E^{\mathrm{GS}}(Q) = \frac{1}{2}K(Q-Q_0)^2, \tag{3.9}$$

where Q denotes the normal mode coordinate, K is the force constant, and Q_0 corresponds to the equilibrium position of the system, i.e., the minimum of the PES. Next, we consider the introduction of an extra electron into the system. Initially, assuming that the added electron remains delocalized, the resulting energy gain is denoted by ΔE_{el}. Under this condition, the corresponding potential energy surface is described by the expression

$$E^{\mathrm{deloc}}(Q) = \Delta E_{\mathrm{el}} + \frac{1}{2}K(Q-Q_0)^2, \tag{3.10}$$

that is represented in blue in Fig. 3.10. However, the added electron can be localized forming a polaron. In this scenario, the electron exerts a force F in its local region, adding a contribution $-F(Q-Q_0)$ to the energy. This interaction effectively lowers the system's energy, leading to a modified PES given by

$$E^{\text{loc}}(Q) = \Delta E_{\text{el}} + \frac{1}{2}K(Q-Q_0)^2 - F(Q-Q_0), \tag{3.11}$$

which is represented in red in Fig. 3.10. Attending to the new PES, the distortion in the system reduces the energy, stabilizing the electron and keeping it localized in a region of space. The force exerted by the electron induces a local distortion in the lattice, resulting in a displacement of the surrounding atomic positions. It is through this distortion—arising from the electron–lattice interaction—that vibronic coupling emerges. This geometric distortion results in a new minimum of the potential energy described by Eq. (3.11). The equilibrium atomic positions are consequently shifted, reflecting the structural deformation induced by the localization of the electron.

The contributions to the total energy of the polaron are schematically illustrated in Fig. 3.10. The first contribution corresponds to the electronic energy gain, ΔE_{el}, associated with the introduction of an extra electron into the system. The second contribution, denoted as ΔE_{lat},

$$\Delta E_{\text{lat}} = |E^{\text{deloc}}(Q') - E^{\text{deloc}}(Q_0)| = \frac{1}{2}K(Q'-Q_0)^2, \tag{3.12}$$

represents the lattice distortion energy required to accommodate the localized charge. This energy cost is associated with the structural modification from the undistorted configuration at $E^{\text{deloc}}(Q_0)$ to the distorted geometry at $E^{\text{deloc}}(Q')$. Thirdly, the energy gain arising from the force performed by the electron leading to the electron–lattice coupling,

$$\Delta E_{\text{el-lat}} = |E^{\text{loc}}(Q') - E^{\text{deloc}}(Q')| = F(Q'-Q_0). \tag{3.13}$$

This energy is associated with lowering of the gap energy. In the case of strong coupling between the electron and the lattice distortions—typically associated with small polarons—the formation of a polaron leads to a highly localized electronic state, resulting in the emergence of a strongly localized mid-gap level. Conversely, for weakly coupled (large) polarons, the effect manifests as a slight lowering of the conduction band edge [88; 133].

Then, the energy of the polaron with respect the ground state is

$$E^{\text{loc}}(Q') - E^{\text{GS}}(Q_0) = \Delta E_{\text{el}} + \Delta E_{\text{lat}} - \Delta E_{\text{el-lat}} = \Delta E_{\text{el}} + \frac{1}{2}K(Q'-Q_0)^2 - F(Q'-Q_0). \tag{3.14}$$

Whereas, the energy associated to the stabilization of the polaron with respect the delocalized state

$$\Delta E_{\text{pol}} = \Delta E_{\text{el-lat}} - \Delta E_{\text{lat}} = F(Q'-Q_0) - \frac{1}{2}K(Q'-Q_0)^2 = \frac{1}{2}\frac{F^2}{K}. \tag{3.15}$$

In the last step of the equation, we apply the equilibrium condition, for which $\partial(\Delta E_{\text{pol}})/\partial Q' = 0$, obtained at $Q' = Q_0 + F/K$.

In the context of polarons the terms *self-trapped carrier* or *self-trapping* are often used. These terms describe the physical scenario in which the carrier becomes localized within a potential well, whose depth is given by $\Delta E_{\text{pol}} = F^2/2K$ [Eq. (3.15)], formed by the displacements of the surrounding atoms. This potential well lowers the energy of the system, energetically favoring localization of the carrier over the delocalized (free) state [133].

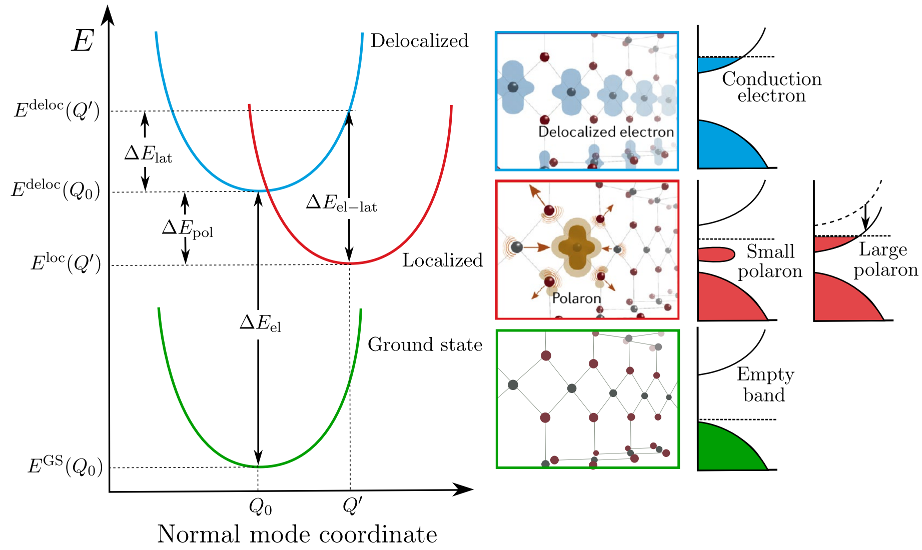


Figure 3.10: Configuration coordinate diagram illustrating the energy balance as a function of lattice distortion for both a delocalized conduction electron and a localized polaron. E_{ST} represents the structural energy, E_{POL} denotes the polaron binding energy, and E_{EL} corresponds to the electronic energy. The right side of the panel displays charge density isosurfaces alongside schematic band structures: the upper panel depicts a delocalized conduction electron, while the lower panel illustrates a small polaron with its characteristic polaron peak localized below the Fermi energy. Figure adapted from [88].

Historically, the idea behind self-localization effect, which emerges from lattice distortions in a perfect crystal, was first introduced by Landau in 1933 [140]. However, it was not until 1946 that Pekar [139] formally coined the term *polaron* to describe this phenomenon. Several experimental observations confirmed the presence of polarons in various systems [89; 90; 91]. Alongside these experiments, numerous theoretical frameworks were developed to investigate the nature, properties and dynamics of these quasiparticles. These advancements culminated in the classification of polarons into two types, distinguished by the spatial extent of the lattice distortions, which determines the radius of the polaron (r_{P}): large-radius polarons, formulated by Fröhlich [141; 142], and small-radius polarons, studied by Holstein [143; 144]. In their theories, they introduced quantum-field Hamiltonians establishing the foundational theoretical framework for subsequent developments [88].

Polarons formed due to short-range vibronic coupling are addressed by Holstein's theory [143; 144]. In this framework, the lattice is considered discrete. This model is primarily used to describe small polarons. Although it has been established that the most distinctive feature for the applicability of this model is the presence of strong electron–lattice coupling. Highly ionic solids, such as AgCl and alkali halides, are typical examples where this model is relevant [128].

In ionic systems, carriers exhibit self-trapping due to the strong lattice distortion that localizes the electron or hole. As a result, the movement of the polaron through the crystal is driven by hoppings between lattice sites, resulting in an incoherent motion. These polarons are characterized by reduced mobility, typically lower than $1\mathrm{cm}^2\mathrm{V}^{-1}\mathrm{s}^{-1}$ [88; 133; 135]. Notably, this mobility increases with temperature because thermal-induced distortions lead to phonon-assisted hopping.

Fröhlich polarons, which correspond to large-radius polarons, extend over multiple lattice cells. In the Fröhlich model, the discretization of the lattice is neglected, and the system is treated as a continuum. Such polarons are prevalent in ionic, polar, and ferroelectric materials due to the long-range electron-phonon coupling they exhibit. Large polarons are characterized by high mobility [145; 146; 147; 148], which is influenced by temperature. As the temperature increases, scattering with the phonon field becomes more significant, leading to a reduction in mobility.

The presence of large polarons can be observed in the band diagram. A polaron leads to a reduction in band dispersion, resulting in band narrowing. However, in contrast to small polarons, it does not generate a deep gap state. In this case, the polaron band is weakly bound to either the conduction or valence band, reflecting a less localized character compared to small polarons. This effect is schematically represented in Fig. 3.10, where the formation of the large polaron leads to a descend of the conduction band reducing the band gap.

3.2.4 BCS superconductivity

Superconductivity is a quantum phenomenon observed in certain materials at temperatures below a critical threshold T_c. This phase is characterized by the complete absence of electrical resistance and the manifestation of the Meissner effect [149], wherein the material exhibits perfect diamagnetism. Specifically, the magnetic susceptibility becomes $\chi = -1$, leading to a magnetic permeability of $\mu = 1 + \chi = 0$, which implies the expulsion of any externally applied magnetic field from the interior of the material.

The emergence of superconductivity lies in the manifestation of the instability of the conventional Fermi liquid ground state. In their seminal paper, Bardeen, Cooper and Schrieffer (BCS) [150] showed this instability drives the system into a new quantum ground state characterized by the formation of correlated electron pairs, known as *Cooper pairs*. The stabilization of this new state arises from an effective attractive interaction between electrons, which—despite their natural Coulomb repulsion—can occur through the mechanism of electron-phonon coupling. As an electron propagates through the crystal lattice, it induces a local distortion due to its interaction with the positively charged ions. This lattice distortion creates a potential that can attract a second electron, resulting in an effective attraction mediated by lattice vibrations, or phonons.

This electron-phonon-mediated attraction modifies the electronic states of the system and leads to the formation of an energy gap Δ at the Fermi surface. The behavior of the energy gap is governed by the BCS gap equation [151]

$$1 = \lambda \int_0^{\hbar\omega_D} \frac{d\varepsilon}{\sqrt{\varepsilon^2 + |\Delta|^2}} \tanh\left(\frac{\sqrt{\varepsilon^2 + |\Delta|^2}}{2k_BT}\right), \tag{3.16}$$

where ω_D is the Debye frequency and $\lambda = |g_{\text{eff}}|^2 g(\varepsilon_F)$ is the dimensionless electron-phonon coupling constant, with $g(\varepsilon_F)$ being the electronic density of states at the Fermi level. This equation implicitly determines the temperature dependence of the energy gap $\Delta(T)$. Physically, this gap represents the minimum energy required to break a Cooper pair. The excitation spectrum in the superconducting state is shifted such that the energy to add or remove a quasiparticle is $\pm E_k$, where $E_k > \Delta$. Consequently, the minimum energy required to create excitations is 2Δ, which suppresses low-energy scattering processes due to vibrations or impurities that would otherwise contribute to electrical resistance, thereby enabling the material to conduct electric current without dissipation.

In the original formulation of BCS theory [150], the electron-phonon interaction is treated within a weak-coupling approximation. In this regime, an increase in the electron-phonon coupling strength λ leads to a corresponding increase in both the superconducting energy gap Δ and the critical temperature T_c. Thus, the strength of this coupling plays a central role in determining the superconducting properties of a material.

For conventional superconductors, the electron-phonon interaction alone is sufficient to explain the formation of Cooper pairs. However, in the case of *high-*T_c superconductors, such as the cuprates, the standard BCS framework is inadequate because its strong electron-phonon coupling. Multiple theoretical models have been proposed to account for high-temperature superconductivity, but none has been universally accepted. Among these models, some highlight the potential importance of strong vibronic interactions, suggesting that lattice distortions may play a more dominant role in the pairing mechanism. In particular, polaronic and bipolaronic theories [87; 152; 153] propose that strong electron-lattice coupling can lead to localized bound states that contribute to unconventional superconducting behavior beyond the scope of the traditional BCS theory.

3.3 Electron-phonon interaction and its effect on the force constant

In the preceding section, we have presented several examples where vibronic coupling emerges as a fundamental interaction governing various physical phenomena. In particular, we discussed the important role played by the forces that electrons create when changing the state. In this Section, we establish the theoretical framework necessary to rigorously describe and analyze the interactions. Specifically, we examine how electron–vibration coupling influences the energy of the system and force constants. This understanding is crucial for interpreting results from first-principles simulations aimed at addressing the aforementioned problems.

The first step involves the definition of the vibronic Hamiltonian, which explicitly incorporates the electron-vibration interaction.

3.3.1 Vibronic Hamiltonian

One of the principal goals of quantum mechanics is to solve the Schrödinger equation and ascertain the electronic structures of atoms and molecules. This task is inherently complex because the wave function depends on the positions of numerous electrons, complicating the derivation of straightforward solutions. Until a few decades ago, the only feasible approach involved semi-empirical methods, which simplified the problem by incorporating extensive experimental data. However, with the advent of advanced computational techniques, it is now possible to solve the Schrödinger equation using only fundamental constants and nuclear atomic numbers, without relying on experimental results. These methods are called *ab initio*.

We operate under the assumption of the Born–Oppenheimer approximation. Firstly, we consider the time-independent electronic Schrödinger equation (stationary situation), which can be expressed in operator form as

$$H_{\mathrm{e}}(\vec{\tau})\Psi_i(\vec{r};\vec{\tau}) = E_i(\vec{\tau})\Psi_i(\vec{r};\vec{\tau}) \quad i = 0, 1, 2, \ldots \tag{3.17}$$

Here, the eigenstates $\{\Psi_i\}$ are multielectronic (many-body) states relevant to the problem (ground state and excitations) and E_i are the associated energies. The ground state is denoted as Ψ_0 and Ψ_i ($i = 1, 2, \ldots$) represent the excited states for the nuclear positions $\vec{\tau}$ ordered in increasing energy. The electronic coordinates are denoted by the vectors $\vec{r}$, whereas the atomic positions are represented by the vectors $\vec{\tau}$.

Since we are working on the Born–Oppenheimer approximation the wavefunction is confined to a single electronic potential energy surface, leading to an *adiabatic* description of the molecular system. As a result, within the framework of this approximation, the nuclei are assumed to move on an *adiabatic* potential energy surface (APES), $E_i(\vec{\tau})$, where the energy depends parametrically on the atomic positions.

In constructing the electronic Hamiltonian of the system, it is important to recognize that the spatial scale under consideration is on the order of angstroms, where electrostatic interactions constitute the predominant forces governing the system's behavior. Thus, the Hamiltonian associated with the problem will include the kinetic energy of the electrons (T_{e}), the electrostatic energy between nuclei (V_{nn}), electrons (V_{ee}), and between electrons and nuclei (V_{en}), as well as the external potential (V_{ext}). Therefore, the resulting electronic Hamiltonian H_{e} is given by

$$H_{\mathrm{e}} = T_{\mathrm{e}}(\vec{r}) + V_{\mathrm{ee}}(\vec{r}) + V_{\mathrm{en}}(\vec{r},\vec{\tau}) + V_{\mathrm{nn}}(\vec{\tau}). \tag{3.18}$$

It depends parametrically on the nuclear positions that appear in V_{en} and V_{nn} but not in their momenta, that are contained in the kinetic energy of the nuclei. The expressions for the electrons kinetic energy is

$$T_{\mathrm{e}}(\vec{r}) = -\hbar^2 \sum_{\boldsymbol{a}} \frac{\vec{\nabla}_{\boldsymbol{a}}^2}{2m}. \tag{3.19}$$

On the other hand, the repulsive Coulomb interaction between nuclei and between electrons is obtained from the following expressions, respectively

$$V_{\mathrm{nn}}(\vec{\tau}) = \frac{1}{4\pi\varepsilon_0} \sum_{\lambda} \sum_{\upsilon>\lambda} \frac{Z_\lambda Z_\upsilon e^2}{|\vec{\tau}_\lambda - \vec{\tau}_\upsilon|}, \tag{3.20}$$

$$V_{\mathrm{ee}}(\vec{r}) = \frac{1}{4\pi\varepsilon_0} \sum_{a} \sum_{b>a} \frac{e^2}{|\vec{r}_a - \vec{r}_b|}. \tag{3.21}$$

Here, Z_λ and Z_υ are the atomic numbers of the nuclei, e is the elementary charge, $\vec{\tau}_\lambda$ and $\vec{\tau}_\upsilon$ are the positions of the nuclei, and $\vec{r}_a$ and $\vec{r}_b$ are the positions of the electrons.

Finally, the electrostatic attractive interaction between electrons and nuclei is

$$V_{\mathrm{en}}(\vec{r},\vec{\tau}) = -\frac{1}{4\pi\varepsilon_0} \sum_{\lambda} \sum_{a} \frac{Z_\lambda e^2}{|\vec{\tau}_\lambda - \vec{r}_a|}. \tag{3.22}$$

Let us now assume a system in the reference configuration. Firstly, the electronic Schrödinger equation [Eq. (3.17)] at zero order of perturbations (without including the vibronic coupling corrections) is solved for the reference geometry $\vec{\tau}^{(0)}$ (a local minimum in the APES) as explained in Sec. 1.2.1,

$$H_{\mathrm{e}}^{(0)}(\vec{r};\vec{\tau}^{(0)})\Psi_i^{(0)}(\vec{r};\vec{\tau}^{(0)}) = E_i^{(0)}(\vec{\tau}^{(0)})\Psi_i^{(0)}(\vec{r};\vec{\tau}^{(0)}). \tag{3.23}$$

Secondly, the perturbed problem is addressed by introducing the perturbative correction. To define the perturbation, we identify and isolate the terms in the electronic Hamiltonian that explicitly depend on the nuclear coordinates,

$$V(\vec{r},\vec{\tau}) = V_{\mathrm{en}}(\vec{r},\vec{\tau}) + V_{\mathrm{nn}}(\vec{\tau}). \tag{3.24}$$

Then, it is possible to define the *vibronic operator*, $W(\vec{r},\vec{\tau})$ that represents the perturbation in H_{e} due to the motion of the nuclei with respect to the reference structure $\vec{\tau}^{(0)}$

$$W(\vec{r},\vec{\tau}) = V_{\mathrm{en}}(\vec{r},\vec{\tau}) - V_{\mathrm{en}}(\vec{r},\vec{\tau}^{(0)}) + V_{\mathrm{nn}}(\vec{\tau}) - V_{\mathrm{nn}}(\vec{\tau}^{(0)}). \tag{3.25}$$

Once defined the vibronic term, the electronic Hamiltonian can be expressed as

$$H_{\mathrm{e}}(\vec{r};\vec{\tau}) = H_{\mathrm{e}}(\vec{r};\vec{\tau}^{(0)}) + W(\vec{r},\vec{\tau}) \equiv H_{\mathrm{e}}^{(0)}(\vec{r}) + W(\vec{r},\vec{\tau}). \tag{3.26}$$

The operator $W(\vec{r},\vec{\tau})$ is typically expressed in terms of normal modes Q, rather than atomic displacements. These normal modes are labeled as Q_Γ, where the subscript Γ does not denote a mode at the high-symmetry Γ point but rather refers to any mode of the system. Upon performing the basis set transformation, the vibronic operator can be rewritten as $W(\vec{r},Q)$. Assuming small displacements of the nuclei around their equilibrium positions, the vibronic operator $W(Q)$ can be expanded as a Taylor series, with the equilibrium positions taken as the origin,

$$W(\vec{r},Q) = \sum_{\Gamma} \left(\frac{\partial W}{\partial Q_\Gamma}\right)_0 Q_\Gamma + \frac{1}{2} \sum_{\Gamma} \sum_{\Gamma'} \left(\frac{\partial^2 W}{\partial Q_\Gamma \partial Q_{\Gamma'}}\right)_0 Q_\Gamma Q_{\Gamma'} + \dots \tag{3.27}$$

Once the vibronic operator has been defined, we know the final expression of the electronic Hamiltonian. Then, it is possible to compute the energy of the electronic states considering the Schrödinger equation, Eq. (3.26). Now the objective consists on solving the electronic Schrödinger equation considering the Hamiltonian corrected by the perturbation,

$$E_i(Q) = \langle \Psi_i(Q) | H_\mathrm{e}(\vec{r};Q) | \Psi_i(Q) \rangle = \langle \Psi_i(Q) | H_\mathrm{e}^{(0)}(\vec{r};Q_0) + W(Q) | \Psi_i(Q) \rangle \,. \tag{3.28}$$

Here, the electronic wavefunctions $\Psi_i(Q)$ correspond to the system in the presence of the applied perturbation. To solve the eigenvalue problem directly, it is necessary to work with a complete, infinite-dimensional basis set in the Hilbert space. However, such an approach is generally infeasible. To address this issue the basis set is restricted to a finite-dimensional subspace of the Hilbert space and the Rayleigh-Ritz variational method [154] is employed to obtain the best solution. This method is based on the variational principle, which states that for a Hermitian operator, the lowest eigenvalue can be approximated by minimizing the energy expectation value within the chosen subspace.

The set $\{\Psi_i^{(0)}\}_{i=0}^{N_\Psi-1}$ defines a basis for the Hilbert space. A subset of this basis, $\{\Psi_i^{(0)}\}_{i=0}^{N_\Psi-1}$, consisting of the first N_Ψ wavefunctions, ordered by their respective energies (known as an adiabatic basis), can be used to represent the total wavefunction. This expansion accounts for nuclear coordinates and incorporates the variations in the electronic wavefunctions caused by atomic distortions,

$$\Psi_i(\vec{r},\vec{\tau}) = \sum_j c_{ij}(\vec{\tau}) \Psi_j^{(0)}(\vec{r};\vec{\tau}). \tag{3.29}$$

Hence, the energy of the i-th state, Eq.(3.28), is considered and expressed according to the expansion of Ψ_i in Eq. (3.29). Thus, the electronic energy is

$$E_i(Q) = \langle \Psi_i(Q) | H_\mathrm{e}(\vec{r};Q) | \Psi_i(Q) \rangle = \sum_j \sum_k c_{ij}^*(\vec{\tau}) c_{ik}(\vec{\tau}) \left\langle \Psi_j^{(0)} \middle| H_\mathrm{e} \middle| \Psi_k^{(0)} \right\rangle . \tag{3.30}$$

By applying the Rayleigh-Ritz variational principle, the energy is minimized with respect to the coefficients and their complex conjugates. In matrix notation, the obtained solution is

$$[H_\mathrm{e}(\vec{r};Q) - E_i(Q)]\, \vec{c}_i(Q) = 0. \tag{3.31}$$

Now, the hamiltonian H_e is expressed in the basis set $\{\Psi_i^{(0)}\}$. In this context, the overlap matrix is the identity since the states are orthonormal. The term $\vec{c}_i$ denotes the vector of coefficients associated with the i-th state. In the preceding equation, both the energy and the coefficients are unknown. With respect to $\vec{c}_i$, this constitutes a homogeneous system of linear equations, which admits a solution when the following determinant condition is satisfied,

$$\det |H_\mathrm{e}(\vec{r};Q) - E_i(Q)| = 0. \tag{3.32}$$

This corresponds to an eigenvalue problem, from which the energies $E_i(Q)$ can be determined, while the vectors $\vec{c}_i$ correspond to the eigenvectors. The vector $\vec{c}_i$ associated with the lowest eigenvalue E_i provides the coefficients that define the optimal $\Psi_i(\vec{r};Q)$ wavefunction, Eq. (3.29).

3.3.2 Vibronic constants

The energy of the i-th multielectronic state, $E_i(Q) = \langle \Psi_i(Q) | H_\mathrm{e}(Q) | \Psi_i(Q) \rangle$ [Eq. (3.31)], along with its associated wavefunction Ψ_i, can be determined by solving the variational problem outlined in Eq. (3.32). As previously discussed, the energy $E_i(Q)$ defines the adiabatic potential energy surface (APES) associated with electronic state i, upon which the nuclei move. To achieve this, the Hamiltonian matrix elements $\langle \Psi_j^{(0)} | H_\mathrm{e} | \Psi_k^{(0)} \rangle$ in the selected basis—frozen wavefunctions of the reference geometry, $\{\Psi_i\}_{i=0}^{N_\Psi - 1}$— must be calculated, which can be expressed as a sum of several contributing terms. The terms arising in this context can be categorized based on their symmetry properties: those involving totally symmetric modes are referred to as "elastic terms", while terms involving other modes are classified as "vibronic terms".

At this stage, the tools required to evaluate the Hamiltonian matrix elements are presented. The electronic Hamiltonian of the system, $H_\mathrm{e}(\vec{r}, \vec{\tau})$, incorporates vibronic coupling and is expressed in Eq. (3.26). Assuming that atomic displacements around the equilibrium geometry are small, H_e and its eigenfunctions can be expanded in a Taylor series following perturbation theory.

The electronic Hamiltonian $H_\mathrm{e}(\vec{r}, Q)$ can be expressed according to perturbation theory, up to second order, as

$$H_\mathrm{e}(\vec{r}, Q) = H_\mathrm{e}^{(0)} + \sum_{\Gamma} \left(\frac{\partial H_\mathrm{e}}{\partial Q_\Gamma} \right)_0 Q_\Gamma + \frac{1}{2} \sum_{\Gamma} \sum_{\Gamma'} \left(\frac{\partial^2 H_\mathrm{e}}{\partial Q_\Gamma \partial Q_{\Gamma'}} \right)_0 Q_\Gamma Q_{\Gamma'} + \dots \tag{3.33}$$

In this context, the wavefunction $\Psi_i(\vec{r}, Q)$ can be expressed according to perturbation theory, up to second order, as

$$\Psi_i(\vec{r}, Q) = \Psi_i^{(0)} + \sum_{\Gamma} \left(\frac{\partial \Psi_i}{\partial Q_\Gamma} \right)_0 Q_\Gamma + \frac{1}{2} \sum_{\Gamma} \sum_{\Gamma'} \left(\frac{\partial^2 \Psi_i}{\partial Q_\Gamma \partial Q_{\Gamma'}} \right)_0 Q_\Gamma Q_{\Gamma'} + \dots \tag{3.34}$$

By considering the electronic Hamiltonian expansion in Eq. (3.26) and the Taylor expansion of the vibronic operator around the reference atomic geometry, Eq. (3.27), the following relationships hold: $\partial H_\mathrm{e} / \partial Q_\Gamma = \partial W / \partial Q_\Gamma$ and $\partial^2 H_\mathrm{e} / \partial Q_\Gamma \partial Q_{\Gamma'} = \partial^2 W / \partial Q_\Gamma \partial Q_{\Gamma'}$.

After the aforementioned clarifications, the next step in solving Eq. (3.32) involves calculating the matrix elements of the Hamiltonian [see Eq. (3.28)] in the $\{\Psi_i^{(0)}\}_{i=0}^{N_\Psi - 1}$ basis,

$$\begin{aligned} \left\langle \Psi_j^{(0)} \middle| H_\mathrm{e} \middle| \Psi_k^{(0)} \right\rangle = \left\langle \Psi_j^{(0)} \middle| H_\mathrm{e}^{(0)} + W(Q) \middle| \Psi_k^{(0)} \right\rangle = E_j^{(0)} + \left\langle \Psi_j^{(0)} \middle| \sum_{\Gamma} \left(\frac{\partial W(Q)}{\partial Q_\Gamma} \right)_0 Q_\Gamma \middle| \Psi_k^{(0)} \right\rangle + \\ + \left\langle \Psi_j^{(0)} \middle| \sum_{\Gamma} \sum_{\Gamma'} \left(\frac{\partial^2 W(Q)}{\partial Q_\Gamma \partial Q_{\Gamma'}} \right)_0 Q_\Gamma Q_{\Gamma'} \middle| \Psi_k^{(0)} \right\rangle . \end{aligned} \tag{3.35}$$

Various terms arise in connection with the derivatives of the vibronic operator, that will be connected to the change of the energy with the displacement. The matrix elements of the vibronic operator in the electronic basis are called to as *vibronic constants*. These constants quantify the strength of interaction between nuclear displacements and the electronic structure. In essence, they describe the extent to which the electronic structure is altered due to nuclear motion and, conversely, how nuclear dynamics are influenced by variations in electronic distribution.

3.3.2.1 Linear vibronic constants

In the context of the Hamiltonian matrix elements computed in Eq. (3.35), two distinct types of vibronic coupling constants arise. The first type corresponds to the linear coupling terms, associated with the first derivative of the Hamiltonian with respect to a normal mode coordinate Q_Γ, while the second type involves quadratic terms, which are linked to the second derivative. In the present Section, we focus exclusively on the linear contributions.

When analyzing vibronic coupling constants, it is essential to distinguish between diagonal and off-diagonal matrix elements. Diagonal elements refer to cases in which both electronic states involved are identical, whereas off-diagonal elements pertain to interactions between different electronic states.

We begin by considering the diagonal linear terms in Eq. (3.35). For a specific normal mode Γ, the corresponding linear vibronic coupling constant takes the form:

$$F_\Gamma^{(j)} = \left\langle \Psi_j^{(0)} \middle| \left(\frac{\partial W}{\partial Q_\Gamma} \right)_0 \middle| \Psi_j^{(0)} \right\rangle . \tag{3.36}$$

Applying the Hellman-Feynman theorem [155], we find

$$F_\Gamma^{(j)} = \left\langle \Psi_j^{(0)} \middle| \left(\frac{\partial W}{\partial Q_\Gamma} \right)_0 \middle| \Psi_j^{(0)} \right\rangle = \left(\frac{\partial}{\partial Q_\Gamma} \left\langle \Psi_j^{(0)} \middle| H \middle| \Psi_j^{(0)} \right\rangle \right)_0 . \tag{3.37}$$

The diagonal constants of the linear coupling, $F_\Gamma^{(j)}$, represent minus the force with which the electrons in state $\Psi_j^{(0)}$ distort the nuclear configuration in the directions defined by the Q_Γ mode.

If we examine a system in a non-degenerate state and in the reference configuration $\vec{\tau}^{(0)}$ (associated to the $Q = 0$ configuration) which represents a critical point of the APES, the first derivative of the adiabatic potential energy surface (APES) with respect the Q_Γ mode vanishes. As a result $F_\Gamma^{(j)} = 0$.

The coefficients $F_\Gamma^{(j)}$ can be expressed in terms of *orbital vibronic constants*. In particular, representing the vibronic constants in a molecular orbital basis is especially useful for comparison with values obtained from other computational approaches, such as second-principles calculations, which constitute one of the primary methodologies employed in this work. A detailed analysis of the vibronic constants in second-principles calculations will be presented in Sec. 3.5.

To define these orbital vibronic constants, we consider the additive nature of the vibronic operator $W(\vec{r}, \vec{\tau})$, Eq.(3.25), with respect to the electrons. It allows to express $W(\vec{r}, \vec{\tau})$ as a sum of one-electron operators,

$$\begin{aligned} & W(\vec{r}, \vec{\tau}) = \sum_{a=1}^{N} w_a(\vec{r}_a, \vec{\tau}) \\ & \text{where} \\ & w_a(\vec{r}_a, \vec{\tau}) = -\sum_\lambda \frac{e^2 Z_\lambda}{|\vec{r}_a - \vec{\tau}_\lambda|} + \sum_\lambda \frac{e^2 Z_\lambda}{|\vec{r}_a - \vec{\tau}_\lambda^{(0)}|} + \frac{1}{N} \sum_{\lambda \neq \upsilon} \frac{e^2 Z_\lambda Z_\upsilon}{|\vec{\tau}_\lambda - \vec{\tau}_\upsilon|} - \frac{1}{N} \sum_{\lambda \neq \upsilon} \frac{e^2 Z_\lambda Z_\upsilon}{|\vec{\tau}_\lambda^{(0)} - \vec{\tau}_\upsilon^{(0)}|} \end{aligned} \tag{3.38}$$

and N is the number of electrons. This additivity of the vibronic operator allows the simplification of the vibronic terms. We define the *linear orbital vibronic constants* as

$$f_\Gamma^{ab} = \langle \phi_a | \left(\frac{\partial w_a(Q)}{\partial Q_\Gamma} \right)_0 | \phi_b \rangle \tag{3.39}$$

where the functions ϕ_a and ϕ_b represent one-electron molecular orbitals (MOs).

In order to show the relation between the linear constants $F_\Gamma^{(j)}$ and the orbital vibronic constants f_Γ^{ab}, we assume $\Psi_i^{(0)}$ can be described by one only determinant. Expressing it in a multiplicative format, we obtain

$$\left| \Psi_i^{(0)} \right\rangle = \det|\phi_1 \phi_2 \dots \phi_n|. \tag{3.40}$$

It is composed by one-electron molecular orbitals ϕ_a (MOs) which conform an orthonormal basis set. At this point, substituting Eq. (3.40) in Eq. (3.36) we can express the vibronic operator as the sum of its one-electron operators [156],

$$F_\Gamma^{(j)} = \left\langle \Psi_j^{(0)} \middle| \left(\frac{\partial W}{\partial Q_\Gamma} \right)_0 \middle| \Psi_j^{(0)} \right\rangle = \sum_{a=1}^{n} o_a \langle \phi_a | \frac{\partial w_a}{\partial Q_\Gamma} | \phi_a \rangle = \sum_{a=1}^{n} o_a f_\Gamma^{(aa)} \tag{3.41}$$

where f_Γ^{aa} represent the diagonal linear orbital vibronic constants and o_a is the occupation number of the molecular orbital ϕ_a.

Thus, we can see that the elements f_Γ^{aa} represent the force exerted by the electrons in the molecular orbital ϕ_a on the nuclear configuration along the directions defined by the vibrational mode Q_Γ. In order to calculate the total force $F_\Gamma^{(j)}$ we have to correct it by subtracting the contribution from the internuclear repulsion along the same direction. Finally, the physical interpretation of Eq. (3.41) indicates that the total force exerted by all the electrons in the system can be represented as the sum of the forces exerted by each individual one-electron molecular orbital.

Let's now move to the off-diagonal linear vibronic constants.

The off-diagonal linear coupling constants, $F_\Gamma^{(jk)}$, are defined as

$$F_\Gamma^{(jk)} = \left\langle \Psi_j^{(0)} \middle| \left(\frac{\partial W}{\partial Q_\Gamma} \right)_0 \middle| \Psi_k^{(0)} \right\rangle. \tag{3.42}$$

When analyzing the linear terms for which the electronic states $\Psi_j^{(0)}$ and $\Psi_k^{(0)}$ correspond to two distinct configurations, a particularly simple scenario arises if these states are single-determinant wavefunctions related by a one-electron excitation, $j \rightarrow k$, under the frozen orbital approximation. Specifically, this transition is described by a single excitation from molecular orbital ϕ_a to molecular orbital ϕ_b, denoted $a \rightarrow b$, while the remaining molecular orbitals remain unaffected. In this context, the electronic wavefunctions $\Psi_j^{(0)}$ and $\Psi_k^{(0)}$ are Slater determinants that differ solely by the occupation of the molecular orbitals ϕ_a and ϕ_b. Using the usual rules [156] of evaluation of matrix elements with Slater determinants we find,

$$F_\Gamma^{(jk)} = \left\langle \Psi_j^{(0)} \middle| \left(\frac{\partial W}{\partial Q_\Gamma} \right)_0 \middle| \Psi_k^{(0)} \right\rangle = \sum_{a',b'} c_{a'j} c_{b'k} \langle \phi_{a'} | \left(\frac{\partial W}{\partial Q_\Gamma} \right)_0 | \phi_{b'} \rangle = \langle \phi_a | \left(\frac{\partial W}{\partial Q_\Gamma} \right)_0 | \phi_b \rangle, \tag{3.43}$$

where the indexes a' and b' run over the one-electron molecular orbitals. Here, the only term that remains is the one associated with the orbitals ϕ_a and ϕ_b, as they are the only molecular orbitals involved in the excitation. All other orbitals remain unchanged and, therefore, do not contribute to the integral.

The linear vibronic coupling constants are directly related to the force constants of the system. In the following section, this relationship will be formally established.

3.3.2.2 Force constants

The vibronic linear coupling constants, analyzed in the preceding Section, are intrinsically related to the force constants of the system. These force constants are associated with the curvature of the adiabatic potential energy surface (APES), and thus provide insight into the stability of the system at a given configuration. In this section, we will explicitly examine the relationship between the vibronic linear constants and the force constants.

Considering $\vec{\tau}^{(0)}$ a critical point of the energy surface, we can study its stability with respect to displacements along Q_Γ by observing the curvature of the APES with respect to that mode. In that configuration, the system will be stable if the curvature is positive (a local minimum in the APES) and it will be unstable if the curvature is negative (a local maximum in the APES).

At this point, we will assume that we are dealing with a system without electronic degeneracy, and therefore, we will place particular emphasis on the diagonal terms concerning vibrational modes. Then, the curvature of the APES $K_\Gamma^{(i)}$ is defined as,

$$K_\Gamma^{(i)} = \left(\frac{\partial^2 E_i}{\partial Q_\Gamma^2}\right)\Bigg|_{Q=0} \tag{3.44}$$

where

$$E_i(Q) = \langle \Psi_i(Q)|H_\mathrm{e}(Q)|\Psi_i(Q)\rangle\,. \tag{3.45}$$

In order to compute the first derivative, the Hellman-Feynman [155; 157] theorem is applied resulting

$$\frac{\partial E_i}{\partial Q_\Gamma} = \langle \Psi_i|\frac{\partial H_\mathrm{e}}{\partial Q_\Gamma}|\Psi_i\rangle\,. \tag{3.46}$$

Knowing the first derivative, the second derivative is computed taking into account the Leibniz integral rule [158],

$$\frac{\partial^2 E_i}{\partial Q_\Gamma^2} = \frac{\partial}{\partial Q_\Gamma}\langle \Psi_i|\frac{\partial H_\mathrm{e}}{\partial Q_\Gamma}|\Psi_i\rangle = \left\langle \frac{\partial \Psi_i}{\partial Q_\Gamma}\middle|\frac{\partial H_\mathrm{e}}{\partial Q_\Gamma}\middle|\Psi_i\right\rangle + \langle \Psi_i|\frac{\partial^2 H_\mathrm{e}}{\partial Q_\Gamma^2}|\psi_i\rangle + \left\langle \Psi_i\middle|\frac{\partial H_\mathrm{e}}{\partial Q_\Gamma}\middle|\frac{\partial \Psi_i}{\partial Q_\Gamma}\right\rangle. \tag{3.47}$$

The Hamiltonian operator, H_e, is a Hermitian operator, which implies that $H^\dagger = H$. This fundamental property ensures,

$$\left\langle \Psi_i\middle|\frac{\partial H_\mathrm{e}}{\partial Q_\Gamma}\middle|\frac{\partial \Psi_i}{\partial Q_\Gamma}\right\rangle = \left\langle \Psi_i\middle|\frac{\partial H_\mathrm{e}^\dagger}{\partial Q_\Gamma}\middle|\frac{\partial \Psi_i}{\partial Q_\Gamma}\right\rangle = \left\langle \Psi_i\middle|\left(\frac{\partial H_\mathrm{e}}{\partial Q_\Gamma}\right)^\dagger\middle|\frac{\partial \Psi_i}{\partial Q_\Gamma}\right\rangle = \left\langle \frac{\partial \Psi_i}{\partial Q_\Gamma}\middle|\frac{\partial H_\mathrm{e}}{\partial Q_\Gamma}\middle|\Psi_i\right\rangle^* \tag{3.48}$$

where the hermitian property has been applied in the firts step. By applying the aforementioned equality to Eq. (3.47), the expression reduces to the sum of a scalar and its complex conjugate. Consequently, the imaginary part cancels out, leaving a result that is twice the real part,

$$\frac{\partial^2 E_i}{\partial Q_\Gamma^2} = \langle \Psi_i | \frac{\partial H_\mathrm{e}}{\partial Q_\Gamma} | \Psi_i \rangle + 2\mathrm{Re}\left(\left\langle \frac{\partial^2 \Psi_i}{\partial Q_\Gamma^2} \middle| \frac{\partial H_\mathrm{e}}{\partial Q} \middle| \Psi_i \right\rangle \right). \tag{3.49}$$

The last step is the evaluation of the previous expression in the equilibrium position considering for the wavefunction the expression from the perturbation theory, Eq.(3.34) with $\Psi_i(Q=0) \equiv \Psi_i^{(0)}$

$$\left(\frac{\partial^2 E_i}{\partial Q_\Gamma^2}\right)_0 = \left\langle \Psi_i^{(0)} \middle| \left(\frac{\partial^2 H_\mathrm{e}}{\partial Q_\Gamma^2}\right)_0 \middle| \Psi_i^{(0)} \right\rangle + 2\mathrm{Re}\left(\left\langle \Psi_i^{(0)} \middle| \left(\frac{\partial H_\mathrm{el}}{\partial Q_\Gamma}\right)_0 \middle| \left(\frac{\partial \Psi_i}{\partial Q_\Gamma}\right)_0 \right\rangle \right). \tag{3.50}$$

According to perturbation theory around the equilibrium position ($Q = 0$), the first correction of the wavefunction satisfies

$$\left| \left(\frac{\partial \Psi_i}{\partial Q}\right)_0 \right\rangle = \sum_{j \neq i} \frac{\left\langle \Psi_j^{(0)} \middle| (\partial H_\mathrm{e}/\partial Q)_0 \middle| \Psi_i^{(0)} \right\rangle}{E_i^{(0)} - E_j^{(0)}} \left| \Psi_j^{(0)} \right\rangle. \tag{3.51}$$

Inserting this equality in Eq. (3.50),

$$\left(\frac{\partial^2 E_i}{\partial Q_\Gamma^2}\right)_0 = \left\langle \Psi_i^{(0)} \middle| \left(\frac{\partial^2 H_\mathrm{e}}{\partial Q_\Gamma^2}\right)_0 \middle| \Psi_i^{(0)} \right\rangle - 2\sum_{j \neq i} \frac{\left| \left\langle \Psi_i^{(0)} \middle| (\partial H_\mathrm{e}/\partial Q)_0 \middle| \Psi_j^{(0)} \right\rangle \right|^2}{E_j^{(0)} - E_i^{(0)}}. \tag{3.52}$$

Considering the definition of H_e in Eq.(3.26), it is the sum of $H_\mathrm{e}^{(0)}$ which does not depend on Q and the vibronic interaction, W which is a function of Q. As a result, the previous expression can be expressed employing the vibronic operator instead of the electronic Hamiltonian,

$$\left(\frac{\partial^2 E_i}{\partial Q_\Gamma^2}\right)_0 = \left\langle \Psi_i^{(0)} \middle| \left(\frac{\partial^2 W}{\partial Q_\Gamma^2}\right)_0 \middle| \Psi_i^{(0)} \right\rangle - 2\sum_{j \neq i} \frac{\left| \left\langle \Psi_i^{(0)} \middle| (\partial W/\partial Q)_0 \middle| \Psi_j^{(0)} \right\rangle \right|^2}{E_j^{(0)} - E_i^{(0)}}. \tag{3.53}$$

According to the obtained expression for the curvature of the APES [Eq. (3.53)], $K_\Gamma^{(i)}$, it can be expressed as the sum of the following terms,

$$K_\Gamma^{(i)} = K_{0,\Gamma}^{(i)} + K_{\mathrm{v},\Gamma}^{(i)}. \tag{3.54}$$

The main effect of the first term

$$K_{0,\Gamma}^{(i)} = \left\langle \Psi_i^{(0)} \middle| \left(\frac{\partial^2 W}{\partial Q_\Gamma^2}\right)_0 \middle| \Psi_i^{(0)} \right\rangle, \tag{3.55}$$

is to describe the force constant created by the initial density when this density is not allowed to relax to change the geometry of the system. The second term represents the vibronic contribution to the curvature,

$$K_{\mathrm{v},\Gamma}^{(i)} = -2\sum_{j \neq i} \frac{\left| \left\langle \Psi_i^{(0)} \middle| (\partial W/\partial Q)_0 \middle| \Psi_j^{(0)} \right\rangle \right|^2}{E_j^{(0)} - E_i^{(0)}}. \tag{3.56}$$

This term is associated with the pseudo-Jahn–Teller effect, which describes the mixing of the ground electronic state with an excited state. The corresponding force constant is related to the linear vibronic coupling constants $F_\Gamma^{(ij)}$ defined in Eq. (3.36),

$$K_{v,\Gamma}^{(i)} = -2\sum_{j\neq i} \frac{\left|F_\Gamma^{(ij)}\right|^2}{E_j^{(0)} - E_i^{(0)}}. \tag{3.57}$$

The constant K_v is referred to as the *vibronic term* because it accounts for the coupling between ground electronic state and excited electronic states due to lattice vibrations to express the charge of the electron density with the vibrations. This mixing of electronic states leads to the correction of the wavefunction under the atomic distortions perturbation $\partial\Psi/\partial Q$.

3.3.2.3 Quadratic vibronic coupling

Once the linear coupling constants from the Taylor expansion in Eq. (3.35) and their relationship to the force constants have been established, the next step is to introduce the quadratic terms of the same expansion. These quadratic vibronic coupling terms are represented by

$$G_{\Gamma\Gamma'}^{ij} = \left\langle \Psi_i^{(0)} \middle| \left(\frac{\partial^2 W}{\partial Q_\Gamma \partial Q_{\Gamma'}} \right)_0 \middle| \Psi_j^{(0)} \right\rangle, \tag{3.58}$$

If Q_Γ and $Q_{\Gamma'}$ are non-degenerate modes, $G_{\Gamma\Gamma'}^{(ij)}$ is non-zero only if $\Gamma = \Gamma'$.

The term $G_{\Gamma\Gamma}^{ii}$ represents K_0, the first term of Eq. (3.55) elastic constant defined in the pseudo Jahn Teller theory. K_0 is referred to as the elastic term (non-vibronic) because it does not mix states.

Similar to Eq. (3.39), if the second order correction to the vibronic coupling is expressed projecting onto the one-electron molecular orbitals $\{\phi_a\}$, the resulting elements are the *quadratic orbital vibronic constants*,

$$g_{\Gamma\Gamma'}^{ab} = \frac{1}{2} \langle \phi_a | \left(\frac{\partial^2 W(Q)}{\partial Q_\Gamma \partial Q_{\Gamma'}} \right)_0 | \phi_b \rangle. \tag{3.59}$$

3.4 Electron-phonon interaction from first-principles

In previous sections, the linear terms of the vibronic operator (Sec. 3.3.2.1) and the force constants (Sec. 3.3.2.2) have been derived by applying the Hellmann-Feynman theorem. These results hold under the assumption that the employed wavefunction corresponds to the exact solution of the Schrödinger equation. However, in practical applications, such as first-principles simulations, the wavefunction is generally not exact but rather incomplete, since completeness is only attained in the theoretical limit of an infinite basis set of wavefunctions. In the framework of an incomplete basis set, the difference between the derivative of the energy and the expectation value of the derivative of the Hamiltonian (equality hold in Hellman-Feynman theorem) is often non-negligible [102; 159]. Consequently, to address this issue in computational calculations, it is essential to include an additional correction term to ensure accurate results when the wavefunction is approximate.

This Section is devoted to analyzing the implications of basis set incompleteness and introducing the corresponding corrections, commonly referred to as Pulay terms, which bridge the gap between the ideal Hellmann-Feynman framework and practical implementations. This discussion provides a foundation for connecting the ideal setting described and the working equations typically used in first-principles simulations. In particular, within the scope of this work, first-principles calculations are carried out using density functional theory.

Finally, in this Section, we extend the treatment of the electron-lattice interaction to periodic systems, enabling the analysis of problems in solid-state physics through a formulation in reciprocal space. Additionally, we examine the representation of these interactions within the Wannier function formalism.

Accordingly, the first step for the described study involves introducing the Hamiltonian used in density functional theory.

3.4.1 Hamiltonian in Density Functional Theory

In first-principles calculations, the behavior of the system can be studied by employing a Hamiltonian composed of three contributions:

$$H = H_{\mathrm{e}} + H_{\mathrm{p}} + H_{\mathrm{e-p}}, \tag{3.60}$$

where H_{e} represents the electronic contribution, H_{p} accounts for the lattice effects, and $H_{\mathrm{e-p}}$ encompasses the electron-phonon interaction.

Electronic Hamiltonian

In the framework of density functional theory, the Kohn-Sham density is [160]

$$n(\vec{r}) = \sum_{\boldsymbol{a}} |\phi^{S}_{\boldsymbol{a}}(\vec{r})|^2. \tag{3.61}$$

where the molecular orbitals $\phi^{S}_{\boldsymbol{a}}$ yield the same electron density as the original interacting system. The associated one electron Hamiltonian is [160]

$$h_{\boldsymbol{a}} = -\frac{\hbar^2}{2m_e}\nabla^2_{\boldsymbol{a}} + v^{\mathrm{KS}}(\vec{r}_{\boldsymbol{a}};\{\vec{\tau}_{\lambda}\}) \tag{3.62}$$

and the corresponding Kohn-Sham Schrödinger-like equations are

$$h_{\boldsymbol{a}}\phi^{S}_{\boldsymbol{a}}(\vec{r}) = \varepsilon_{\boldsymbol{a}}\phi^{S}_{\boldsymbol{a}}(\vec{r}). \tag{3.63}$$

The orbitals $\{\phi_i^S\}$ in the Kohn-Sham system are solutions to one-electron equations similar to the Hartree equations, but including, in this case, the exchange and correlation terms integrated in the $v^{\mathrm{KS}}(\vec{r}_{\boldsymbol{a}},\{\vec{\tau}_{\boldsymbol{\lambda}}\})$ potential. The exact solution to the Schrödinger equation in the case of non-interacting electron is given as a Slater determinant composed of those molecular orbitals, ϕ_i, $\Psi = |\phi_1^S \ldots \phi_n^S|$. The scalars ε_i^S are the eigenvalues associated with the molecular orbitals $\phi_{\boldsymbol{a}}^S$ and the potential $v^{\mathrm{KS}}(\vec{r}_{\boldsymbol{a}},\{\vec{\tau}_{\boldsymbol{\lambda}}\})$ is an effective potential for the non-interacting system. The obtained v^{KS} potential follows

$$v^{\mathrm{KS}}(\vec{r}_{\boldsymbol{a}};\{\vec{\tau}_{\boldsymbol{\lambda}}\}) = \frac{e^2}{4\pi\varepsilon_0}\int \frac{n(\vec{r}\,';\{\vec{\tau}_{\boldsymbol{\lambda}}\})}{|\vec{r}_{\boldsymbol{a}} - \vec{r}\,'|} d\vec{r}\,' - \frac{e^2}{4\pi\varepsilon_0}\sum_{\boldsymbol{\lambda}} \frac{Z_{\boldsymbol{\lambda}}}{|\vec{r}_{\boldsymbol{a}} - \vec{\tau}_{\boldsymbol{\lambda}}|} + v^{\mathrm{xc}}(\vec{r}_{\boldsymbol{a}},\{\vec{\tau}_{\boldsymbol{\lambda}}\}). \tag{3.64}$$

It is the sum of the the Hartree electronic screening, the electronic-nuclear potential, and the exchanges and correlation potential, respectively.

Lattice vibrations

In order to study the lattice vibrations, the total potential energy, U, is expanded around the RAG, which represents an equilibrium point of the APES. The considered displacements from the RAG are sufficiently small to handle the harmonic approximations, expanding the potential energy until second order. lattice vibrations

$$U = U_0 + \frac{1}{2}\sum_{\boldsymbol{\lambda}\alpha}\sum_{\boldsymbol{\upsilon}\beta} \left.\frac{\partial^2 U}{\partial\tau_{\boldsymbol{\lambda}\alpha}\partial\tau_{\boldsymbol{\upsilon}\beta}}\right|_{\vec{\tau}^{(0)}} (\tau_{\boldsymbol{\lambda}\alpha} - \tau^{(0)}_{\boldsymbol{\lambda}\alpha})(\tau_{\boldsymbol{\upsilon}\beta} - \tau^{(0)}_{\boldsymbol{\upsilon}\beta}), \tag{3.65}$$

where the value U_0 represents the potential energy of the ions in the equilibrium positions, $\vec{\tau}^{(0)}$ and α and β the cartesian coordinates. Moreover, the Hamiltonian describing the phonons is,

$$H_{\mathrm{p}} = \frac{1}{2}\sum_{\boldsymbol{\lambda}\alpha}\sum_{\boldsymbol{\upsilon}\beta} \left.\frac{\partial^2 U}{\partial\tau_{\boldsymbol{\lambda}\alpha}\partial\tau_{\boldsymbol{\upsilon}\beta}}\right|_{\vec{\tau}^{(0)}} (\tau_{\boldsymbol{\lambda}\alpha} - \tau^{(0)}_{\boldsymbol{\lambda}\alpha})(\tau_{\boldsymbol{\upsilon}\beta} - \tau^{(0)}_{\boldsymbol{\upsilon}\beta}) - \sum_{\boldsymbol{\lambda}\alpha}\frac{\hbar^2}{2M_\lambda}\frac{\partial^2}{\partial\tau^2_{\boldsymbol{\lambda}\alpha}}. \tag{3.66}$$

Electron-phonon interaction

The electron-phonon coupling is computed following the previous defined potential $v^{\mathrm{KS}}(\vec{r}_{\boldsymbol{a}},\{\vec{\tau}_{\boldsymbol{\lambda}}\})$. This term represents the vibronic operator in Eq. (3.27),

$$H_{\mathrm{ep}} = \sum_{\boldsymbol{a}}\sum_{\boldsymbol{\lambda}\alpha} \left.\frac{\partial v^{\mathrm{KS}}}{\partial\tau_{\boldsymbol{\lambda}\alpha}}\right|_{\vec{\tau}^{(0)}} (\tau_{\boldsymbol{\lambda}\alpha} - \tau^{(0)}_{\boldsymbol{\lambda}\alpha}) + \sum_{\boldsymbol{a}}\sum_{\boldsymbol{\upsilon}\beta}\sum_{\boldsymbol{\lambda}\alpha} \left.\frac{\partial^2 v^{\mathrm{KS}}}{\partial\tau_{\boldsymbol{\upsilon}\beta}\partial\tau_{\boldsymbol{\lambda}\alpha}}\right|_{\vec{\tau}^{(0)}} (\tau_{\boldsymbol{\lambda}\alpha} - \tau^{(0)}_{\boldsymbol{\lambda}\alpha})(\tau_{\boldsymbol{\upsilon}\beta} - \tau^{(0)}_{\boldsymbol{\upsilon}\beta}). \tag{3.67}$$

3.4.2 Force constants from first-principles

As noted in the introduction to this Section, the theoretical evaluation of vibronic coupling constants typically invokes the Hellmann–Feynman theorem. However, in practical first-principles simulations, this theorem is not strictly satisfied due to the finite and incomplete nature of the employed basis set. This limitation necessitates the development of alternative approaches to accurately compute vibronic coupling terms. To circumvent the limitations associated with the direct application of the Hellmann–Feynman theorem, the vibronic force constants are obtained without applying the Hellman-Feynman theorem. As a result, the derivative of the energy with respect to a mode Q is computed taking into account the Leibniz integral rule and the hermitian property of the Hamiltonian.

$$\frac{\partial E}{\partial Q} = \left\langle \Psi_i \middle| \frac{\partial H_\mathrm{e}}{\partial Q} \middle| \Psi_i \right\rangle + 2\mathrm{Re} \left\langle \Psi_i \middle| H_\mathrm{e} \middle| \frac{\partial \Psi_i}{\partial Q} \right\rangle. \tag{3.68}$$

Furthermore, if the wavefunction is a real-valued function (in conjunction with the Hermitian nature of the Hamiltonian),

$$\frac{\partial E}{\partial Q} = \left\langle \Psi_i \middle| \frac{\partial H_\mathrm{e}}{\partial Q} \middle| \Psi_i \right\rangle + 2 \left\langle \Psi_i \middle| H_\mathrm{e} \middle| \frac{\partial \Psi_i}{\partial Q} \right\rangle. \tag{3.69}$$

Now, by evaluating the preceding expression at the reference structure ($Q = 0$),

$$\left(\frac{\partial E_i}{\partial Q}\right)_{Q=0} = \left\langle \Psi_i^{(0)} \middle| \left(\frac{\partial H_\mathrm{e}}{\partial Q}\right)_{Q=0} \middle| \Psi_i^{(0)} \right\rangle + 2 \left\langle \Psi_i^{(0)} \middle| H_\mathrm{e}^{(0)} \middle| \left(\frac{\partial \Psi_i}{\partial Q}\right)_{Q=0} \right\rangle. \tag{3.70}$$

In the application of the Hellmann-Feynman theorem the second term in the summation vanishes as the wavefunction is assumed to be exact, thereby upholding the theorem. However, in practical applications, such as first-principles calculations, the non-exact nature of the wavefunction introduces an error that is not negligible. To correct this discrepancy, the second term, commonly referred to as Pulay forces [161], is incorporated.

Considering the first derivative of the energy in Eq. (3.47), we compute the second derivative. In the study of the second derivative, we build a Hessian matrix with off-diagonal elements where the derivative has been performed with respect two different modes, as example, Q_Γ and $Q_{\Gamma'}$ and diagonal term. Assuming the of modes, $\{Q\}$ has been build considering nondegenerate normal modes, the Hessian matrix is diagonal. As a result, in order to study the second derivative, the only terms that we have to take into account are the ones in the diagonal,

$$\frac{\partial^2 E_i}{\partial Q^2} = \left\langle \Psi_i \middle| \frac{\partial^2 H_\mathrm{e}}{\partial Q^2} \middle| \Psi_i \right\rangle + 4 \left\langle \Psi_i \middle| \frac{\partial H_\mathrm{e}}{\partial Q} \middle| \frac{\partial \Psi_i}{\partial Q} \right\rangle + 2 \left\langle \frac{\partial \Psi_i}{\partial Q} \middle| H_\mathrm{e} \middle| \frac{\partial \Psi_i}{\partial Q} \right\rangle + 2 \left\langle \Psi_i \middle| H_\mathrm{e} \middle| \frac{\partial^2 \Psi_i}{\partial Q^2} \right\rangle. \tag{3.71}$$

Since the energy E_i is a real number and an eigenstate of the Hamiltonian H_e,

$$\frac{\partial^2 E_i}{\partial Q^2} = \left\langle \Psi_i \middle| \frac{\partial^2 H_\mathrm{e}}{\partial Q^2} \middle| \Psi_i \right\rangle + 4 \left\langle \Psi_i \middle| \frac{\partial H_\mathrm{e}}{\partial Q} \middle| \frac{\partial \Psi_i}{\partial Q} \right\rangle + 2 \left\langle \frac{\partial \Psi_i}{\partial Q} \middle| H_\mathrm{e} \middle| \frac{\partial \Psi_i}{\partial Q} \right\rangle + 2E_i \left\langle \Psi_i \middle| \frac{\partial^2 \Psi_i}{\partial Q^2} \right\rangle. \tag{3.72}$$

Evaluating in the reference structure, $Q = 0$

$$\begin{aligned} \left(\frac{\partial^2 E_i}{\partial Q^2}\right)_0 = {} & \left\langle \Psi_i^{(0)} \middle| \left(\frac{\partial^2 H_\mathrm{e}}{\partial Q^2}\right)_0 \middle| \Psi_i^{(0)} \right\rangle + 4 \left\langle \Psi_i^{(0)} \middle| \left(\frac{\partial H_\mathrm{e}}{\partial Q}\right)_0 \middle| \left(\frac{\partial \Psi_i}{\partial Q}\right)_0 \right\rangle + \\ & + 2 \left\langle \left(\frac{\partial \Psi_i}{\partial Q}\right)_0 \middle| H_\mathrm{e}^{(0)} \middle| \left(\frac{\partial \Psi_i}{\partial Q}\right)_0 \right\rangle + 2E_i^{(0)} \left\langle \Psi_i^{(0)} \middle| \left(\frac{\partial^2 \Psi_i}{\partial Q^2}\right)_0 \right\rangle. \end{aligned} \tag{3.73}$$

3.4.3 2n+1 Theorem

To calculate $\left(\partial^2 E_i / \partial Q^2\right)_0$ as defined in Eq. (3.73), the first- and second-order corrections to the wavefunction are required. However, to simplify the computational complexity, the $2n+1$ theorem [52] can be applied. The theorem states that knowing the n-th order term of the wavefunction, it is possible to compute until the $2n+1$ order of the energy corrections. Specifically, in this case, we will demonstrate that, by knowing the first-order correction to the wavefunction ($n = 1$), it is possible to compute the second-order correction to the energy avoiding the calculation of $\partial^2\Psi_i/\partial Q^2$.

The starting point involves applying the normalization condition to the wavefunction obtained from perturbation theory, as given by Eq. (3.34). By enforcing the normalization constraint,

$$\begin{aligned}\langle\Psi_i|\Psi_i\rangle = 1 = &\left\langle\Psi_i^{(0)}|\Psi_i^{(0)}\right\rangle + \\ &+Q\left[\left\langle\Psi_i^{(0)}|\left(\frac{\partial\Psi_i}{\partial Q}\right)_0\right\rangle + \left\langle\left(\frac{\partial\Psi_i}{\partial Q}\right)_0|\Psi_i^{(0)}\right\rangle\right] + \\ &+Q^2\left[\frac{1}{2}\left\langle\left(\frac{\partial^2\Psi_i}{\partial Q^2}\right)_0|\Psi_i^{(0)}\right\rangle + \frac{1}{2}\left\langle\Psi_i^{(0)}|\left(\frac{\partial^2\Psi_i}{\partial Q^2}\right)_0\right\rangle + \left\langle\left(\frac{\partial\Psi_i}{\partial Q}\right)_0|\left(\frac{\partial\Psi_i}{\partial Q}\right)_0\right\rangle\right].\end{aligned} \tag{3.74}$$

The electronic wavefunctions in the reference geometry satisfy $\langle\Psi_i^{(0)}|\Psi_i^{(0)}\rangle = 1$. Since equality Eq. 3.74 is holds for any mode Q

$$\left\langle\Psi_i^{(0)}|\frac{\partial\Psi_i}{\partial Q_\Gamma}\right\rangle + \left\langle\left(\frac{\partial\Psi_i}{\partial Q}\right)_0|\Psi_i^{(0)}\right\rangle = 0, \tag{3.75}$$

$$\frac{1}{2}\left\langle\left(\frac{\partial^2\Psi_i}{\partial Q^2}\right)_0|\Psi_i^{(0)}\right\rangle + \frac{1}{2}\left\langle\Psi_i^{(0)}|\left(\frac{\partial^2\Psi_i}{\partial Q^2}\right)_0\right\rangle + \left\langle\left(\frac{\partial\Psi_i}{\partial Q}\right)_0|\left(\frac{\partial\Psi_i}{\partial Q}\right)_0\right\rangle = 0. \tag{3.76}$$

Considering the assumption that the wavefunction Ψ_i is real-valued, $\langle\Psi_i^{(2)}|\Psi_i^{(0)}\rangle = \langle\Psi_i^{(0)}|\Psi_i^{(2)}\rangle$

$$\left\langle\Psi_i^{(0)}|\left(\frac{\partial^2\Psi_i}{\partial Q^2}\right)_0\right\rangle = -\left\langle\left(\frac{\partial\Psi_i}{\partial Q}\right)_0|\left(\frac{\partial\Psi_i}{\partial Q}\right)_0\right\rangle. \tag{3.77}$$

The previous equality is substitute in Eq.(3.73)

$$\begin{aligned}\left(\frac{\partial^2E_i}{\partial Q^2}\right)_0 = &\left\langle\Psi_i^{(0)}\middle|\left(\frac{\partial^2H_\mathrm{e}}{\partial Q^2}\right)_0\middle|\Psi_i^{(0)}\right\rangle + 4\left\langle\Psi_i^{(0)}\middle|\left(\frac{\partial H_\mathrm{e}}{\partial Q}\right)_0\middle|\left(\frac{\partial\Psi_i}{\partial Q}\right)_0\right\rangle + \\ &+2\left\langle\left(\frac{\partial\Psi_i}{\partial Q}\right)_0\middle|H_\mathrm{e}^{(0)}\middle|\left(\frac{\partial\Psi_i}{\partial Q}\right)_0\right\rangle - 2E_i^{(0)}\left\langle\left(\frac{\partial\Psi_i}{\partial Q}\right)_0\middle|\left(\frac{\partial\Psi_i}{\partial Q}\right)_0\right\rangle.\end{aligned} \tag{3.78}$$

Analysis of the expression reveals that only the zeroth- and first-order corrections to the wavefunction appear. Consequently, knowledge of the first derivative of the wavefunction is sufficient to determine the second-order energy. This result exemplifies the so-called $2n+1$ theorem for the total energy, which also holds in the case of the one-particle eigenvalue problem. The theorem was rigorously demonstrated up to third order in energy in Ref. [52], and the Appendix provides a general proof for any arbitrary order n.

3.4.4 First-order correction to wavefunction

Once the second-order correction to the energy is expressed as a function of the first-order correction to the wavefunction, the next step involves evaluating this term explicitly. The value of $\partial\Psi/\partial Q$ could be obtained by means of the Sternheimer equation [162],

$$(H_\mathrm{e}^{(0)} - E_i^{(0)})\left(\frac{\partial\Psi_i}{\partial Q}\right)_0 = -\left(\frac{\partial H_\mathrm{e}}{\partial Q}\right)_0\Psi_i^{(0)}, \tag{3.79}$$

which is the result of solving the perturbation problem. In Eq. (3.79), the first-order energy does not appear since we assume the reference structure is a minimum of the APES, as we discussed before. In first-principles simulations all equations associated to to the electronic structure are solved within a basis. Let's see how the basis choice affects the solution of Eq. (3.79).

3.4.4.1 Adiabatic basis $\{\Psi_i^{(0)}\}$

The term $\partial\Psi_i/\partial Q$ can be expanded within a basis of the Hilbert space. In particular, the set of eigenfunctions of $H_\mathrm{e}^{(0)}$ represents a basis of the Hilbert space. The previous basis set is called *adiabatic* since the functions composing the basis set remain unchanged under variations in geometry. Expressing the function $\partial\Psi_i/\partial Q$ as a linear combination of basis set $\{\Psi_i^{(0)}\}$

$$\left|\left(\frac{\partial\Psi_i}{\partial Q}\right)_0\right\rangle = \sum_j c_{ij}\left|\Psi_j^{(0)}\right\rangle. \tag{3.80}$$

The zeroth order basis is divided into two subsets. Considering the state $\Psi_i^{(0)}$, the index j associated to state $\Psi_j^{(0)}$ belongs to I, $j \in I$, if $\Psi_i^{(0)}$ and $\Psi_j^{(0)}$ are eigenstates associated to degenerate states, $E_j^{(0)} = E_i^{(0)}$. Conversely, the index j belongs to the orthonormal complement $j \in I^\perp$, in other words, if $E_j^{(0)} \neq E_i^{(0)}$.

$$\left|\left(\frac{\partial\Psi_i}{\partial Q}\right)_0\right\rangle = \sum_j c_{ij}\left|\Psi_j^{(0)}\right\rangle = \sum_{j\in I} c_{ij}\left|\Psi_j^{(0)}\right\rangle + \sum_{j\in I^\perp} c_{ij}\left|\Psi_j^{(0)}\right\rangle. \tag{3.81}$$

The left part of the Sternheimer equation can thus be rewritten as

$$\begin{aligned}\left(H_\mathrm{e}^{(0)} - E_i^{(0)}\right)\left|\left(\frac{\partial\Psi_i}{\partial Q}\right)_0\right\rangle &= \sum_{j\in I} c_{ij}\left(H_\mathrm{e}^{(0)} - E_i^{(0)}\right)\left|\Psi_j^{(0)}\right\rangle + \sum_{j\in I^\perp} c_{ij}\left(H_\mathrm{e}^{(0)} - E_i^{(0)}\right)\left|\Psi_j^{(0)}\right\rangle = \\ &= \sum_{j\in I^\perp} c_{ij}\left(E_j^{(0)} - E_i^{(0)}\right)\left|\Psi_j^{(0)}\right\rangle.\end{aligned} \tag{3.82}$$

The Sternheimer equation yields the following expression upon substituting the left-hand side of the equation with the previously derived expression,

$$\sum_{j\in I^\perp} c_{ij}^{(1)}\left(E_j^{(0)} - E_i^{(0)}\right)\left|\Psi_j^{(0)}\right\rangle = -\left(\frac{\partial H_\mathrm{e}}{\partial Q}\right)_0\left|\Psi_i^{(0)}\right\rangle. \tag{3.83}$$

Premultiplying by $\langle\Psi_k^{(0)}|$ with $k \in I^\perp$ and applying the orthonormality of the basis set, $\langle\Psi_k^{(0)} \mid \Psi_j^{(0)}\rangle = \delta_{kj}$

$$c_{ij}^{(1)} = \frac{1}{E_i^{(0)} - E_j^{(0)}}\left\langle\Psi_j^{(0)}\right|\left(\frac{\partial H_\mathrm{e}}{\partial Q}\right)_0\left|\Psi_i^{(0)}\right\rangle \qquad \text{for } j \in I^\perp. \tag{3.84}$$

To compute c_{ij} for $j \in I$, premultiplication by $\langle\Psi_k^{(0)}|$ with $k \in I$ causes the left-hand side of Eq. (3.83) to vanish, as $\langle\Psi_k^{(0)} \mid \Psi_j^{(0)}\rangle = 0$. In fact, the presence of gauge freedom permits c_{ij} for $j \in I$ to be set to zero. As a result, the so-called sum-over-states expression can be expressed as

$$\left|\left(\frac{\partial\Psi_i}{\partial Q}\right)_0\right\rangle = \sum_{j\in I^\perp}\frac{\left\langle\Psi_j^{(0)}\right|(\partial H_\mathrm{e}/\partial Q)_0\left|\Psi_i^{(0)}\right\rangle}{E_i^{(0)} - E_j^{(0)}}\left|\Psi_j^{(0)}\right\rangle, \tag{3.85}$$

which requires knowledge of all the zeroth order wavefunctions and energies. The resulting expression is equivalent to that derived from the Hellmann–Feynman theorem, as shown in Eq. 3.51, and also corresponds to the expression obtained within the framework of the pseudo-Jahn–Teller theory [73].

For an exact result, the summation over the excited states extends to infinity. In practice, it is required to truncate this sum and consider only a finite subset of excited states. However, depending on the nature of the basis set, this approach may yield highly inaccurate results. This inaccuracy arises because a large number of excited states is typically required to adequately capture the density changes.

In particular, localized adiabatic basis functions yield highly inaccurate results due to the limited number of elements in the basis set in practice. For instance, if the basis set is constructed at $Q = 0$, it may fail to describe the system accurately at $Q \neq 0$ because it lacks the necessary size and flexibility to represent the system away from the reference geometry. However, if the basis set were large enough, the localized adiabatic basis set would, in principle, be able to reproduce the results, but this calculation is not computationally efficient computationally. Conversely, plane-wave basis functions, which also constitute an adiabatic basis, perform well when the geometry is perturbed. This accuracy arises from the extremely large number of basis functions in the plane wave set, which collectively allow for an effective representation of the system under geometric deviations from the reference configuration.

In localized basis sets this limitation can be overcome employing a basis set that varies with position commonly referred to as the *diabatic basis*. In practice, due to the reduced number of functions which composed a localized basis, the diabatic basis is much more exact than the adiabatic one.

3.4.4.2 Diabatic basis $\{\chi_a\}$

A basis set is referred to as diabatic when its constituent functions vary with the positions of the atoms. An example of a diabatic basis is a floating basis set, in which the orbitals dynamically follow the motion of the nuclei. It must be emphasized that the term "diabatic" refers solely to the basis set changing with the atomic positions, while the framework remains within the adiabatic scheme described at the beginning of the chapter.

Employing the previous basis set, the wavefunction is expanded employing a basis set $\{\chi_a\}$ which depends on the atomic displacements associated to Q,

$$\Psi_i^{(0)} = \sum_a \tilde{c}_{ia} \chi_a. \tag{3.86}$$

We compute the derivative of the wavefunction with respect to the Q coordinates

$$\frac{\partial \Psi_i}{\partial Q} = \sum_a \left[\frac{\partial \tilde{c}_{ia}}{\partial Q} \chi_a + c_{ia} \frac{\partial \chi_a}{\partial Q} \right]. \tag{3.87}$$

Substituting in the Sternheimer equation by the previous expansion,

$$\sum_a \left(\frac{\partial \tilde{c}_{ia}}{\partial Q} \right)_0 (H_\mathrm{e}^{(0)} - E_i^{(0)}) \left| \chi_a \right\rangle + \sum_a \tilde{c}_{ia} (H_\mathrm{e}^{(0)} - E_i^{(0)}) \left| \left(\frac{\partial \chi_a}{\partial Q} \right)_0 \right\rangle = - \sum_a \tilde{c}_{ia} \left(\frac{\partial H_\mathrm{e}}{\partial Q} \right)_0 \left| \chi_a \right\rangle. \tag{3.88}$$

Now, the equation is expressed in terms of matrix elements applying the bra $\langle \chi_b |$,

$$\begin{aligned}
&\sum_a \left[H_{ab}^{(0)} - E_i^{(0)} S_{ab} \right] \left(\frac{\partial \tilde{c}_{ia}}{\partial Q} \right)_0 = \\
&= - \left[\sum_a \left\langle \chi_b \middle| \left(\frac{\partial H_{\mathrm{e}}}{\partial Q} \right)_0 \middle| \chi_a \right\rangle - \left\langle \chi_b \middle| H_{\mathrm{e}}^{(0)} \middle| \frac{\partial \chi_a}{\partial Q} \right\rangle - E_i^{(0)} \left\langle \chi_b \middle| \frac{\partial \chi_a}{\partial Q} \right\rangle \right] \tilde{c}_{ia}.
\end{aligned} \tag{3.89}$$

The derivatives $(\partial \tilde{c}_{ia} / \partial Q)_0$ which measure the changes of the coefficients $\{\tilde{c}_{ia}\}$ with respect the atomic distortions are obtained solving the previous system of linear equations. Considering the expression presented in Eq. (3.89), there are two terms which depend on $\partial \chi_a / \partial Q$. These terms represent the Pulay corrections.

There exist several examples of diabatic bases. In this context, some examples are presented studying the obtained expression in Eq.(3.89)

- Basis is not nuclear coordinate dependent (plane waves):
  This is the simplest case and the easier to interpret

$$\sum_a \left(H_{ab} - E_i^{(0)} \right) c_{ia}^{(1)} = - \sum_a \left[\left\langle \chi_b \middle| \left(\frac{\partial H_{\mathrm{e}}}{\partial Q} \right)_{Q=0} \middle| \chi_a \right\rangle \right] c_{ia}. \tag{3.90}$$

- Basis displaces rigidly (SIESTA orbitals):
  The orbitals have the form $\chi_a \equiv \chi_a(\vec{r} - \vec{\tau}(Q))$, resulting

$$\frac{\partial \chi_a}{\partial Q_\alpha} = \frac{\partial \chi_a}{\partial \tau_\alpha} \frac{\partial \tau_\alpha}{\partial Q_\alpha}, \tag{3.91}$$

  where α represents a cartesian coordinate. As a consequence, we obtain

$$\begin{aligned}
\sum_a \left(H_{ab}^{(0)} - E_i^{(0)} \right) \left(\frac{\partial \tilde{c}_{ia}}{\partial Q_\alpha} \right)_0 = - \sum_a \Bigg[\left\langle \chi_b \middle| \left(\frac{\partial H_{\mathrm{e}}}{\partial Q_\alpha} \right)_0 \middle| \chi_a \right\rangle - \\
- \langle \chi_b | \left(H_{\mathrm{e}}^{(0)} - E_i^{(0)} \right) \left| \left(\frac{\partial \chi_a}{\partial \tau_\alpha} \right)_0 \right\rangle \left(\frac{\partial \tau_\partial}{\partial Q_\alpha} \right)_0 \Bigg] \tilde{c}_{ia}.
\end{aligned} \tag{3.92}$$

- Basis is orthogonal at all geometries (Wannier orbitals):
  Since the Wannier functions are orthonormal, we can derive the orthonormalization condition, $\langle \boldsymbol{\chi_b} | \boldsymbol{\chi_a} \rangle = \boldsymbol{\delta_{ab}}$, to obtain

$$\left\langle \left(\frac{\partial \chi_b}{\partial Q} \right)_0 \middle| \chi_a \right\rangle + \langle \chi_b \left| \left(\frac{\partial \chi_a}{\partial Q} \right)_0 \right\rangle = 0. \tag{3.93}$$

  As we are working with real functions, the following relationship is satisfied

$$\left\langle \left(\frac{\partial \chi_b}{\partial Q} \right)_0 \middle| \chi_a \right\rangle = 0, \tag{3.94}$$

$$\sum_a \left(H_{ab}^{(0)} - E_i^{(0)} \right) \left(\frac{\partial c_{ia}}{\partial Q} \right)_0 = \sum_a \left[\left\langle \chi_b \middle| \left(\frac{\partial H_{\mathrm{e}}}{\partial Q} \right)_0 \middle| \chi_a \right\rangle + \langle \chi_b | H_{\mathrm{e}}^{(0)} \left| \left(\frac{\partial \chi_a}{\partial Q} \right)_0 \right\rangle \right] c_{ia}. \tag{3.95}$$

Considering the systems of equations obtained for each case, it can be observed that in both diabatic and adiabatic scenarios, the results converge to those obtained when a complete basis set is employed. On the one hand, when using an adiabatic basis, the equation derived for the wavefunction change converges to the exact result, provided that the infinite set of excited states is considered. On the other hand, in the case of a diabatic basis, if an exact scenario is assumed where Pulay forces are not relevant, the resulting equation corresponds to that obtained by the exact scheme.

At first glance, it may seem contradictory that, in the complete basis set limit, an adiabatic basis would yield the same results as a diabatic basis. However, in the exact case, both basis sets lead to the same results as they can be related through a basis transformation, as demonstrated in the study presented in Ref. [159].

3.4.5 Variational energy

The previous discussion focused on computing force constants in a general framework [Eq. (3.78)], which naturally leads to evaluating the variation of the wavefunction induced by nuclear distortions (Sec. 3.4.4). This last step is performed by applying the Sternheimer equation, Eq. (3.79).

A major advantage of the Sternheimer equation is that it avoids the explicit summation over excited states—an operation that is computationally demanding in large systems. Instead, the method directly yields the first-order correction to the electronic wavefunction, $\partial\Psi_i/\partial Q$, which can subsequently be used to evaluate various response properties, such as dipole moments, polarizabilities, and electric-field gradients. The Sternheimer formulation is often introduced directly through perturbation theory, although it can also be derived variationally from the second-order energy functional, as we will present in this section.

The variational principle states that the total energy of a system is minimized with respect to the wavefunction of the system in its ground state. When a perturbation is applied, the variational principle still holds in the sense that the second-order correction to the energy is stationary with respect to the first-order correction to the wavefunction. For example, in Hartree–Fock theory, the ground-state energy is obtained by minimizing the energy functional with respect to the molecular orbitals. Under a perturbation, the problem is addressed using Coupled Perturbation Hartree–Fock (CPHF), where the second-order energy correction remains variational.

The second-order energy correction represents the change in energy resulting from an external perturbation which distorts the lattice inducing a redistribution of the electron density. As expressed in Eq. (3.78), the second-order energy correction is a functional which explicitly depends on $\Psi_i^{(0)}$ and $(\partial\Psi_i/\partial Q)_0$, where Q is the perturbation parameter. This functional captures the system's energy response to the perturbation and provides a framework for understanding how energy evolves under external influences.

At this point, we are going to show that the correction to the energy is variational with respect $\Psi_i^{(1)} \equiv \partial\Psi_i/\partial Q$. In other words, we are going to prove the following statement:

$$\begin{aligned} E^{(2)} \equiv \frac{\partial^2 E_i}{\partial Q^2} = \min_{\partial\Psi_i/\partial Q} \Bigg\{ & \left\langle \Psi_i \middle| \frac{\partial^2 H_\mathrm{e}}{\partial Q^2} \middle| \Psi_i \right\rangle + 4 \left\langle \Psi_i \middle| \frac{\partial H_\mathrm{e}}{\partial Q} \middle| \frac{\partial \Psi_i}{\partial Q} \right\rangle + \\ & + 2 \left\langle \frac{\partial \Psi_i}{\partial Q} \middle| H_\mathrm{e} \middle| \frac{\partial \Psi_i}{\partial Q} \right\rangle - 2E_i \left\langle \frac{\partial \Psi_i}{\partial Q} \middle| \frac{\partial \Psi_i}{\partial Q} \right\rangle \Bigg\}. \end{aligned} \tag{3.96}$$

If the previous equality is satisfied, then $\delta E_i^{(2)}/\delta\Psi_i^{(1)} = 0$ where the quantity $\delta\Psi_i^{(1)}$ represents an infinitesimal variation in $\Psi_i^{(1)}$. The second order energy $E_i^{(2)}$ changes under this variation as

$$\begin{aligned} \delta\left(\frac{\partial^2 E_i}{\partial Q^2}\right) = & 4 \left\langle \Psi_i \middle| \frac{\partial H_\mathrm{e}}{\partial Q} \middle| \delta\left(\frac{\partial \Psi_i}{\partial Q}\right) \right\rangle + 2 \left\langle \delta\left(\frac{\partial \Psi_i}{\partial Q}\right) \middle| H_\mathrm{e} \middle| \frac{\partial \Psi_i}{\partial Q} \right\rangle + 2 \left\langle \frac{\partial \Psi_i}{\partial Q} \middle| H_\mathrm{e} \middle| \delta\left(\frac{\partial \Psi_i}{\partial Q}\right) \right\rangle - \\ & - 2E_i \left\langle \delta\left(\frac{\partial \Psi_i}{\partial Q}\right) \middle| \frac{\partial \Psi_i}{\partial Q} \right\rangle - 2E_i \left\langle \frac{\partial \Psi_i}{\partial Q} \middle| \delta\left(\frac{\partial \Psi_i}{\partial Q}\right) \right\rangle = 0. \end{aligned} \tag{3.97}$$

Knowing the hamiltonian is hermitian and grouping terms

$$\begin{aligned} \delta\left(\frac{\partial^2 E_i}{\partial Q^2}\right)_0 = & 4 \left\langle \delta\left(\frac{\partial \Psi_i}{\partial Q}\right)_0 \middle| \left(\frac{\partial H_\mathrm{e}}{\partial Q}\right)_0 \middle| \Psi_i^{(0)} \right\rangle + 4 \left\langle \delta\left(\frac{\partial \Psi_i}{\partial Q}\right)_0 \middle| H_\mathrm{e}^{(0)} \middle| \left(\frac{\partial \Psi_i}{\partial Q}\right)_0 \right\rangle - \\ & - 4E_i \left\langle \delta\left(\frac{\partial \Psi_i}{\partial Q}\right)_0 \middle| \left(\frac{\partial \Psi_i}{\partial Q}\right)_0 \right\rangle = \\ = & 4 \left\langle \delta\left(\frac{\partial \Psi_i}{\partial Q}\right)_0 \middle| \left\{ \left(\frac{\partial H_\mathrm{e}}{\partial Q}\right)_0 \middle| \Psi_i^{(0)} \right\rangle + H_\mathrm{e}^{(0)} \left| \left(\frac{\partial \Psi_i}{\partial Q}\right)_0 \right\rangle - E_i \left| \left(\frac{\partial \Psi_i}{\partial Q}\right)_0 \right\rangle \right\} = 0. \end{aligned} \tag{3.98}$$

As a result, the expression within braces corresponds to the Sternheimer equation when the variation is null, Eq. (3.79), which is satisfied. Consequently, the second-order term of the energy is variational with respect to $\Psi_i^{(1)}$. This implies that the energy can be minimized by optimizing the second-order term with respect to $\Psi_i^{(1)}$, while maintaining the normalization of the total wavefunction. Furthermore, this property provides an alternative approach for determining the first-order correction to the wavefunction, as it reduces the problem to finding the value of $\Psi^{(1)}$ that minimizes $E_i^{(2)}$.

The variational nature of $E_i^{(2)}$ ensures that it is stationary when $\Psi_i^{(1)}$ is properly optimized to describe the system's response to the perturbation. This guarantees that the wavefunction response accurately reflects both the redistribution of electron density and the contributions arising from atomic displacements.

Given the established connection between the Sternheimer equation and the variational principle, the Coupled Perturbed Hartree–Fock (CPHF) method can be understood as the explicit realization of the Sternheimer formalism within the framework of Hartree–Fock theory. Conceptually, CPHF is equivalent to the Sternheimer approach. Both methods solve a linear response problem to determine how the wavefunction (or orbitals) respond to a perturbation, and both avoid the need for summing over unoccupied states.

These methods differ only in notation and the conventions of their respective fields. The Sternheimer approach, rooted in the physics literature—particularly in the study of periodic systems using plane-wave basis sets—is implemented in electronic structure codes such as ABINIT[163] and QUANTUM EXPRESSO[164]. In contrast, the CPHF method is the standard formulation in quantum chemistry and is widely implemented in software packages like Gaussian [165], ORCA [166], and Psi4 [167].

3.4.6 Effect of the reciprocal space: response to incommesurate perturbations of periodic systems

Thus far, vibronic coupling, which describes the interaction between electronic and vibrational states, has been primarily described for molecular systems. However, this concept can also be extended to periodic systems, such as crystalline solids, where vibronic coupling manifests as the interaction between electrons and phonons (lattice vibrations).

In the context of periodic systems, the Bloch theorem provides a fundamental framework for studying and modeling the behavior of electrons under a periodic external potential. In the absence of perturbations, the ground-state external potential operator exhibits translational periodicity in the direct lattice, such that for any lattice vector $\vec{R}$,

$$v_{\text{ext}}^{(0)}(\vec{r}+\vec{R},\vec{r}'+\vec{R}) = v_{\text{ext}}^{(0)}(\vec{r},\vec{r}'), \tag{3.99}$$

where $\vec{r}$ and $\vec{r}'$ represent the electronic positions. According to the Bloch theorem, the wavefunction $\psi_{n\vec{k}}^{(0)}(\vec{r})$ can be expressed as the product of a plane wave characterized by a vector $\vec{k}$ of the reciprocal lattice and a periodic function labeled by the band number, n and the $\vec{k}$-vector, $u_{n\vec{k}}^{(0)}$, with the same periodicity as the external potential

$$\psi_{n\vec{k}}^{(0)}(\vec{r}) = e^{i\vec{k}\vec{r}} u_{n\vec{k}}^{(0)}(\vec{r}). \tag{3.100}$$

Now the effects induced by a perturbation over the periodic external potential $v_{\text{ext}}^{(0)}$ are studied. Specifically, we analyze the response of a periodic system to a perturbation associated with a phonon characterized by an incommensurate wave vector $\vec{q}$ and branch index ν, as such a perturbation breaks the translational symmetry of the system. This perturbation gives rise to a structural distortion, which can be described in terms of local vibrational modes in real space, denoted by $Q_\nu(\vec{R})$, where $\vec{R}$ labels the lattice vectors. The corresponding mode in reciprocal space is obtained via a Fourier transform,

$$Q_\nu(\vec{q}) = \sum_{\vec{R}} Q_\nu(\vec{R})\, e^{i\vec{q}\cdot\vec{R}}. \tag{3.101}$$

The introduction of an incommensurate wave vector results in a non-local perturbation, in other words, the perturbation acts as a linear operator rather than a simple local function in real space. Furthermore, the intrinsic periodicity of the lattice imposes constraints on the perturbing potential operator, ensuring that it satisfies the following condition,

$$v_{\text{ext},\vec{q}\nu}^{(1)}(\vec{r}+\vec{R},\vec{r}'+\vec{R}) = e^{i\vec{q}\vec{R}} v_{\text{ext},\vec{q}\nu}^{(1)}(\vec{r},\vec{r}'). \tag{3.102}$$

This perturbation reflects a Bloch-like modulation characterized by the wave vector $\vec{q}$ and the branch index ν. As consequence, the first order term of the hamiltonian,

$$h^{(1)} = e^{i\vec{q}\vec{r}} v_{\text{ext,P},\vec{q}\nu}^{(1)}(\vec{r},\vec{r}'), \tag{3.103}$$

which represents a linear perturbation, can be expressed as the product of two elements: a plane wave factor that captures the modulation as a consequence to the perturbation and a periodic part, represented by the subscript “P”, which keeps the periodicity of the lattice satisfying then

$$v^{(1)}_{\text{ext,P},\vec{q}\nu}(\vec{r}+\vec{R},\vec{r}'+\vec{R}) = v^{(1)}_{\text{ext,P},\vec{q}\nu}(\vec{r},\vec{r}'). \tag{3.104}$$

Analyzing the perturbing hamiltonian in Eq. (3.103) it is complex-valued. To ensure that the Hamiltonian remains real (i.e. Hermitian) regardless of the specific value of $\vec{q}$, it must incorporate the Hermitian conjugate counterpart of the perturbing potential,

$$h^{(1)} = e^{i\vec{q}\vec{r}} v^{(1)}_{\text{ext,P},\vec{q}\nu}(\vec{r},\vec{r}') + e^{-i\vec{q}\vec{r}} v^{(1)}_{\text{ext,P},-\vec{q}\nu}(\vec{r},\vec{r}'), \tag{3.105}$$

where $v^{(1)}_{\text{ext,P},\vec{q}\nu}$ and $v^{(1)}_{\text{ext,P},-\vec{q}\nu}$ are mutually complex conjugate. The resulting first order corrections to the wavefunction is obtained applying the Sternheimer equation, Eq. (3.79),

$$\left(H_0 - E_{n\vec{k}}\right) \psi^{(1)}_{n\vec{k}}(\vec{r}) = -h^{(1)} \psi^{(0)}_{n\vec{k}}(\vec{r}). \tag{3.106}$$

Expanding in Bloch states, the first-order correction takes the form

$$\psi^{(1)}_{n\vec{k}}(\vec{r}) = \sum_{m,\vec{k}'} C_{m\vec{k}',n\vec{k}} \psi^{(0)}_{m\vec{k}'}(\vec{r}). \tag{3.107}$$

Projection onto Bloch states is performed. In addition, perturbation matrix elements are considered

$$C_{m\vec{k}',n\vec{k}} = \frac{\langle \psi^{(0)}_{m\vec{k}'} | h^{(1)} | \psi^{(0)}_{n\vec{k}} \rangle}{E_{n\vec{k}} - E_{m\vec{k}'}}. \tag{3.108}$$

Matrix elements $\langle \psi^{(0)}_{m\vec{k}'} | h^{(1)} | \psi^{(0)}_{n\vec{k}} \rangle$ are nonzero only when the momentum is conserved. As a result, perturbation $h^{(1)}$ characterized by $\vec{q}$ couples states satisfying $-\vec{k}' + \vec{k} + \vec{q} = 0$, and then $\vec{k}' = \vec{k} + \vec{q}$. The perturbation shifts the wavevector of the original state from $\vec{k}$ to $\vec{k}+\vec{q}$. Finally, the obtained first order correction for the wavefunction is

$$\psi^{(1)}_{n\vec{k}}(\vec{r}) = \sum_{m} \frac{\langle \psi^{(0)}_{m,\vec{k}+\vec{q}} | h^{(1)} | \psi^{(0)}_{n,\vec{k}} \rangle}{E_{n\vec{k}} - E_{m,\vec{k}+\vec{q}}} \psi^{(0)}_{m,\vec{k}+\vec{q}}(\vec{r}). \tag{3.109}$$

It follows the Bloch shape and verifies

$$\psi^{(1)}_{n\vec{k}}(\vec{r}+\vec{R}) = e^{i(\vec{k}+\vec{q})\vec{R}} \psi^{(1)}_{n\vec{k}}(\vec{r}). \tag{3.110}$$

This study of a non-periodic perturbation in a periodic system is connected to electron-phonon [74] theory through the equivalence between the Eq. (3.105) and Eq. (33) in Ref. [74]. From this point, it is possible to determine the electron-lattice linear coupling constant shown in Eq. (38) of Ref. [74]

$$f_{mn\nu}(\vec{k},\vec{q}) = \left\langle u_{m\,\vec{k}+\vec{q}} \middle| v^{(1)}_{\text{ext,P},\vec{q}}(\vec{r},\vec{r}') \middle| u_{n\vec{k}} \right\rangle_{\text{uc}}, \tag{3.111}$$

where the subscript “uc” denotes integration over the unit cell. This term is related to linear orbital vibronic constant in Eq. (3.39) but expressed in reciprocal space. Similarly, the second order term is computed following

$$g_{mn\nu\nu'}(\vec{k},\vec{q},\vec{q}') = \frac{1}{2} \left\langle u_{m\vec{k}+\vec{q}+\vec{q}'} \middle| v^{(2)}_{\text{ext,P},\vec{q},\vec{q}'} \middle| u_{n\vec{k}} \right\rangle_{\text{uc}}. \tag{3.112}$$

The second order term in reciprocal space is related to Eq. (40) attending to Ref. [74].

3.4.7 Finite-differences

The study of electron-phonon coupling can be performed employing finite differences [76], and alternatively, density functional perturbation theory (DFPT) [52; 53; 54].

The main advantages of finite differences methods are the following ones:

- The finite difference method facilitates the study of electron-phonon coupling without relying on complex theoretical frameworks. The approach solely involves finite difference formulas, making it more straightforward in implementation. In comparison to DFPT [74], this method does not necessitate an elaborate theoretical basis while achieving equivalent precision.
- The finite differences procedure is both versatile and transferable. This method can be seamlessly applied across various computational approaches without the requirement of a new implementation. Specifically, within the framework of Density Functional Theory (DFT), the finite differences approach can be employed effectively regardless of the chosen functional. This versatility allows for its application to systems exhibiting either weak or strong electronic correlation, considering the proper methods to address these interactions effectively. As an example of this versatility, the finite differences technique has been successfully utilized to investigate systems containing the bismuthate anion transition-metal chloronitrides complexes, enabling the study of vibronic coupling and going beyond the LDA and GGA approximations in order to take into account the dominant role of electron-electron correlations, which are fundamental to understanding the properties of these strongly correlated systems [168].
- Finite-difference methods are able to incorporate electron-phonon coupling terms extending beyond the limitations of lowest-order perturbation theory (linear terms). This approach enables the capture of non-linear effects in the electron–phonon interaction. An example where the treatment of multiphonon interactions is critical is the investigation of temperature-dependent absorption phenomena in hybrid perovskites [169].

However, the finite differences procedure is computationally demanding presenting certain challenges. The primary issues involve establishing a balance between time and length scales.

- Length Scale: The finite difference method requires multiple first-principles calculations involving the explicit displacement of atoms. For simulations of modes characterized by long wavelengths, it is necessary to employ large supercells in which the desired mode can be properly commensurate.
- Time Scale: The computation of finite differences necessitates a large number of calculations to account for all possible nuclear displacements, each characterized by the three Cartesian coordinates. For instance, in the case of the STO system using a $2 \times 2 \times 2$ supercell to include the Γ, X, M, and R modes, the number of required calculations scales with several factors: the number of atoms in the simulation cell, the directions of displacement, the sense of the displacement (positive or negative). The number of calculation scales when we are interest in quadratic constants.

To overcome this computational problem a proposed solution entails the incorporation of symmetry. Finite differences calculations involve several calculations that are related through the symmetry operations of the space group associated with the system under investigation. By imposing symmetry constraints, it is possible to significantly reduce the number of first-principles calculations required to carry out the finite-difference procedure.

3.4.8 Electron-phonon formalism in a Wannier functions basis set

In Sec. 3.4.6 electron-phonon interaction has been introduced to tackle solid systems exhibiting periodicity. In crystalline materials, electrons are commonly represented by Bloch waves, as given in Eq. (2.3). However, as discussed in Sec. 1.4, electrons can also be described by localized orbitals, in particular by Wannier functions. As mentioned above, these functions are obtained by a Fourier transform of the Bloch functions and span the same Hilbert space. Consequently, the electron-phonon matrix elements can be expressed in terms of Wannier functions, wherein the relationship between the electron-phonon coupling terms in the Bloch and Wannier representations is established through the Fourier transform that connects them. In this way, the first order term expressed in the Wannier functions representation, following Ref.[74], is

$$f_{mn\lambda\alpha}\left(\vec{R}_\Lambda,\vec{R}_A\right) = \left\langle \chi_{m\vec{0}}(\vec{r}) \left| \left(\frac{\partial V^{\mathrm{KS}}}{\partial \tau_{\lambda\alpha}}\right)_{\vec{r}-\vec{R}_\Lambda} \right| \chi_{n\vec{0}}\left(\vec{r}-\vec{R}_A\right) \right\rangle_{\mathrm{sc}}, \tag{3.113}$$

where $\partial V^{\mathrm{KS}}/\partial \tau_{\lambda\alpha}$ denotes the first order correction to the Kohn-Sham effective potential with respect to the nuclear displacements $\Delta\vec{\tau}_\lambda$ and the subscript “sc” indicated integration over the Born-von Kármán supercell. As defined in Sec. 1.2.1, the Greek index refers to atomic indices, while Latin indices are related to the electrons. Then, $\vec{R}_\Lambda$ represents the lattice vector associated to the atom and $\vec{R}_A$ is the lattice vector related to the Wannier function. The α component determines the Cartesian coordinate. The term expressed in Eq. (3.113) is related to electron-phonon interaction matrix element $f_{mn\nu}(\vec{k},\vec{q})$ by substituting the periodic part of the Bloch functions with its representation in terms of Wannier functions [equation related to Eq. (2.3)],

$$u_{n\vec{k}}(\vec{r}) = \sum_m \sum_{\vec{R}} e^{-i\vec{k}\cdot(\vec{r}-\vec{R})} U^{\dagger}_{mn\vec{k}} \chi_{m\vec{R}}(\vec{r}). \tag{3.114}$$

The first order correction to the KS potential expressed in reciprocal space, Eq. (3.111), can be rewritten in terms of Wannier functions expression as

$$\begin{aligned} f_{mn\nu}(\vec{k},\vec{q}) = &\sum_\Lambda \sum_A e^{i\left(\vec{k}\cdot\vec{R}_A+\vec{q}\cdot\vec{R}_\Lambda\right)} \\ &\times \sum_{m'n'\lambda\alpha} U_{mm'\vec{k}+\vec{q}}\, f_{m'n'\lambda\alpha}\left(\vec{R}_A,\vec{R}_\Lambda\right) U^{\dagger}_{n'n\vec{k}} \mathrm{u}_{\lambda\alpha,\vec{q}\nu}, \end{aligned} \tag{3.115}$$

where $\mathrm{u}_{\lambda\alpha,\vec{q}\nu} = \left(\hbar/2M_\lambda\omega_{\vec{q}\nu}\right)^{1/2} e_{\lambda\alpha,\nu}(\vec{q})$ and $e_{\kappa\alpha,\nu}(\vec{q})$ are the vibrational eigenmodes. The employed Born-von Kármán supercell for electronic band structure, $N_{\vec{R}_A}$ and phonon dispersions ($N_{\vec{R}_\Lambda}$) may be different. The inverse relation is

$$
\begin{aligned}
f_{mn\lambda\alpha}(\vec{R}_A,\vec{R}_\Lambda) =& \frac{1}{N_{\vec{R}_A} N_{\vec{R}_\Lambda}} \sum_A \sum_\Lambda e^{-i\left(\vec{k}\cdot\vec{R}_A+\vec{q}\cdot\vec{R}_\Lambda\right)} \\
&\times \sum_{m'n'\lambda\alpha} \mathrm{u}^{-1}_{\lambda\alpha,\vec{q}\nu} U^\dagger_{mm'\vec{k}+\vec{q}} \, f_{m'n'\nu}(\vec{k},\vec{q}) \, U^\dagger_{n'n\vec{k}} \, .
\end{aligned}
\tag{3.116}
$$

The elements $f_{mn\lambda\alpha}(\vec{R}_\Lambda,\vec{R}_A)$ decay rapidly with increasing $\vec{R}_\Lambda$ and $\vec{R}_A$ due to the localized nature of Wannier functions. Consequently, when computing the elements $f_{mn\nu}(\vec{k},\vec{q})$ from the real-space matrix elements $f_{mn\lambda\alpha}(\vec{R}_\Lambda,\vec{R}_A)$, only a limited number of terms are necessary to accurately reconstruct $f_{mn\nu}(\vec{k},\vec{q})$ across the Brillouin zone. By analyzing the decay behavior of the electron-lattice matrix elements $f_{m'n'\lambda\alpha}(\vec{R}_A,\vec{R}_\Lambda)$, it is observed that they decay with respect to $\vec{R}_A$ and $\vec{R}_\lambda$ at least as rapidly as the Wannier functions. In the special case of $\vec{R}_A=\vec{0}$, the decay with $\vec{R}_\Lambda$ is governed by the decay of the screened electric potential produced by atomic displacement $\Delta\tau_{\lambda\alpha}$. Specifically, in nonpolar semiconductors and insulators the decay follows a quadrupolar pattern, scaling as $|\vec{R}_\Lambda|^{-3}$, due to the analytical properties of the dielectric matrix [170]. Regarding metals, the decay rate follows $|\vec{R}_\Lambda|^{-4}$ scaling because of Friedel oscillations due to Fermi-surface nesting [171]. The rapid decay of the linear coupling constants $f_{mn\lambda\alpha}(\vec{R}_\Lambda,\vec{R}_A)$ with distances $\vec{R}_\Lambda$ and $\vec{R}_A$ plays a crucial role in ensuring the efficiency and feasibility of second-principles calculations.

Because of the localized nature of MLWFs, a basis set composed by these functions is particularly well-suited for the Slater-Koster interpolation of different magnitudes, as demonstrated in Ref. [69]. In particular, the calculation of $f_{mn\lambda\alpha}(\vec{R}_\Lambda,\vec{R}_A)$ from first-principles can be employed to determine the first order terms in reciprocal space $f_{mn\nu}(\vec{k},\vec{q})$ by considering the interpolation technique. Furthermore, the rapid decay of $f_{mn\lambda\alpha}(\vec{R}_\Lambda,\vec{R}_A)$ with respect to lattice vectors, as discussed earlier, is a crucial factor that makes this interpolation technique particularly efficient and practical for such calculations. The first step involves obtaining the matrix elements $f_{mn\nu}(\vec{k},\vec{q})$ using DFPT, with a coarse grid over the Brillouin zone. Secondly, maximally localized Wannier functions are derived using the method outlined in Refs.[69; 68] , which provides the unitary matrices $U_{\vec{k}}$. These matrices enable the calculation of $f_{mn\lambda\alpha}(\vec{R}_\Lambda,\vec{R}_A)$ values by means of Eq. (3.116). Once these values are obtained, interpolation can be performed to compute $f_{mn\nu}(\vec{k},\vec{q})$ on finer grids according to Eq. (3.115). To apply this equation, it is necessary to obtain the $U_{\vec{k}+\vec{q}}$ matrices on those finer grids. This can be achieved by employing the interpolation procedure described in Ref.[69].

3.5 Electron-lattice interaction from Second-Principles

After having examined electron-lattice interactions within the framework of first-principles methods, we now turn to the implementation of vibronic coupling in the second-principles formalism for electrons. This approach will be further elaborated in Chapter 4, within the context of the SCALE-UP code.

Electron-phonon interactions are ubiquitous in condensed matter systems; consequently, their inclusion in second-principles methods is essential for accurately capturing a wide range of physical phenomena. In particular, ABO$_3$ systems are prone to antiferrodistortive (AFD) phases, where the oxygen octahedra surrounding the B cation undergo tilting and rotation along the z-axis, leading to a slight deformation of the unit cell and resulting in a tetragonal lattice. These systems can also exhibit ferroelectric (FE) phases, characterized by an off-center cooperative effect in which cations displace oppositely with respect to the oxygen cage, giving rise to parallel dipoles within the unit cells. Additionally, they can display anti-ferroelectric (AFE) phases, distinguished by ferroelectric distortions that generate antiparallel dipoles. Furthermore, numerous compounds exhibit the Jahn-Teller effect, wherein geometric distortions arise due to the presence of a spatially degenerate electronic ground state. In other systems, such as those doped with an electron or a hole, lattice distortions can occur due to the formation of polarons. As a final example, in one-dimensional crystals with one electron per ion, a Peierls transition is observed. In this case, the system undergoes dimerization, as an equidistant one-dimensional crystal is inherently unstable. These geometric modifications alter bonding and orbital hybridization, necessitating the inclusion of electron-lattice coupling corrections in the Hamiltonian.

As mentioned before, the system's geometry determines its Hamiltonian. In the second-principles framework, the dependence of the model parameters on the atomic configuration is captured by the electron-lattice coupling terms. To calculate these parameters, it is recognized that in crystals, nuclei undergo low-amplitude oscillatory displacements around their equilibrium positions. This motion, close to their equilibrium positions, motivates a Taylor expansion of the Hamiltonian matrix elements, $h_{\mathbf{ab}}$. Considering the one-electron parameters γ_{ab} of the Hamiltonian matrix elements, which contain the part of the hamiltonian which does not take into account the electron-electron interaction,

$$\begin{aligned}\gamma_{ab}(\{\vec{\tau}_\lambda\}) = \gamma_{ab}(\{\vec{\tau}_\lambda^{(0)}\}) + \sum_{\lambda,\alpha} \left.\frac{\partial \gamma_{ab}}{\partial \tau_{\lambda\alpha}}\right|_{\text{RAG}} (\tau_{\lambda\alpha} - \tau_{\lambda\alpha}^{(0)}) + \\ + \frac{1}{2}\sum_{\lambda,\alpha}\sum_{\upsilon,\beta} \left.\frac{\partial^2 \gamma_{ab}}{\partial \tau_{\lambda,\alpha}\partial \tau_{\upsilon\beta}}\right|_{\text{RAG}} (\tau_{\lambda\alpha} - \tau_{\lambda\alpha}^{(0)})(\tau_{\upsilon\beta} - \tau_{\upsilon\beta}^{(0)})\end{aligned} \tag{3.117}$$

where the indices α and β correspond to the direction of the displacement in Cartesian coordinates, with $\alpha, \beta \in x, y, z$. Using the notation employed in Eq. (1.2) and neglecting the changes in the lattice vectors (homogeneous strains), the nuclear positions $\{\tau_{\lambda\alpha}\}$ can be replaced by the relative displacements $\{\vec{u}_\lambda\}$, for which $\vec{u} = \vec{0}$ represents the crystal's reference structure,

$$\gamma_{ab}(\{\vec{u}_\lambda\}) = \gamma_{ab}^{\text{RAG}} + \sum_{\lambda,\alpha} \left.\frac{\partial \gamma_{ab}}{\partial \tau_{\lambda\alpha}}\right|_{\text{RAG}} u_{\lambda\alpha} + \frac{1}{2}\sum_{\lambda,\alpha}\sum_{\upsilon,\beta} \left.\frac{\partial^2 \gamma_{ab}}{\partial \tau_{\lambda\alpha}\partial \tau_{\upsilon\beta}}\right|_{\text{RAG}} u_{\lambda\alpha}\, u_{\upsilon\beta}, \tag{3.118}$$

Attending to the previous expression, the zeroth-order term corresponds to the value of the Hamiltonian matrix element for the chosen reference geometry, while the contribution containing the first and second-order derivatives evaluated in the RAG represent the electron-lattice coupling corrections. The first-order derivative is denoted as

$$f_{ab,\lambda\alpha} = -\left.\frac{\partial \gamma_{ab}}{\partial \tau_{\lambda\alpha}}\right|_{\text{RAG}}, \tag{3.119}$$

A negative sign has been introduced to ensure that $f_{ab,\lambda\alpha}$ is positive when the system favors the displacement of the λ atom in the positive direction.

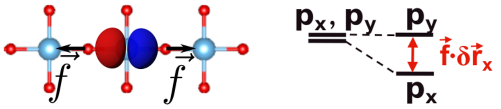
The linear term $\vec{f}_{ab,\lambda}$ is a three-dimensional vector and depends on the displacement of one atom. Meanwhile, the second-order derivative is expressed by

$$g_{ab,\lambda\alpha\ \upsilon\beta} = \left.\frac{\partial^2 \gamma_{ab}}{\partial \tau_{\lambda\alpha} \partial \tau_{\upsilon\beta}}\right|_{\text{RAG}} \tag{3.120}$$

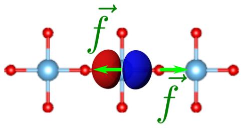
where $\overleftrightarrow{g}_{ab,\lambda\upsilon}$ is a second-rank tensor in three-dimensions which is associated to the displacement of two atoms. The terms $\vec{f}_{ab,\lambda}$ and $\overleftrightarrow{g}_{ab,\lambda\upsilon}$ represent first- and second-order tensors, respectively, that are closely associated with the concept of *orbital vibronic constants* as introduced in Eq. (3.39) for the linear case and in Eq. (3.59) for the quadratic contributions. These tensors are also intimately related to the electron-phonon coupling constants expressed in the Wannier function representation discussed in Sec. 3.4.8.

The $\vec{f}_{ab,\lambda}$ and $\overleftrightarrow{g}_{ab,\lambda\upsilon}$ elements are the electron-lattice parameters of the second-principles model. Rewriting Eq.(3.118) in terms of these parameters yields

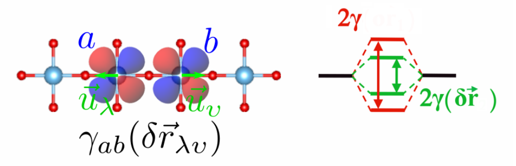
$$\gamma_{ab}(\{\vec{u}_\lambda\}) = \gamma_{ab}^{\text{RAG}} - \sum_{\lambda} \vec{f}^{\,T}_{ab,\lambda} \vec{u}_\lambda + \frac{1}{2} \sum_{\lambda} \sum_{\upsilon} \vec{u}^{\,T}_{\lambda} \overleftrightarrow{g}_{ab,\lambda\upsilon} \vec{u}_\upsilon, \tag{3.121}$$



(a) Diagonal term: from lattice to electrons



(b) Diagonal term: from electrons to lattice



(c) Off diagonal term

Figure 3.11: Examples of electron-lattice coupling effects. The diagonal term $\vec{f}^{\,T}_{aa,\lambda}$ produce a reduction of the on-site energy of a p_x orbital when a displacement on the x-direction is applied, (a) and the modification of the occupation of Wannier orbitals can induce atomic force in the neighbor atoms, (b). The off-diagonal term $\vec{f}^{\,T}_{ab,\lambda}$ generates alterations on the bandwidth when there is a relative displacement between atoms and thus, between orbitals, (c). In the picture we have neglected the Sr atoms since it does not participate in the active set. Figure extracted from Ref. [37] with permissions.

In second-principles, the linear terms $\vec{f}_{ab,\lambda}$ are the linear orbital vibronic constants presented in Eq. (3.39) expressed in atomic coordinates instead of modes. And, in the same way, this terms is equivalent to Eq. (3.113) in electron-phonon formulation described in Sec. 3.4.8. According to the second order terms, $\overleftrightarrow{g}_{ab}$, these elements correspond to the quadratic orbital vibronic constants in Eq. (3.59).

These electron-lattice terms influence the electronic structure by modulating the electrostatic interactions in several ways. In order to understand the physics of the $\vec{f}$ terms we study the following particular cases:

DIAGONAL INTERACTIONS, γ_{aa}. The interaction γ_{aa} controls the average energy (center of mass) of the corresponding band. Hereunder, we separate the effect produced by the lattice on the electrons, and the alterations generated by the electrons in the lattice.

From the lattice to electrons: the parameters depend on the underlying geometry

This interaction is related to $\vec{f}^{T}_{aa,\lambda}$ terms. If $\vec{f}^{T}_{aa,\lambda} \neq 0$, then the displacement of one atom, $\delta\vec{u}_{\lambda}$, might modify the center, shape and spreading of the Wannier χ_a, and therefore the center of gravity of the bands. An example is illustrated in Fig. 3.11a. It shows a p_x and a p_y-like Wannier orbitals and the presence of forces applied on one atom (spheres). If atoms are displaced along the x-direction, a change in the on-site energy of the p_x orbital is produced. In Fig. 3.11a the selected convention considers positive value of the $\vec{f}^{T}_{aa,\lambda}$ parameter. As a result, the energy of the p_x orbital is lowered while the energy of p_y does not change.

From the electrons to lattice: the forces depend on the electronic density

The force which suffers an atom λ is computed by

$$\vec{F}_{\lambda} = -\vec{\nabla}_{\lambda}E = -\vec{\nabla}_{\lambda}E^{(0)} - \sum_{ab} D^{U}_{ab}\vec{\nabla}_{\lambda}\gamma_{ab}. \tag{3.122}$$

The term of the deformation occupation matrix, D^{U}_{ab}, quantifies the difference between the density matrix and the reference density matrix (RED), which use to be identify with the ground state density of the neutral system. Then, if the occupation of a particular Wannier χ_a is changed by the injection of extra charge or by charge extraction, the matrix element $D^{U}_{aa} \neq 0$ and we are out of Born-Oppenheimer surface. The modification of the occupation can induce atomic forces on atom λ, Fig. 3.11b, and therefore a displacement. As a result, the force field is corrected by the second term in Eq. (3.122). In SPDFT model, these distortions are included in $\vec{f}_{aa}$.

OFF-DIAGONAL INTERACTIONS, γ_{ab}. The off-diagonal one-electron terms, γ_{ab}, govern the hopping mechanism, encompassing hybridization, covalency, and the bandwidth of electronic states. If $\vec{f}^{T}_{ab,\lambda\upsilon} \neq 0$, a relative displacement of atom λ can modify the overlap between the Wannier functions χ_a and χ_b, thereby altering the hopping terms. In other words, $\vec{f}$ quantifies the variation in bandwidth induced by atomic displacements. Figure 3.11c illustrates the interaction between two orbitals: the Wannier function χ_a, centered on atom λ, and χ_b, associated with atom υ, when atom λ undergoes displacement. If the displacement $\vec{u}_{\lambda}$ increases the separation between the two atoms, the Wannier functions move farther apart, reducing their overlap and consequently diminishing the hopping amplitude. As a result, the bandwidth narrows (depicted in green) compared to the undistorted case (depicted in red).

Chapter 4

Second-Principles: electronic contribution

The SPDFT method [37] is based on a expansion of the full DFT energy of a material in terms of the electron density, so that all the properties of the system like bands, spectra or geometries can be accurately predicted when perturbing the system. In order to reach large-scale simulations it is assumed that no strong changes in the bond topology of the system are going to take place [37]. For example, the method allows describing the typical ferroelectric transitions in materials like $BaTiO_3$ or $PbTiO_3$, where the underlying perovskite structure is retained, but would not be valid to take into account what happens during the melting of a solid into a liquid. The basic geometry of the system is given by the Reference Atomic Geometry (RAG), which was further explained in Sec. 1.2.1. As seen before, the RAG is typically selected as a high-symmetry critical point in the energy surface of the system. The atomic position of a λ atom in the RAG is defined by Eq. (1.1) while its position in a distorted geometry relative to the RAG is described by Eq. (1.2).

4.1 Electron density

From second-principles, the electronic density is treated as a perturbed quantity by splitting it into two parts: the *Reference Electron Density* (RED) and the *deformation charge density*, which represents the perturbation. The *Reference Electron Density*, $n_0(\vec{r})$, is defined for each possible atomic configuration. It is important to note that the concepts of RAG and RED are completely independent: in this methodology the RED is defined for every atomic structure accessible by the system, and not only for the RAG.

Similarly to perturbation theory, the RED would correspond to a simple electron density, as close as possible to the real density, $n(\vec{r})$. In order to determine $n_0(\vec{r})$, for the sake of simplicity, we consider a nonmagnetic system. In particular, we will focus on an insulator or semiconductor compound. Under these assumptions, the RED, $n_0(\vec{r})$, can be represented as the ground state and it can be obtained from DFT calculations. Now, if the system is perturbed a small *deformation charge density*, $\delta n(\vec{r})$, will appear, where the perturbation changes the density as,

$$n(\vec{r}) = n_0(\vec{r}) + \delta n(\vec{r}). \tag{4.1}$$

Once the foundation of the method has been established, the next step involves analyzing the energy. To do this, it is essential to define a basis to express the density. By employing a basis set of Bloch functions, which are characterized by a wave vector $\vec{k}$ and a band index j, the density can be written as

$$n(\vec{r}) = \sum_{j\vec{k}} o_{j\vec{k}} \left| \psi_{j\vec{k}}(\vec{r}) \right|^2. \tag{4.2}$$

In this context, $o_{j\vec{k}}$ denotes the occupation of the $|\psi_{j\vec{k}}\rangle$ state. At scales of tens of nanometers, where the number of atoms reaches hundreds of thousands, a localized basis set in real space that conveys the same information as the Bloch functions becomes essential to control the range of the interactions and help create a method that scales linearly with the number of atoms. Consequently, a Wannier function basis is employed [65]. This basis is not only localized but it is also orthonormal. Equation (1.37) allows expressing the Wannier functions in terms of the Bloch basis.

In this section and the subsequent ones, we introduce a change in notation relative to the previous chapters to simplify the formalism and align with the notation used in the reference work [37]. A Wannier function, originally denoted as $|\chi_{a\vec{R}_A}\rangle$, is centered at the Bravais lattice vector $\vec{R}_A$ and labeled within that unit cell by the discrete index a. In this section, we adopt the more compact notation $|\chi_{\boldsymbol{a}}\rangle$, where the subscript $\boldsymbol{a}$ represents both the unit cell and the discrete index a, such that $a\vec{R}_A \leftrightarrow \boldsymbol{a}$. Throughout this work, latin indices are used to label all physical quantities associated with electrons. According to this convention, the Wannier functions are expressed as Eq. (1.37)

$$|\chi_{\boldsymbol{a}}\rangle = \frac{V_{\text{cell}}}{(2\pi)^3} \int_{\text{BZ}} e^{-i\vec{k}\cdot\vec{R}_A} \sum_{j=1}^{J} U_{ja}(\vec{k}) \left| \psi_{j\vec{k}}^{(0)} \right\rangle d\vec{k}, \tag{4.3}$$

where the superscript "(0)" in the Bloch functions $|\psi_{j\vec{k}}^{(0)}\rangle$ indicates the Bloch states of the system corresponding to the RAG in the RED configuration, i.e. the *unperturbed* system. Note that the inverse transformation from Wannier to Bloch functions reads

$$\left| \psi_{jk}^{(0)} \right\rangle = \sum_{a} c_{j\vec{k},a}^{(0)} e^{i\vec{k}\cdot\vec{R}_A} |\chi_{\boldsymbol{a}}\rangle. \tag{4.4}$$

The Wannier functions corresponding to the RED form a complete basis of the Hilbert space. Hence, we can use them to represent any perturbed electronic configuration of the system as

$$|\psi_{jk}\rangle = \sum_{a} c_{j\vec{k},a} e^{i\vec{k}\cdot\vec{R}_A} |\chi_{\boldsymbol{a}}\rangle, \tag{4.5}$$

where the sum can be extended to as many bands as needed to accurately describe the phenomenon of interest.

In this way, Eq. (4.2) can be reformulated using real Wannier functions, eliminating the need for complex conjugates. The resulting expression is

$$n(\vec{r}) = \sum_{j\vec{k}} o_{j\vec{k}} \psi^*_{j\vec{k}}(\vec{r}) \psi_{j\vec{k}}(\vec{r}) = \sum_{j\vec{k}} \sum_{\boldsymbol{ab}} o_{j\vec{k}} c^*_{j\vec{k},a} c_{j\vec{k},b} e^{i\vec{k}\cdot(\vec{R}_B - \vec{R}_A)} \chi_{\boldsymbol{a}}(\vec{r}) \chi_{\boldsymbol{b}}(\vec{r}) = \sum_{\boldsymbol{ab}} d_{\boldsymbol{ab}} \chi_{\boldsymbol{a}}(\vec{r}) \chi_{\boldsymbol{b}}(\vec{r}). \tag{4.6}$$

There, $d_{\boldsymbol{ab}}$ defines the charge distribution and is called as the *density matrix*,

$$d_{\boldsymbol{ab}} = \sum_{j\vec{k}} o_{j\vec{k}} c^*_{j\vec{k},a} c_{j\vec{k},b} e^{i\vec{k}\cdot(\vec{R}_B - \vec{R}_A)}. \tag{4.7}$$

Similarly to the total density $n(\vec{r})$ expressed in terms of Wannier functions in Eq. (4.6), the reference density $n_0(\vec{r})$ can be written in this basis as follows,

$$n_0(\vec{r}) = \sum_{\boldsymbol{ab}} d^{(0)}_{\boldsymbol{ab}} \chi_{\boldsymbol{a}}(\vec{r}) \chi_{\boldsymbol{b}}(\vec{r}), \tag{4.8}$$

where $d^{(0)}_{\boldsymbol{ab}}$ is defined as the *reference density matrix*.

Once the total density, $n(\vec{r})$, and the reference density, $n_0(\vec{r})$, have been defined in terms of Wannier functions, we introduce a fundamental quantity that underpins the development of the electronic energy within this methodology. This quantity is the *deformation charge density*,

$$\delta n(\vec{r}) = \sum_{\boldsymbol{ab}} D_{\boldsymbol{ab}} \chi_{\boldsymbol{a}}(\vec{r}) \chi_{\boldsymbol{b}}(\vec{r}), \tag{4.9}$$

which is expressed in terms of the *difference density matrix* $D_{\boldsymbol{ab}}$. This matrix quantifies the difference between the total and reference densities, as given by Eq. (4.6) and Eq. (4.8),

$$D_{\boldsymbol{ab}} = d_{\boldsymbol{ab}} - d^{(0)}_{\boldsymbol{ab}}. \tag{4.10}$$

Its elements are supposed to be small in the sense that the trace of the difference density is much smaller than the number of electrons in the system, $\sum_{\boldsymbol{a}} |D_{\boldsymbol{aa}}| \ll \sum_{\boldsymbol{a}} d^{(0)}_{\boldsymbol{aa}}$.

This deformation charge density may contain interactions of the following nature:

- The charge doping by the addition of electrons to the bottom of the conduction band or by extracting electrons from the top of the valence band. This phenomenon is related with the simulation of *polarons*.
- The *excitation* of the electronic charge density in the neutral system by the transference of electrons from the valence to the conduction band. In this case, this effect is related to the optical excitations including exciton creation.
- For each one of the commented phenomenons, $\delta n(\vec{r})$ must capture the charge redistribution (response) of the remaining electrons to the induced change, *screening* phenomenon. It can be viewed as resulting from an admixture of occupied and unoccupied states of the reference neutral configuration since the eigenstates represent a basis of the Hilbert space.

It is important to highlight that the three quantities, $n(\vec{r})$, $n_0(\vec{r})$, and $\delta n(\vec{r})$, are, in fact, parametric functions of the atomic positions $\{\vec{\tau}_\lambda\}$. However, in the case of a non-magnetic insulator, the ground state, characterized by the reference density matrix $d^{(0)}_{\boldsymbol{ab}}$, remains independent of the atomic structure of the lattice and characterizes the reference adiabatic potential energy surface (APES). As an illustrative example, we consider the insulator SrTiO$_3$ in its cubic paraelectric configuration, where a Wannier function basis associated with the Bloch states at the top of the valence band and the bottom of the conduction band is employed. When the ground state is taken as the reference electronic configuration, the reference density matrix $d^{(0)}$ assumes a diagonal form, consisting of twos and zeros along the diagonal, independently of the geometry of the system. The alteration of the geometry is implicitly accounted in the change of the shape of the WFs, including its spread.

4.2 Approximate expressions for the energy

The primary goal of this Section is to derive a functional form that enables an accurate approximation of the DFT total energy while minimizing computational expense. The DFT energy can be expressed as

$$E_{\mathrm{DFT}} = \sum_{j\vec{k}} o_{j\vec{k}} \left\langle \psi_{j\vec{k}} \middle| t + v_{\mathrm{ext}} \middle| \psi_{j\vec{k}} \right\rangle + \frac{1}{2} \iint \frac{n(\vec{r})n(\vec{r}')}{|\vec{r}-\vec{r}'|} d^3r d^3r' + E_{\mathrm{xc}}[n] + E_{\mathrm{nn}}. \tag{4.11}$$

From the previous expression, the DFT energy is formulated as a functional of the electronic density, which contributes to both the Coulomb interaction and the exchange-correlation energy. By applying the decomposition of the electronic density, outlined in Eq. (4.1), the linear part of the Coulomb energy term can be addressed straightforwardly. Moreover, an appropriate choice of the RED, that ensures $D_{\boldsymbol{ab}}$ remains small, enables an accurate approximation of the exchange-correlation energy. This approximation is achieved through a low-order Taylor expansion of δn around the RED, as proposed in reference [172],

$$E_{\mathrm{xc}}[n] = E_{\mathrm{xc}}[n_0] + \int \left. \frac{\delta E_{\mathrm{xc}}}{\delta n(\vec{r})} \right|_{n_0} \delta n(\vec{r}) d^3r + \frac{1}{2} \iint \left. \frac{\delta^2 E_{\mathrm{xc}}}{\delta n(\vec{r}) \delta n(\vec{r}')} \right|_{n_0} \delta n(\vec{r}) \delta n(\vec{r}') d^3r d^3r' + \cdots . \tag{4.12}$$

Under the aforementioned approximation, the total energy can be expressed as the sum of contributions arising from the reference density (zeroth order on density) corrected by first and second orders on the density, related to the deformation density. By performing the expansion up to second order, the final expression closely resembles that of the Hartree-Fock method. Formally, this is written as

$$E_{\mathrm{DFT}} \approx E = E^{(0)} + E^{(1)} + E^{(2)}. \tag{4.13}$$

The method is systematically improvable by incorporating higher-order energy terms in the expansion, while consistently seeking an optimal balance between accuracy and computational cost. Now, we study the different terms of the expansion step by step in order to elucidate the specific phenomena they represent.

To achieve a better understanding of the concepts of RED, deformation density and the involved energy correction terms, an example is illustrated in Fig. 4.1, where the different densities are represented in the top cartoons. Here, the proposed splitting of the electronic density is applied on a doped semiconductor composed by two different types of atoms: large green/blue and small red balls in a 2D-square geometry and 3 atoms as motif. The RAG corresponds to the high-symmetry configuration in which the large atom is located at the center of the square, while the small atoms lie at the centers of the sides. The neutral system (undoped) is associated to the RED, Figs. 4.1(b). Here, the system is in the ground state: the valence band is completely occupied while the conduction band is empty. At this moment, if some charge is added (or removed) Figs. 4.1(c), 4.1(f), the electronic cloud will be affected and will tend to screen the field caused by the extra charge. The doping electron (respectively, hole) will occupy the states at the bottom of the conduction band (respectively, top of the valence band). Such a deformation density, captures both the doping and the system's response to it, as previously exposed. The doping is represented by green balls in the electronic band diagrams while the the screening is shown in orange.

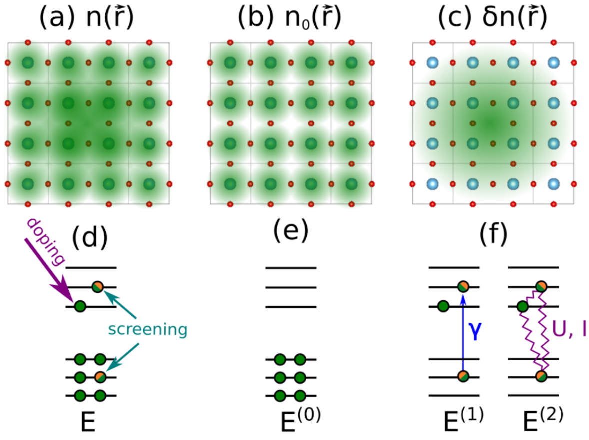


Figure 4.1: Schematic representation of the reference atomic structure and the reference and deformation electron densities. Illustrations (a)–(c) show the total density, the RAG and the deformation density for a semiconductor. Illustrations (d)–(f) exhibit the energy levels obtained for the neutral ground state. The occupation of a given state is determined by the full green circles. And the partial occupation is given by half filled orange/green circles. The notations $E^{(0)}$, $E^{(1)}$ and $E^{(0)}$ for the energies is related to the reference energy, one-electron contribution to the energy and two-electron contribution respectively. Figure extracted from Ref. [37] with permissions.

4.2.1 Lattice models

$E^{(0)}$ - Energy of the reference state

The zeroth order term corresponds, without approximation, to the full DFT energy for the reference density $n_0(\vec{r})$

$$E^{(0)} = \sum_{j\vec{k}} o_{j\vec{k}}^{(0)} \left\langle \psi_{j\vec{k}}^{(0)} \right| t + v_{\text{ext}} \left| \psi_{j\vec{k}}^{(0)} \right\rangle + \frac{1}{2} \iint \frac{n_0(\vec{r}) n_0'(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r d^3r' + E_{\text{xc}}[n_0] + E_{nn}. \tag{4.14}$$

It is the dominant contribution to the total energy. We can compute $E^{(0)}$ as a function of the variables controlling the geometry: the strain $\overset{\leftrightarrow}{\eta}$ and the displacements from the RAG $\{\vec{u}\}$, by employing a model potential that depends only on the atomic positions, and where the electrons are assumed to remain in the RED and are integrated out. Here is where the lattice models, explained in Sec. 1.2.2, come into play. This approach offers a significant advantage over other methods, such as standard DFTB techniques, which require both an explicit and precise treatment of electronic interactions to obtain the RED, as well as the numerical solution for $E^{(0)}$ and n_0 for each atomic configuration within the simulation.

The zeroth-order energy exclusively accounts for the lattice degrees of freedom and does not include the electronic degrees of freedom. Therefore, to analyze the band structure of a system, it is essential to incorporate the relevant electronic degrees of freedom in an appropriate manner. In particular, this methodology includes the electronic degrees of freedom through a tight-binding-like model (see Sec. 1.3) while ensuring the avoidance of double-counting with the effective atomic potentials.

4.2.2 Electronic models

As discussed above, the electronic band structure of a system can be computed by a simple tight binding model, Eq. (1.65). This model presents several advantages, as the low computational cost and the high efficiency. The main ingredient for its application are the Hamiltonian matrix elements expressed in the Wannier functions basis set shown in Eq. (1.65). Thus, the objective of the method is the construction of these matrix elements parametrically.

The tight-binding-like Hamiltonian matrix elements $h^s_{\boldsymbol{ab}}$, defined within this framework to account for electronic degrees of freedom, are determined by the following expression for the spin channel s,

$$h^s_{\boldsymbol{ab}} = \langle \chi^s_{\boldsymbol{a}} | h | \chi^s_{\boldsymbol{b}} \rangle \tag{4.15}$$

where h is the Kohn-Sham [160] one-electron Hamiltonian defined for the total density $n(\vec{r})$,

$$h = t + v_{\text{ext}} - v_{\text{H}}(n;\vec{r}) + v_{\text{xc}}[n;\vec{r}]\,. \tag{4.16}$$

The term $v_H(n;\vec{r})$ represents the Hartree potential

$$v_H(n_0;\vec{r}) = -\int \frac{n_0(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|} d^3r', \tag{4.17}$$

and v_{xc} is the exchange-correlation potential

$$v_{xc}[n;\vec{r}] = \left.\frac{\delta E_{xc}[n]}{\delta n(\vec{r})}\right|_n . \tag{4.18}$$

Electrostatics

Regarding the tight-binding Hamiltonian terms, the methodology distinguishes between short- and long-range contributions, denoted as $h^{s,\text{sr}}_{\boldsymbol{ab}}$ and $h^{s,\text{lr}}_{\boldsymbol{ab}}$, respectively,

$$h^s_{\boldsymbol{ab}} = h^{s,\text{sr}}_{\boldsymbol{ab}} + h^{s,\text{lr}}_{\boldsymbol{ab}}. \tag{4.19}$$

In the Hamiltonian, the long-range contribution corresponds to the Coulomb interaction, represented by the electrostatic potential generated by both electrons and nuclei acting on the Wannier functions $\chi_{\boldsymbol{a}}$ and $\chi_{\boldsymbol{b}}$. The long-range contribution is computed using a far-field electrostatic potential, where the lattice effects are described through a multipole expansion representing them as effective charges and dipoles.

The remaining contributions to the system, such as kinetic energy, exchange-correlation effects, and external applied fields, are local or semi-local. The short-range component, which captures these local interactions, must be explicitly modeled within the second-principles framework.

$\boldsymbol{E^{(1)}}$ - One-electron excitations

The term $E^{(1)}$ contains the differences in one-electron energies. It involves the one-electron excitations as captured by the deformation density

$$E^{(1)} = \sum_{j\vec{k}} \left[o_{j\vec{k}} \left\langle \psi_{j\vec{k}} \middle| h_0 \middle| \psi_{j\vec{k}} \right\rangle - o^{(0)}_{j\vec{k}} \left\langle \psi^{(0)}_{j\vec{k}} \middle| h_0 \middle| \psi^{(0)}_{j\vec{k}} \right\rangle \right], \tag{4.20}$$

where h_0 is the one-electron Hamiltonian, Eq. (4.16), defined for the RED, n_0. Attending to Eq. (4.20), it is important to emphasize that the summation is performed over the difference between the one-electron energies of the actual system and those of the reference system. In contrast, standard DFTB schemes typically incorporate a direct summation of one-electron energies (for example, see references [61] and [172]). By focusing on the difference between these values rather than their absolute magnitudes, this approach yields a much smaller quantity, which is more suitable for accurate calculations.

The first-order term of the energy $E^{(1)}$ can be expressed in the Wannier functions basis set as

$$E^{(1)} = \sum_{\boldsymbol{ab}} D_{\boldsymbol{ab}} \gamma_{\boldsymbol{ab}}, \tag{4.21}$$

where $\gamma_{\boldsymbol{ab}}$ represents the one-electron operator corresponding to the tight-binding-like interaction

$$\gamma_{\boldsymbol{ab}} = \langle \chi_{\boldsymbol{a}} | h_0 | \chi_{\boldsymbol{b}} \rangle = \int \chi_{\boldsymbol{a}}(\vec{r}) h_0(\vec{r}) \chi_{\boldsymbol{b}}(\vec{r}) \, d^3r. \tag{4.22}$$

If we consider the division in short-range and long-range interaction,

$$\gamma_{\boldsymbol{ab}} = \gamma^{\text{sr}}_{\boldsymbol{ab}} + \gamma^{\text{lr}}_{\boldsymbol{ab}}. \tag{4.23}$$

The short-range term on the one-electron integral, $\gamma^{\text{sr}}_{\boldsymbol{ab}}$ is highly sensitive to the system's geometry. Changes in geometry influence the spread and centers of the Wannier functions, reflecting modifications in hybridization. In this method, small geometric distortions are considered, allowing for the incorporation of electron-lattice coupling corrections, $\delta\gamma^{\text{sr,el}}_{\boldsymbol{ab}}$, to the one-electron parameters obtained for the RAG, $\gamma^{\text{sr,RAG}}_{\boldsymbol{ab}}$. These corrections, which account for first- and second-order modifications in the one-electron integrals, effectively describe variations in the Wannier functions and their interaction with the lattice. To analyze this coupling, $\gamma^{\text{sr}}_{\boldsymbol{ab}}$ is expanded in terms of atomic displacements through a Taylor series around the RAG, under the assumption of small deformations,

$$\gamma^{\text{sr}}_{\boldsymbol{ab}} = \gamma^{\text{sr,RAG}}_{\boldsymbol{ab}} + \delta\gamma^{\text{sr,el}}_{\boldsymbol{ab}} = \gamma^{\text{sr,RAG}}_{\boldsymbol{ab}} + \sum_{\boldsymbol{\lambda v}} \left[-\vec{f}^{\,T}_{\boldsymbol{ab,\lambda v}} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda v}} + \sum_{\boldsymbol{\lambda' v'}} \delta\vec{\tau}^{\,T}_{\boldsymbol{\lambda v}} \cdot \overset{\leftrightarrow}{g}_{\boldsymbol{ab,\lambda v}} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda v}} + \ldots \right], \tag{4.24}$$

Here, $\delta\vec{\tau}_{\boldsymbol{\lambda}\boldsymbol{\upsilon}}$ quantifies the relative displacement between atoms $\boldsymbol{\lambda}$ and $\boldsymbol{\upsilon}$,

$$\delta\vec{\tau}_{\boldsymbol{\lambda}\boldsymbol{\upsilon}} = \overleftrightarrow{\eta}\left(\vec{R}_{\Upsilon} - \vec{R}_{\Lambda} + \vec{\tau}^{(0)}_{\upsilon} - \vec{\tau}^{(0)}_{\lambda}\right) + \vec{u}_{\boldsymbol{\upsilon}} - \vec{u}_{\boldsymbol{\lambda}}. \tag{4.25}$$

The previous expression for $\gamma^{\mathrm{sr}}_{\boldsymbol{ab}}$ term is related to Eq. (3.121). A comparison of both expressions reveals that the linear terms involve two atoms, with displacements represented as relative differences between atomic positions rather than absolute displacements of individual atoms. The methodology presented in [37] follows this approach to automatically enforce the ASR which can easily be violated due to small numerical errors. In the following Chapter, we will introduce a new methodology aimed at addressing some limitations in the formula Eq. (4.24).

Regarding Eq. (4.24), $\vec{f}$ and $\overleftrightarrow{g}$ are the first- and second-rank tensors that characterize the electron-lattice coupling. These terms correspond to those in Sec. 3.5, and their relationships will be established in Sec. 5.6.3.

$E^{(2)}$ - Two-electron excitations

$E^{(2)}$ contains the two-electrons interactions

$$E^{(2)} = \frac{1}{2}\int d^3r \int d^3r' g\left(\vec{r},\vec{r}'\right)\delta n(\vec{r})\delta n\left(\vec{r}'\right) \tag{4.26}$$

where the screened electron-electron interaction operator, $g\left(\vec{r},\vec{r}'\right)$, is given by

$$g\left(\vec{r},\vec{r}'\right) = \frac{1}{|\vec{r}-\vec{r}'|} + \left.\frac{\delta^2 E_{\mathrm{xc}}}{\delta n(\vec{r})\delta n\left(\vec{r}'\right)}\right|_{n_0} \tag{4.27}$$

where δn represents the deformation density exposed in Eq. (4.9). This expression can be extended to deal with strongly-correlated and magnetic systems taking into account in a differentiate way the two spin channels. Thus, the deformation density can be expressed in a similar way for each one of the spin channels s defining $D^s_{\boldsymbol{ab}}$. As a result, the deformation density $\delta n(\vec{r})$ is given by

$$\delta n(\vec{r},s) = \sum_{\boldsymbol{ab}} D^s_{\boldsymbol{ab}}\chi_{\boldsymbol{a}}(\vec{r})\chi_{\boldsymbol{b}}(\vec{r}) \tag{4.28}$$

where

$$D^{\uparrow}_{\boldsymbol{ab}} = d^{\uparrow}_{\boldsymbol{ab}} - \frac{1}{2}d^{(0)}_{\boldsymbol{ab}} \qquad\qquad D^{\downarrow}_{\boldsymbol{ab}} = d^{\downarrow}_{\boldsymbol{ab}} - \frac{1}{2}d^{(0)}_{\boldsymbol{ab}} \tag{4.29}$$

The difference density matrix $D_{\boldsymbol{ab}}$ (including both spin channels) can be expressed as the sum of the two spin components and is renamed by $D^U_{\boldsymbol{ab}}$. Conversely, the system's magnetization charge, $D^I_{\boldsymbol{ab}}$, is obtained as the difference between the two spin components,

$$D^U_{\boldsymbol{ab}} = D^{\uparrow}_{\boldsymbol{ab}} + D^{\downarrow}_{\boldsymbol{ab}} \qquad\qquad D^I_{\boldsymbol{ab}} = D^{\uparrow}_{\boldsymbol{ab}} - D^{\downarrow}_{\boldsymbol{ab}}. \tag{4.30}$$

The $D^I_{\boldsymbol{ab}}$ matrix only plays a role in the spin-polarized systems and vanishes in non-spin-polarized compounds because a non-spin-polarized density satisfies $D^{\uparrow}_{\boldsymbol{ab}} = D^{\downarrow}_{\boldsymbol{ab}}$.

Once the difference density matrix has been defined as a function of the spin, $D^s_{\boldsymbol{ab}}$, the general expression for the energy second order correction expressed in a Wannier function basis set is

$$E^{(2)} = \frac{1}{2}\sum_{\boldsymbol{ab}}\sum_{\boldsymbol{a'b'}}\left\{D^U_{\boldsymbol{ab}}D^U_{\boldsymbol{a'b'}}U_{\boldsymbol{ab},\boldsymbol{a'b'}} - D^I_{\boldsymbol{ab}}D^I_{\boldsymbol{a'b'}}I_{\boldsymbol{ab},\boldsymbol{a'b'}}\right\} \tag{4.31}$$

where U and I are the two-electron parameters. As an approximation, fluctuations in the interaction parameters U and I with geometry are neglected, and only the dependence of the one-electron integrals $\gamma_{\boldsymbol{ab}}$ on the lattice distortions is considered.

The associated electron-electron coupling correction added to the RED hamiltonian is studied in terms of these parameters

$$\delta h^{s,\mathrm{ee}}_{\boldsymbol{ab}} = \sum_{\boldsymbol{a'b'}}\left[\left(D^s_{\boldsymbol{a'b'}} + D^{-s}_{\boldsymbol{a'b'}}\right)U_{\boldsymbol{ab},\boldsymbol{a'b'}} - \left(D^s_{\boldsymbol{a'b'}} - D^{-s}_{\boldsymbol{a'b'}}\right)I_{\boldsymbol{ab},\boldsymbol{a'b'}}\right]. \tag{4.32}$$

The two-electron parameters U and I are defined as four-center integrals. Considering $X = U, I$,

$$X_{\boldsymbol{ab},\boldsymbol{a'b'}} = \int \mathrm{d}^3r \int \mathrm{d}^3r' \chi_{\boldsymbol{a}}(\vec{r},s)\chi_{\boldsymbol{b}}(\vec{r},s)\chi_{\boldsymbol{a'}}(\vec{r}\,',s)\chi_{\boldsymbol{b'}}(\vec{r}\,',s)g_X(\vec{r},\vec{r}\,',s,s'), \tag{4.33}$$

where g^U and g^I are the Hubbard and Stoner two-electron constants. The U and I terms match to the four-index integrals of Hartree-Fock theory. The Hubbard two-electron constant is computed as

$$g_U(\vec{r},\vec{r}\,') = \frac{1}{|\vec{r}-\vec{r}\,'|} + \frac{1}{2}\left[\left.\frac{\delta^2 E_{\mathrm{xc}}}{\delta n(\vec{r},\uparrow)\delta n(\vec{r}\,',\uparrow)}\right|_{n_0} + \left.\frac{\delta^2 E_{\mathrm{xc}}}{\delta n(\vec{r},\uparrow)\delta n(\vec{r}\,',\downarrow)}\right|_{n_0}\right]. \tag{4.34}$$

It contains the classical electrostatic Hartree interaction (first term of the expression). This contribution is corrected by the second term where the elements $\delta^2 E_{\mathrm{xc}}/\delta n(\vec{r},s)\delta n(\vec{r}\,',s')$ capture the effective screening of the two-electron interactions due to exchange and correlation. It depends on the total difference density matrix described above, $D^U_{\boldsymbol{ab}}$, and quantifies the energy needed to add or remove electrons.

The Stoner two-electron constant is related to the spin polarization and is given by the expression

$$g_I(\vec{r},\vec{r}\,') = \frac{1}{2}\left[\left.\frac{\delta^2 E_{xc}}{\delta n(\vec{r},\uparrow)\delta n(\vec{r}\,',\downarrow)}\right|_{n_0} - \left.\frac{\delta^2 E_{xc}}{\delta n(\vec{r},\uparrow)\delta n(\vec{r}\,',\uparrow)}\right|_{n_0}\right]. \tag{4.35}$$

This magnitude considers the magnetic interactions providing the difference in interaction between electrons with parallel and antiparallel spins. Its origin is essentially quantum due to its exchange-correlation nature.

Finally, the SPDFT energy in Eq. (4.13), when using the WFs basis, takes the expression,

$$\begin{aligned} E =& E^{(0)} + \sum_{\boldsymbol{ab}}\left(D^{\uparrow}_{\boldsymbol{ab}} + D^{\downarrow}_{\boldsymbol{ab}}\right)\gamma_{\boldsymbol{ab}} + \\ &+ \frac{1}{2}\sum_{\boldsymbol{ab}}\sum_{\boldsymbol{a'b'}}\left\{\left(D^{\uparrow}_{\boldsymbol{ab}} + D^{\downarrow}_{\boldsymbol{ab}}\right)\left(D^{\uparrow}_{\boldsymbol{a'b'}} + D^{\downarrow}_{\boldsymbol{a'b'}}\right)U_{\boldsymbol{ab},\boldsymbol{a'b'}} - \right. \\ &\left. - \left(D^{\uparrow}_{\boldsymbol{ab}} - D^{\downarrow}_{\boldsymbol{ab}}\right)\left(D^{\uparrow}_{\boldsymbol{a'b'}} - D^{\downarrow}_{\boldsymbol{a'b'}}\right)I_{\boldsymbol{ab},\boldsymbol{a'b'}}\right\}. \end{aligned} \tag{4.36}$$

The bands and Bloch wavefunctions can be extracted using the one-electron Hamiltonian Eq. (4.15) whose final expression for the *s* spin-channel considering the electrostatic contributions, electron-lattice coupling $\delta\gamma_{\boldsymbol{ab}}^{\text{sr,el}}$ and the electron-electron correction, $\delta h_{\boldsymbol{ab}}^{s,\text{ee}}$,

$$\begin{aligned} h_{\boldsymbol{ab}}^{s} =& \gamma_{\boldsymbol{ab}}^{\text{sr,RAG}} + \sum_{\boldsymbol{\lambda v}} \left[-\vec{f}_{\boldsymbol{ab},\boldsymbol{\lambda v}}^{T} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda v}} + \delta\vec{\tau}_{\boldsymbol{\lambda v}}^{T} \cdot \overset{\leftrightarrow}{g}_{\boldsymbol{ab},\boldsymbol{\lambda v}} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda v}} \right] + \\ & + \sum_{\boldsymbol{a'b'}} \left[\left(D_{\boldsymbol{a'b'}}^{s} + D_{\boldsymbol{a'b'}}^{-s} \right) U_{\boldsymbol{ab},\boldsymbol{a'b'}} - \left(D_{\boldsymbol{a'b'}}^{s} - D_{\boldsymbol{a'b'}}^{-s} \right) I_{\boldsymbol{ab},\boldsymbol{a'b'}} \right] + h_{\boldsymbol{ab}}^{\text{lr}}. \end{aligned} \tag{4.37}$$

As a result, we have a mean field approach consisting of a tight-binding one-electron Hamiltonian corrected by electron-electron interactions has been derived. It depends parametrically on the parameters $\gamma_{\boldsymbol{ab}}^{\text{sr,RAG}}$, $\vec{f}_{\boldsymbol{ab},\boldsymbol{\lambda v}}$, $\overset{\leftrightarrow}{g}_{\boldsymbol{ab},\boldsymbol{\lambda v}}$, $U_{\boldsymbol{ab},\boldsymbol{a'b'}}$, $I_{\boldsymbol{ab},\boldsymbol{a'b'}}$ and $h_{\boldsymbol{ab}}^{\text{lr}}$. The method aimed to determine these parameters through first-principles simulations, with the primary objective of obtaining a parametrization flexible enough to describe various atomic configurations and electronic states beyond the ground state surface.

The present thesis has focused on the computation of these parameters by refining the existing methodology and proposing alternative approaches.

Chapter 5

MODELMAKER

In the previous Chapter, based on Ref. [37], we introduced *second-principles density-functional-theory* (SPDFT) calculations, designed to address large-scale problems including both atomic and electronic degrees of freedom on the same footing in systems with many atoms and extended simulation times. This method constructs general models capable of describing diverse properties, mechanical, structural, electrical, or magnetic, while maintaining a computational cost significantly lower than that of standard DFT calculations. SPDFT models can capture essential physical properties such as total energy, electronic bands, and magnetic moments for systems containing tens or even hundreds of thousands of atoms. Importantly, this approach retains predictive accuracy, as its parameters are systematically derived from DFT calculations and can, in principle, be refined to converge toward first-principles results.

Since the original publication of second-principles methods concerning the electronic degrees of freedom [37], machine-learning (ML) models in condensed matter physics and chemistry have proliferated [29; 30]. Universal deep neural network approaches to represent the DFT Hamiltonian of crystalline materials, aiming to bypass the computationally demanding self-consistent field iterations of DFT and substantially improve the efficiency of *ab initio* electronic-structure calculations, are already available [31; 32]. Machine learning and artificial intelligence models have proven to be highly accurate in condensed matter physics, capable of predicting material properties, phase transitions, and electronic structures in large systems with remarkable precision. However, the complexity of these models, often involving deep neural networks or high-dimensional feature spaces, makes them inherently opaque. Unlike traditional physics-based approaches, where analytical equations offer clear insights into underlying mechanisms, ML-driven models function as "black boxes", providing results without a transparent explanation of the reasoning behind them. This lack of interpretability poses a challenge, as scientists may obtain precise predictions but struggle to extract fundamental physical principles or causal relationships from the model's inner workings. Moreover, ML-based approaches are often tailored to specific properties (such as energy gap predictions [33]) or materials (for instance, the mechanical properties of zeolites [34]). In contrast, SPDFT offers a physics-driven framework that applies to a broader range of systems and properties, making it a valuable complement to data-driven methodologies.

The methodology developed in Ref. [37] and briefly described in Chapter 4, is based on the construction of the one-electron Hamiltonian matrix elements represented in a Wannier functions basis set [65]. As shown in Eq. (4.37), these Hamiltonian matrix elements are expressed as a function of atomic displacements and density matrix variations; and depend parametrically on a set of parameters which are obtained from a collection of first-principles calculations which conform the *training set*. These parameters define the models and are the key to all future work in second-principles.

In Ref. [37] a first algorithm to generate the different parameters composing the model was described. In this initial version, these Hamiltonians were obtained from DFT calculations on a loosely defined training set, allowing users to incorporate specific configurations relevant to the system under study. However, this approach had several limitations: (i) It did not explicitly enforce the symmetry of the system, leading to discrepancies in equivalent parameters and potential errors in band structure representation. (ii) The lack of a formally defined training set made it difficult to compare different models, as they contained non-equivalent terms derived from disparate DFT calculations. (iii) Model generation required an extensive number of DFT calculations (potentially tens of thousands), incurring significant computational costs, which contradicted the efficiency goals of the method. (iv) There was no systematic way to assess the quality of a given model, making it challenging to refine and improve parameter sets. Considering these limitations, in this work we have improved the SPDFT model generation method to correct these main deficiencies.

To address the previously mentioned problems, we propose a new automated model construction algorithm in which most parameters are either directly calculated or fitted, thereby ensuring model consistency. The user needs to only choose a few computational thresholds that are then used to systematically generate the training set and control the overall quality of the resulting model. This new method requires a significant interaction of the model-generating toolbox developed in python code (MODELMAKER) with the first-principles code. Thus, we have selected SIESTA [38] as the main DFT program to interface with MODELMAKER, although adapting it to other codes that employ WANNIER90 [39] to build WFs should be a simple task. As described later in Sec. 5.3, we have significantly upgraded the capacity of SIESTA to deal with the simultaneous Wannierization of different sets of bands.

During this work and since the original publication of the SPDFT method [37], we have identified ways to address some of the limitations of that work. Below, we report several important changes to the SPDFT Hamiltonian that are now implemented.

Wannierization. An issue that negatively impacted the accuracy of the second-principles models obtained with the previous method was the way the Wannierization procedure was carried out. Typically, all WFs for a system are obtained in a single run where one selects all the bands of interest and produces the corresponding WFs. However, if one applies this procedure, for example, to the valence and conduction bands of an insulator, the resulting WFs are described, to maximize localization, by a mixture of valence and conduction Bloch functions. Since valence bands are fully occupied in an insulator while the conduction ones are empty, the resulting density matrix for the ground state, see Sec. 5.2, is not diagonal nor constant with changes in the geometry, which involves a significant computational penalty when carrying out, for example, molecular dynamics simulations.

Symmetry. One-electron Hamiltonians expressed in the WFs basis often exhibited small numerical inconsistencies. These resulted in symmetry-related terms being treated as distinct, which in turn led to unphysical features in the second-principles models—such as small band splittings where degeneracies were expected.

Training set calculations. The approach in the original SPDFT formulation [37] emphasized the importance of filtering the large volume of data produced by first-principles calculations. However, less attention was given to the construction of the training set, that is, to defining the specific first-principles simulations needed to parameterize a second-principles model.

Description of the electrostatics. The description of electrostatics in Ref. [37] involved local dipoles that only depended on the geometry of the system. However, when calculating optical properties [173] we found that atomic dipoles need to be allowed to vary as a function of time due to the hybridization between WFs that is induced by external electric fields.

Next, we present the methodology on which MODELMAKER is based, we compare it with the algorithm proposed in Ref. [37] and analyze in detail the issues it has encountered, and we discuss the improvements introduced in this work.

Considering the system for which we are going to build the electronic model, the starting point of our methodology is the identification of the Reference Atomic Geometry (RAG) and the Reference Electronic Density (RED).

5.1 Selection of the Reference Atomic Geometry (RAG): the unit cell

The specific criteria for selecting the RAG are outlined in Sec. 1.2.1. Additionally to these characteristics, an important feature that must be incorporated in the selection of the RAG cell, which defines the simulation box for the model, is the designation of the RAG as a conventional cell. This simulation box is referred to as the *unit cell* in the MODELMAKER framework.

The *unit cell* represents the simulation box which serves as the framework for constructing the model and, in particular, the basis set of Wannier functions $\{\chi_a\}$. This basis set is employed to express the Hamiltonian matrix elements h_{ab}, where the subscripts a and b correspond to the Wannier functions χ_a and χ_b, respectively, in the chosen basis. The Wannier function χ_a is centered in the home *unit cell* while the χ_b Wannier function may either reside within the home *unit cell* or within one of its periodic replicas, as determined by the $\vec{k}$-mesh sampling utilized in WANNIER90, see Sec. 2.1.3.

When selecting the *unit cell* of the system in the RAG, it is crucial to ensure that it corresponds to a conventional, as this allows to properly incorporating the system's symmetry in the implementation (see Sec. 5.4). Since, as highlighted in the introduction of this Chapter, one of the key advancements in this improved methodology is the integration of symmetry within the code. In the MODELMAKER implementation, this symmetry is defined according to the space group of the system by the symmetry operations for its conventional cell.

As an example in the selection of the RAG *unit cell*, consider the $\mathrm{SrTiO_3}$ system in the paraelectric cubic geometry. In this case, the most natural choice for the *unit cell* would be the primitive cell, which is a cubic P cell. Here, both the primitive and conventional cells coincide. In contrast, for the solid system LiF with a rock salt structure, the primitive cell consists of two atoms, whereas the conventional cell is a cubic F cell composed by 8 atoms. Therefore, the *unit cell* for this system should be the cubic F cell.

Ultimately, it is important to underscore that although the model is constructed based on the designated *unit cell*—that determines the Wannier functions forming the basis set of the Hamiltonian matrix elements $h_{\boldsymbol{ab}}$— the model generation procedure also considers an additional cell, referred to as the *supercell* within the MODELMAKER framework. This *supercell* is a supercell of the *unit cell* and defines the simulation box employed in the first-principles calculations included in the training set for the computation of model parameters, which will be elaborated upon in subsequent sections.

5.2 Selection of the Reference Electronic Density (RED)

In the original work [37], the RED was defined using a geometry-independent, constant diagonal matrix,

$$d^{(0)}_{\boldsymbol{ab}} = o_{\boldsymbol{a}} \delta_{\boldsymbol{ab}}, \tag{5.1}$$

where $o_{\boldsymbol{a}}$ denotes the occupation number of the corresponding WF. This density matrix is periodic in the sense that $d^{(0)}_{\boldsymbol{ab}} = d^{(0)}_{\boldsymbol{a'b'}}$ if $(\boldsymbol{a},\boldsymbol{b}) \equiv (\boldsymbol{a'},\boldsymbol{b'})$ with "≡" meaning equivalent by translation. For non-magnetic insulators, this definition was shown [37] to accurately reproduce the ground state within the framework of DFT, regardless of the geometry. However, this approach becomes insufficient for metals and magnetic insulators, where many off-diagonal elements of the density matrix ($d_{\boldsymbol{ab}}$ with $\boldsymbol{a} \neq \boldsymbol{b}$) are non-zero in the ground state. Accurately capturing the ground state density matrix in such systems therefore requires a self-consistent approach, which can be both time-consuming and technically demanding.

In the new implementation of SPDFT, the reference density matrix is no longer constrained to be diagonal, although it remains constant with respect to the geometry. This generalization allows us to select the RED as the ground state density matrix of the RAG, accommodating both insulating and metallic systems.

For insulators, this choice remains valid across different geometries. In contrast, for metals—where partially filled bands evolve with the geometry—this results in variations of the ground state density matrix along the Born-Oppenheimer surface. Nonetheless, the current approach leads to smaller variations in the density matrix and enables a more accurate description of its geometry dependence, which is captured through electron-lattice interaction elements (see Sec. 3.5).

5.3 Band manifold selection and Wannierization procedure

The second-principles approach involves computing the matrix elements of the Kohn-Sham one-electron Hamiltonian within a Wannier function basis set $\{|\chi_a\rangle\}$ [69], as expressed in Eq. (4.37). The methodology for establishing the Wannier function basis from Bloch states is initiated by considering the system at the RAG (see Sec. 1.2.1) within the RED framework (selected in Sec. 5.2). The relationship between this basis and the Bloch state representation is elaborated in Sec. 1.4.

As discussed in Sec. 1.5, Wannier functions offer a highly adaptable representation of the electronic band structure, enabling the method to incorporate all chemically active electrons. By employing a Wannier function basis, it is possible to construct a tailored set of Wannier functions that span a selected group of bands, known as the *active set*. The flexibility to select an active set from the complete set of bands enhances computational efficiency. Then, constructing the Wannier function basis set requires a precise definition of the specific problem being addressed. In particular, if the problem under study involves the formation of hole polarons, the selected active set consists of the top of the valence band. Conversely, if the focus is on electron polarons, the bottom of the conduction band plays a fundamental role. In a different scenario, when investigating low-energy electron-hole excitons or optical properties, the active set includes both the top of the valence band and the bottom of the conduction band. Thus, the Wannier functions employed in the proposed methodology correspond to the bands forming the active set.

Once the desired collection of bands for computing the Wannier functions is selected, it is necessary to establish the procedure for obtaining the Wannier functions. The bands within the active set are organized into *manifolds*, as defined in Sec. 1.4.4.

By default, WANNIER90 [39; 40] performs the Wannierization procedure over a single manifold, typically defined by a global energy window. However, in the context of second-principles modeling, it is often desirable to construct WFs from multiple, physically distinct manifolds—for instance, separate treatment of valence and conduction bands, though the approach is not limited to this case.

While one might naively consider applying WANNIER90 independently to each manifold, this presents challenges when matrix elements connecting different manifolds are required. Although Hamiltonian matrix elements in the WF basis,

$$h_{ab} = \langle \chi_a | h | \chi_b \rangle \,, \tag{5.2}$$

vanish for WFs χ_a and χ_b belonging to different orthogonal manifolds. However, some phenomena require to obtain integrals involving two manifolds, like

$$\vec{r}_{ab} = \langle \chi_a | \vec{r} | \chi_b \rangle \,, \tag{5.3}$$

that plays the role of a dipolar transition matrix elements in optical transitions. To address this, we have extended the SIESTA code [174] to enable the simultaneous Wannierization of multiple independent manifolds via WANNIER90. The SIESTA code can now generate the input required for WANNIER90 to construct a combined auxiliary manifold containing all previously obtained WFs, enabling the evaluation of inter-manifold matrix elements such as those described in Eq. (5.3).

In addition, we have identified two practical issues that arise during Wannierization when applying the maximally localized algorithm to obtain the WFs (described in Sec. 1.4.6). First, small geometric perturbations can lead to discontinuous changes in WF shapes, introducing significant errors in electron-lattice coupling elements calculated via finite differences. Second, the initial projectors used to generate WFs may break symmetry when the spread minimization algorithm is employed, leading to less symmetric and less transferable WFs. We found that these issues can be mitigated by disabling the spread optimization in WANNIER90 thereby leading to the calculation of Wannier functions via projection (see Sec. 1.4.5), which results in WFs that are more symmetric and stable under structural changes. Furthermore, we observed improved localization and symmetry when using SIESTA's basis functions as initial projectors, rather than the default hydrogenoid functions typically used by WANNIER90.

5.4 Symmetrization of the parameters

A major limitation in earlier second-principles model construction approaches was the need to explicitly include all symmetry-equivalent configurations in the first-principles training set. On one hand, this led to an unnecessary increase in the number of parameters, as symmetry-related integrals were treated as independent. On the other hand, it required a significantly larger number of first-principles calculations to sample all symmetry-equivalent atomic configurations.

In the present work, we exploit the space group symmetries of the reference atomic geometry (RAG), obtained from the Bilbao Crystallographic Server [175; 176], and apply the corresponding symmetry operations T_o to the Wannier functions (WFs). These transformations yield relations among WFs of the form

$$T_o \left|\chi_{\boldsymbol{a}}\right\rangle = \sum_{\boldsymbol{a}'} T^o_{\boldsymbol{a}\boldsymbol{a}'} \left|\chi_{\boldsymbol{a}'}\right\rangle . \tag{5.4}$$

Here, $T^o_{\boldsymbol{a}\boldsymbol{a}'}$ is the symmetry-matrix element relating the Wannier functions $\boldsymbol{a}$ and $\boldsymbol{a}'$ through symmetry operation T_o. We apply this transformation to the relevant matrix elements that define the second-principles model, including those entering the total energy expression [Eq. (4.36)], the one-electron Hamiltonian [Eq. (4.37)], and the expected position operator [Eq. (5.3)]. In the case of the one-electron integrals, and considering that the symmetry operations are defined to leave the Hamiltonian invariant, we obtain

$$\begin{aligned} T_o\gamma_{\boldsymbol{ab}} &= T_o \left\langle\chi_{\boldsymbol{a}}\right| h(n_0) \left|\chi_{\boldsymbol{b}}\right\rangle = \left\langle T_o\chi_{\boldsymbol{a}}\right| h(n_0) \left|T_o\chi_{\boldsymbol{b}}\right\rangle = \\ &= \sum_{\boldsymbol{a}'\boldsymbol{b}'} T^{o,T}_{\boldsymbol{a}\boldsymbol{a}'} T^o_{\boldsymbol{b}\boldsymbol{b}'} \left\langle\chi_{\boldsymbol{a}'}\right| h(n_0) \left|\chi_{\boldsymbol{b}'}\right\rangle = \sum_{\boldsymbol{a}'\boldsymbol{b}'} T^{o,T}_{\boldsymbol{a}\boldsymbol{a}'} T^o_{\boldsymbol{b}\boldsymbol{b}'} \gamma_{\boldsymbol{a}'\boldsymbol{b}'} = \gamma_{\boldsymbol{ab}}. \end{aligned} \tag{5.5}$$

where the T superindex indicates transpose. The set of γ_{ab} elements that are related through nonzero coefficients $T^{o,T}_{aa'}T^{o}_{bb'}$ are said to form a group of symmetry-equivalent parameters.

Similarly, for the linear electron-lattice coupling constants, we obtain

$$T_o\vec{f}_{ab,\lambda v} = \vec{f}'_{ab,\lambda v} = \sum_{a'b'} T^{o,T}_{aa'} T^{o}_{bb'} \vec{f}_{a'b',\lambda' v'} \tag{5.6}$$

and for the quadratic electron-lattice couplings

$$T_o \overset{\leftrightarrow}{g}_{ab,\lambda v} = \overset{\leftrightarrow}{g}'_{ab,\lambda v} = \sum_{a'b'} T^{o,T}_{aa'} T^{o}_{bb'} \overset{\leftrightarrow}{g}_{a'b',\lambda' v'}. \tag{5.7}$$

For the electron-electron interaction terms, the symmetry relations become

$$T_o U_{ab,a'b'} = U_{ab,a'b'} = \sum_{a''b''}\sum_{a'''b'''} T^{o,T}_{aa''} T^{o}_{bb''} T^{o,T}_{a'a'''} T^{o}_{b'b'''} U_{a''b'',a'''b'''} \tag{5.8}$$

and

$$T_o I_{ab,a'b'} = I_{ab,a'b'} = \sum_{a''b''}\sum_{a'''b'''} T^{o,T}_{aa''} T^{o}_{bb''} T^{o,T}_{a'a'''} T^{o}_{b'b'''} I_{a''b'',a'''b'''}. \tag{5.9}$$

We call Eqs. (5.5)-(5.9) symmetry constraints (S_C), for reasons that will become evident in Sec. 5.6.2.

In order to obtain $T^{o}_{aa'}$ we assume that the WFs transform like the SIESTA basis orbitals that were used as projectors during the Wannierization procedure, Sec. 5.3. Comparing the value of the h_{ab} elements obtained through WANNIER90 we can self-check whether the Wannierization scheme applied in any practical calculation leads to correct symmetrization.

Another important use of symmetry is the reduction in the number of calculations needed to create a training set to determine the value of the various integrals needed in a second-principles model. In the case of calculations where the geometry changes, like those needed to calculate the electron-lattice coupling, see Sec. 3.5, a distortion with respect to the RAG, realized in a particular geometry of a *supercell*, is described by the vectors, $\vec{u}_\lambda$ [see Eq. (1.2)]. Calculations that involve the distortions

$$\vec{u}_{\lambda'} = T_o \vec{u}_\lambda \tag{5.10}$$

do not contain any new information with respect to the carried for $\vec{u}_\lambda$. Using the DFT calculation at the geometry given by $\vec{u}_\lambda$ and applying Eq. (5.5) it is possible to obtain the values of h^{DFT}_{ab} at $\vec{u}_{\lambda'}$ by simply operating,

$$h^{\mathrm{DFT}}_{a'b'}(\vec{u}_{\lambda'}) = T_o^{-1} h^{\mathrm{DFT}}_{a'b'}(\vec{u}_{\lambda'}) = \sum_{ab} T^{o,-1}_{aa'} T^{o,-1}_{b'b} h^{\mathrm{DFT}}_{ab}(\vec{u}_\lambda). \tag{5.11}$$

This procedure allows us reducing the number of calculations to obtain the electron-lattice matrix elements, as we will see later.

5.5 Goal function

To assess the accuracy of the model relative to first-principles simulations, it is essential to establish a robust metric that quantifies the discrepancy introduced by the model, defined as *goal function*. Specifically, we define this metric in terms of the Hamiltonian matrix elements expressed in a Wannier functions basis set, which serve as fundamental quantities for evaluating model performance.

For a given calculation A, characterized by a specific geometric configuration and electronic state, the proposed metric quantifies the quadratic error between the Hamiltonian matrix element computed using the second-principles approach, $h_{\boldsymbol{ab}}(A)$, and the corresponding result obtained from first-principles simulations, $h_{\boldsymbol{ab}}^{\mathrm{DFT}}(A)$. Mathematically, this metric serves as a measure of the deviation of the second-principles model from the benchmark provided by density functional theory (DFT) calculations.

The first-principles matrix elements $h_{\boldsymbol{ab}}^{\mathrm{DFT}}$ are obtained using WANNIER90 code and represent the interaction between the Wannier functions $\chi_{\boldsymbol{a}}$ and $\chi_{\boldsymbol{b}}$. These values serve as a reference for model validation.

As outlined in Eq. (4.37), the second-principles Hamiltonian matrix elements depend on a collection of model parameters, encompassing both electron-lattice interactions and electron-electron contributions. These matrix elements can thus be expressed as functions of a parameter set, $\{p_i\}_{1\leq i\leq P}$, encapsulating both collections of parameters. As a result, the second-principles Hamiltonian matrix element $h_{\boldsymbol{ab}}$ for a defined calculation A and a collection of parameters is expressed as $h_{\boldsymbol{ab}}(A,\{p_i\})$.

To comprehensively evaluate the cumulative error introduced by the model, the defined metric systematically accounts for all pairs of Wannier functions, $\chi_{\boldsymbol{a}}$ and $\chi_{\boldsymbol{b}}$. Moreover, in cases where the error needs to be assessed over a collection of calculations $\{A\}$ rather than a single configuration, the goal function is extended to aggregate errors across all calculations in the dataset.

As a result, the expression of the goal function, Θ, formulated to quantify the quadratic error in the second-principles model, is expressed as follows,

$$\Theta(\{A\},\{p_i\}) = \sum_{A}\sum_{\boldsymbol{ab}} \left[h_{\boldsymbol{ab}}(A,\{p_i\}) - h_{\boldsymbol{ab}}^{\mathrm{DFT}}(A)\right]^2 . \tag{5.12}$$

By defining this goal function, we establish a rigorous framework for quantifying the accuracy of second-principles models. Finally, this goal function serves as a fundamental tool in parameter optimization schemes, where minimizing Θ facilitates the optimal selection of the parameter set for the second-principles model. In fact, this approach represents one of the two methodologies applied to determine the model parameters. A detailed explanation of this technique will be provided in following sections.

5.6 Obtaining parameters

The proposed second-principles DFT methodology facilitates the simulation of large systems by considering extensive supercells under operating conditions, such as finite temperatures or externally applied electric fields. In this framework, the ability to conduct the desired simulations relies on the knowledge of a set of parameters defined in Sec. 4.2 that characterize one-electron interactions which include electron-lattice coupling as well as two-electron interactions. At this point, the critical question to address is the appropriate method for determining these parameters. We propose two alternative approaches for computing the model parameters: direct calculation from first-principles simulations and a fitting procedure. Each method has its own advantages and limitations, with the choice depending on computational feasibility and the required level of accuracy.

Previously, we presented the derivation of the mathematical expressions for each system parameter, as outlined in [37]. Consequently, by applying Eq. (4.22) for the one-electron parameters and Eq. (4.33) for the two-electron parameters, it is theoretically possible to obtain their numerical values with the prior knowledge of the expressions for the Wannier functions used as a basis set, the one-electron Hamiltonian, and the specific integral operators. To ensure predictive accuracy, these parameters should ideally be determined through first-principles calculations. However, existing first-principles approaches tend to be overly restrictive, primarily focusing on single-center, strongly correlated electron systems rather than the broader multicenter interactions considered in this study. This limitation is particularly important in the case of electron-electron interaction parameters. The large number of relevant integrals in a typical system, especially complex four-index integrals, results in a substantial computational burden, making direct first-principles calculations for electron-electron parameters computationally demanding.

An alternative strategy for obtaining model parameters is to employ a fitting procedure. The objective of this approach is to achieve accurate results while significantly improving computational efficiency. This efficiency is achieved by addressing two main aspects. First, as previously discussed, the calculation of certain parameters, in particular those associated with four-center integrals (electron-electron parameters) is computationally demanding. The fitting procedure circumvents this challenge by estimating these parameters using a reduced set of first-principles calculations instead of directly computing them for the entire system. Second, computational efficiency in large-scale second-principles simulations can be improved by building smaller models that select only the most relevant degrees of freedom. An advanced fitting procedure can systematically identify and prioritize the parameters most critical to the problem, ensuring that the resulting models remain computationally manageable without sacrificing accuracy. For instance, in the case of $SrTiO_3$ system, electron-lattice interaction parameters associated with the Sr atom are found to be negligible, whereas those related to Ti and O atoms play a dominant role. By focusing on these relevant parameters, the model becomes significantly more efficient while retaining its predictive power.

Nevertheless, the success of the fitting approach depends critically on the selection of an appropriate collection of first-principles simulations as the foundation for parameter estimation. These calculations must be designed to capture the essential physics of the system, enabling the identification of relevant parameters and the accurate extraction of their values. Without a well-constructed collection of first-principles calculations, the fitting procedure may overlook key interactions, leading to suboptimal model performance.

Both approaches, direct parameter calculation from first-principles simulations and the fitting procedure, present viable alternatives for obtaining the parameters required in second-principles simulations. Direct calculation provides the most accurate parameter values, but it is computationally expensive, especially for large systems. On the other hand, the fitting procedure offers greater computational efficiency by selecting the most relevant parameters, although it does not achieve the same level of precision as direct calculation. The appropriate method is selected by the necessary trade-off between accuracy and efficiency for the given problem.

Supercell for the training set calculations

The two methodologies described previously require a set of single-point first-principles calculations to determine the model parameters. Regardless of the selected approach, this collection of calculations is referred to as the *training set*.

A crucial initial decision in the construction of the training set is the selection of the simulation box where the first-principles calculations are going to be performed. This simulation box can be the *unit cell* (defined in Sec. 5.1) or a cell composed by periodic replicas of the *unit cell*, leading to a supercell of the *unit cell*. In the MODELMAKER code, this simulation box is referred to as the *supercell*.

The role of this supercell in the model construction is crucial. Although the model Hamiltonian matrix elements $h_{\boldsymbol{ab}}$ [Eq. (4.37)] are defined for the conventional *unit cell*—where the Wannier function $\chi_{\boldsymbol{a}}$ is centered within the *unit cell*—the electron-lattice parameters incorporate the contributions of atoms within the *supercell*. Likewise, the electron-electron interaction terms involve Wannier functions $\chi_{\boldsymbol{a}'}$ and $\chi_{\boldsymbol{b}'}$, which also are defined in the *supercell*. It is important to note that the training set can typically be restricted to atomic and electronic configurations that are compatible with small *supercells*. This aligns with the fact that, when expressed in a basis of localized Wannier functions, the non-electrostatic interactions in most materials are short-ranged. Next, we outline the key reasons why the construction of a supercell is necessary.

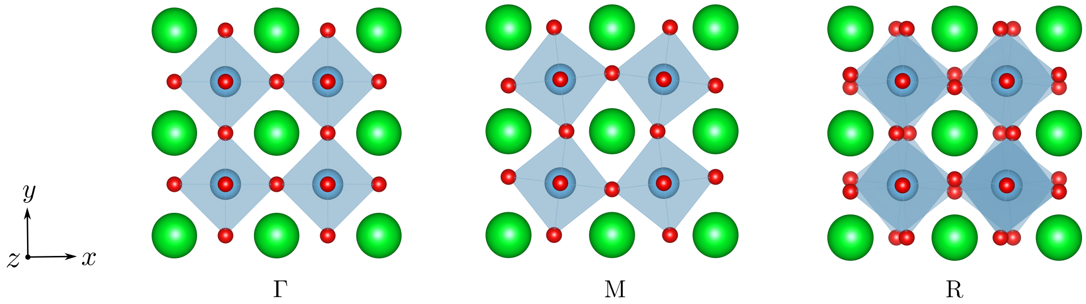


Figure 5.1: Geometric phases of a perovskite associated with structural distortions characterized by phonon modes at the Γ, M, and R points of the Brillouin zone. The Γ mode represents the perovskite in its cubic, centrosymmetric phase. The M mode corresponds to antiphase rotations of the octahedra within the plane. The R mode involves antiphase rotations of the octahedra along all three Cartesian directions.

Firstly, several systems exhibit instabilities in different phases, including those involving phonon modes characterized by $\vec{k} \neq \Gamma$. These phases often require larger supercells to capture atomic distortions that cannot be accommodated within the *unit cell*. Since the resulting models must accurately simulate the system's characteristic phases, the choice of supercell size depends on the phases under study and must be selected to properly account for these structural modulations.

For example, in the case of ABO_3 systems, beyond their ferroelectric properties [177], ABO_3 perovskites exhibit a broader range of structural phase transitions. In particular, they frequently undergo non-polar structural distortions associated with different tilts of the oxygen octahedra. These distortions, which correspond to finite wave vectors $\vec{k} \neq \Gamma$, require larger simulation cells. To accurately capture such structural instabilities, the supercell size must be chosen to match the periodicity dictated by the phonon wave vector.

For example, the antiferrodistortive phase $\vec{k} = M$ characterized by in-phase octahedra rotations along the z-axis and antiphase rotations in the plane, see Fig. 5.1, requires a supercell with 20 atoms, while antiferrodistortive phase at the $\vec{k} = R$ point with antiphase rotations along different axes, see Fig. 5.1, requires a 40-atom supercell. This phase is particularly relevant in $SrTiO_3$. In the specific case of $SrTiO_3$, the material undergoes a phase transition from a paraelectric cubic phase to a non-polar antiferrodistortive (AFD) tetragonal phase ($I4/mcm$, No. 140) at a critical temperature $T_C \approx 105$ K [178]. This AFD phase emerges due to rotations of the oxygen octahedra around the tetragonal axis, preserving the overall non-polar character of the crystal.

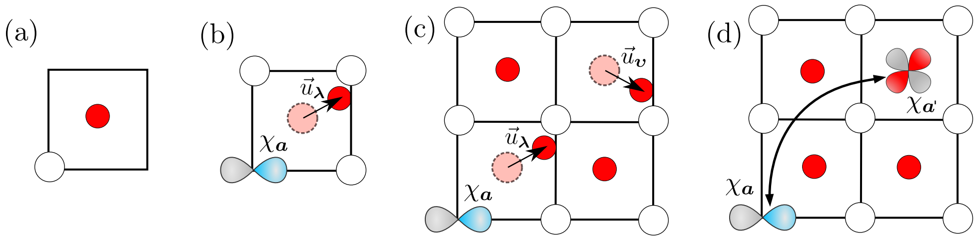


Figure 5.2: a) The simulation box selected as the *unit cell* for constructing the second-principles model using the MODELMAKER code. (b) Displacement of the atom $\boldsymbol{\lambda}$ relative to the Wannier function $\chi_{\boldsymbol{a}}$ within the *unit cell* leading to electron-lattice coupling quantified by the electron-lattice parameter $\vec{f}_{\boldsymbol{aa},\boldsymbol{\lambda}}$. (c) Displacements of atoms $\boldsymbol{\lambda}$ and $\boldsymbol{\upsilon}$ relative to the Wannier function $\chi_{\boldsymbol{a}}$ within a 2×2 supercell, leading to electron-lattice coupling characterized by the electron-lattice parameter $\overset{\leftrightarrow}{g}_{\boldsymbol{aa},\boldsymbol{\lambda\upsilon}}$. (d) Interaction between Wannier functions $\chi_{\boldsymbol{a}}$ and $\chi_{\boldsymbol{a}'}$ in a 2×2 supercell, resulting in the formation of electron-electron interaction parameters $U_{\boldsymbol{aa},\boldsymbol{a}'\boldsymbol{a}'}$ and $I_{\boldsymbol{aa},\boldsymbol{a}'\boldsymbol{a}'}$.

When selecting the *supercell*, it is essential not only to consider the specific phases to be simulated, but also to ensure that the *supercell* is sufficiently large to accurately capture atomic displacements and their effects on the system. To demonstrate this, we consider a two-dimensional schematic system where the primitive cell (which coincides with the conventional cell) is depicted in Fig. 5.2(a) and is defined by a motif consisting of two atoms. The motif comprises a white atom located at (0, 0) and a red atom positioned at (0.5, 0.5) in fractional coordinates. Now, a Wannier function centered on the white atom is selected, characterized by a p-orbital symmetry.

First, we define the *unit cell* of the system as the conventional cell, following the guidelines outlined in Sec. 5.1.

Now, if we consider the *unit cell* itself as the *supercell*, it is possible to account for the effect of the displacement of atom $\boldsymbol{\lambda}$, illustrated in Fig. 5.2(b), on the on-site interaction of Wannier function $\chi_{\boldsymbol{a}}$.

However, if we aim to consider the displacement of the atom $\boldsymbol{\upsilon}$ as depicted in Fig. 5.2(c) (which is not sufficiently distant to disregard its effect on the Wannier function $\chi_{\boldsymbol{a}}$), it is not feasible. In this case, it would not be possible to define $\vec{f}_{\boldsymbol{aa},\boldsymbol{\upsilon}}$ [expressed in accordance with Eq. (3.121)].

Analogously, if instead of considering the displacement of a single atom, we aim to assess the effect of the simultaneous displacement of both $\boldsymbol{\lambda}$ and $\boldsymbol{\upsilon}$ atoms, represented in Fig. 5.2(c), selecting the *unit cell* as the *supercell* is not feasible. At a minimum, a *supercell* of size 2×2 is required. For instance, the quadratic parameter $\overset{\leftrightarrow}{g}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}}$ could not be defined in this context.

A similar rationale applies to the interaction between Wannier functions in the consideration of electron-electron interactions. If the cell is not sufficiently large, certain interactions between Wannier functions, as the one depicted in Fig. 5.2(d) between $\chi_{\boldsymbol{a}}$ and $\chi_{\boldsymbol{a}'}$ leading to the $U_{\boldsymbol{aa},\boldsymbol{a}'\boldsymbol{a}'}$ may be omitted from the model, effectively being approximated as zero. By employing a larger *supercell*, these interactions can be properly accounted for, thus enhancing the accuracy of the model.

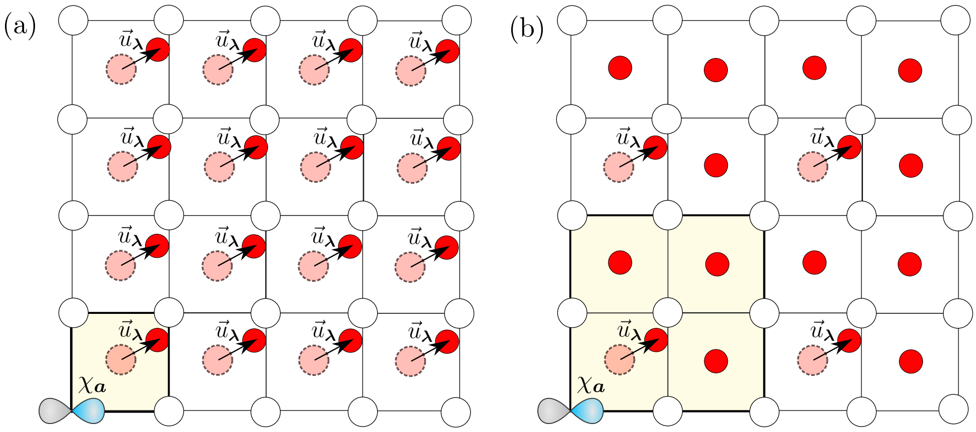


Figure 5.3: Displacement of the atom $\boldsymbol{\lambda}$ and its periodic replicas when: (a) The selected *supercell* (shaded in yellow) is identical to the *unit cell*, i.e., the conventional cell of the system. (b) The selected *supercell* (shaded in yellow) is a 2×2 supercell of the *unit cell*, comprising four conventional cells of the system.

Secondly, if the selected *supercell* is too small, the displacement of an atom and its periodic images can lead to artificial interactions, resulting in erroneous or misleading results. For instance, considering the previous system, we select as *supercell* the primitive cell, shaded in yellow in Fig. 5.3(a). Now, we aim to assess the effect of displacing the atom $\boldsymbol{\lambda}$ on the on-site interaction of the Wannier function $\chi_{\boldsymbol{a}}$. Owing to the implementation of periodic boundary conditions in the simulation, the displacement of atom $\boldsymbol{\lambda}$ simultaneously results in the displacement of all its periodic replicas, as illustrated in Fig. 5.3(a). Consequently, in a small simulation box, when we displace the atom $\boldsymbol{\lambda}$ and evaluate the effect of its displacement on the Wannier function, we are effectively measuring the influence not only of the atom $\boldsymbol{\lambda}$ but also of its periodic replicas. However, by considering a sufficiently large *supercell* for building the training set, in this case, a 2×2 supercell as indicated by the yellow-shaded region in Fig. 5.3(b), displacing the atom $\boldsymbol{\lambda}$ ensures that its periodic images are adequately spaced. Given the inherently localized nature of the Wannier functions, the impact of periodic images of atom $\boldsymbol{\lambda}$ on the Wannier function becomes negligible, thereby mitigating artificial interactions and enhancing the accuracy of the calculations.

In particular, for simulations involving polarons, training the system on large enough cells is essential to avoid the interaction between the defect and its periodic images.

For the $\mathrm{SrTiO_3}$ system studied in this work, a $2 \times 2 \times 2$ *supercell* of the primitive cell has been selected. The use of a 222 supercell for the training set can be justified in light of Machine Learning lattice models showing that, for phonon band structures, such a cell already captures the essential lattice dynamics even for polar materials, including ferroelectric instabilities and its chain-like character, that have been typically though as long-range effects. While longitudinal optical modes are not accurately reproduced, these models show a very good performance in describing the phase diagram of polar materials even when trained on relatively short supercells [179]. These findings, though focused on lattice models, hint that the nearsightedness principle might also extend to our case and suggest that a similarly sized cell could suffice for a qualitatively correct description of the electronic structure. However, further investigation would be required to confirm this in the electronic context.

After determining the size of the *supercell*, i.e., the simulation box used for first-principles calculations to construct the training set, the subsequent step involves the careful selection of atomic and electronic configurations that will comprise the training set. It is crucial to ensure the inclusion of all necessary calculations to comprehensively capture the physical characteristics of the system. These configurations play a fundamental role in the identification and computation of the model parameters.

For instance, if the electronic bands of the system under investigation exhibit pronounced sensitivity to structural distortions, the training set should encompass calculations for various geometric configurations to adequately capture this dependence. As relevant example is systems that exhibit different phases, such as those previously described for $\mathrm{ABO_3}$ perovskites, where atomic distortions influence the band structure, leading to breaking of degeneracies and changes in band curvature.

Similarly, when examining the effects of electron or hole doping on the properties of a material, the training set should incorporate DFT simulations of the system in a charged state to accurately capture these influences. In the case of magnetic systems, it is crucial for the training set to include simulations with various spin configurations, as this will enable the accurate capture of the mechanisms underlying magnetic interactions.

After introducing the methodologies available for computing model parameters and highlighting the necessity of a training set, we now provide a detailed examination of these approaches. We first introduce the parameter fitting methodology previously developed in SCALE-UP [37] to provide context for the existing approach. We will refer to this procedure as *previous methodology*. The present work aims to improve upon it by developing an alternative approach for obtaining model parameters using MODELMAKER. Accordingly, two methodologies have been explored. The first approach, *methodology I*, involves an advanced parameter fitting technique, which, compared to direct parameter computation, offers the advantage of incorporating only the most relevant degrees of freedom while requiring a training set with minimal computational cost. Despite its advantages, challenges encountered in the implementation of this technique required the development of a second methodology. This alternative approach, *methodology II*, is based on the direct calculation of electron-lattice parameters, providing a more explicit and efficient means of parameter determination.

5.6.1 Previous methodology

In this Section, we present the methodology employed in Ref. [37] for the calculation of the parameters that define the model. Here, the objective is to establish the initial state of the approach, highlight its limitations, and identify the challenges that have been addressed in this work.

The original methodology consisted of several sequential steps which are detailed in the subsequent sections. First, the following section presents the employed training set, which comprised *representative atomic and electronic configurations*. These configurations served as the basis for computing the model parameters through a fitting procedure. Once the training set was generated, the next step involved filtering the pairs of Wannier functions, denoted as $\{(\boldsymbol{a},\boldsymbol{b})\}$, that would be considered for constructing the Hamiltonian matrix elements $h_{\boldsymbol{ab}}$. Given the large number of potential parameters, the subsequent step focused on identifying the most relevant ones to be included in the model. This step was crucial in determining the simplest model capable of accurately reproducing the first-principles calculations within the training set. Additionally, this stage involved defining the matrices required for parameter fitting, procedure which constituted the final step of the methodology.

In the concluding part we address the limitations of this approach, providing the necessary context to introduce the two developed methodologies proposed in this study.

5.6.1.1 Training set

Electron-electron training set

The construction of the training set for the calculation of electron-electron parameters was carried out changing the electronic configuration of the system from the RED considering the underlying physics of the compound. The training set was system-dependent, composed by *representative configurations* and there was no established procedure to compute it in a systematic way. For instance, the training set for a magnetic system included first-principles calculations of several spin arrangements to accurately capture the magnetic couplings. In the case of an optically active insulator system, the first-principles calculations had to be associated with excitations from the valence band to the conduction band. Below we show both examples.

As an illustrative example of a magnetic compound, we consider the NiO system with a rock-salt crystal structure. This system belongs to the class of transition metal oxides and exhibits magnetic properties. NiO is a correlated electron system with many competing magnetic phases. Specifically, three magnetic phases are separated by energy differences on the order of tenths of electron-volts per formula unit.

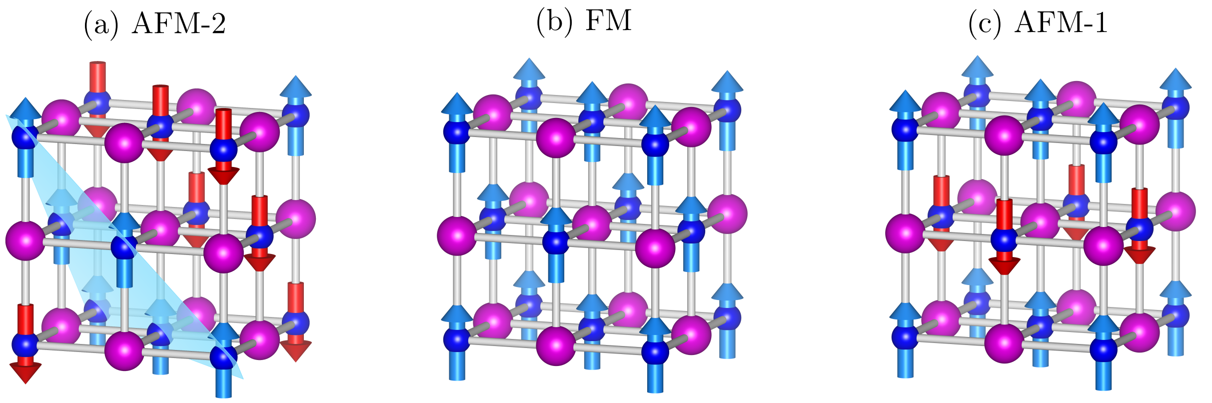


Figure 5.4: Training set used for the calculation of electron–electron interaction parameters in the magnetic NiO system, modeled in its experimentally observed cubic phase. Nickel (Ni) atoms are depicted as blue spheres, while oxygen (O) atoms are shown as magenta spheres. The first-principles simulations in the training set involve altering the magnetization of Ni^{2+} ions by imposing various spin configurations to model the magnetic ordering of the system. The magnetic orderings illustrated include: antiferromagnetic type 2 (AFM-2), in panel (a), characterized by alternating spin polarity across ferromagnetic [111] diagonal planes; ferromagnetic (FM) in panel (b); antiferromagnetic type 1 (AFM-1) in panel (c), similar to AFM-2 but with alternating planes oriented along the [001] direction.

Regarding the spin order, the ground state of NiO is the AFM-2 phase, characterized by alternating planes of spin-up and spin-down polarized nickel atoms along the [111] direction [Fig. 5.4(a)]. However, this crystal structure also allows for other spin arrangements that compete with the AFM-2 phase, such as the ferromagnetic phase [Fig. 5.4(b)] and the AFM-1 phase [Fig. 5.4(c)]. The latter is characterized by alternating planes of spin-up and spin-down polarized nickel atoms along the [001] direction. Due to the small energy differences between these magnetic phases–in particular the antiferromagnetic phase AFM-2 is more stable than the ferromagnetic phase by 89 meV per formula unit (f.u.)–it was necessary to study all three in the fitting of parameters. Consequently, in the training set related to electron-electron interactions, the NiO compound had to include these three spin arrangements with different D_{aa} [Eq. (4.10)] occupations to accurately represent the competing magnetic phases of the system. A more comprehensive study of this example concerning the electron-electron training set and parameters is provided in Ref. [37].

If, instead, the system under study were hBN crystal, an insulator whose optical spectrum exhibits highly localized excitons [180] as shown in Fig. 5.5, the training set selection criteria would differ. The optical spectrum of hBN crystal reveals peaks at energies lower than the bandgap energy. These peaks had been attributed to interactions between holes in the top of the valence band and electrons in the bottom of the conduction band leading to the formation of excitons. These interactions are generated through optical absorption, leading to the appearance of additional spectral lines not predicted by conventional band theory and associated to the excitonic interaction. In SCALE-UP, optical excitations correspond to changes in the electronic configuration: during an absorption event from the valence band to the conduction band, the system transitions from its RED to a new electronic configuration. Consequently, to accurately capture these effects, the training set for the hBN compound had to include calculations that accounted for changes in the electronic configuration in the form of excitations from the valence band to the conduction band, without introducing distortions in the RAG.

To summarize, the selection of the training set for computing electron-electron parameters was inherently system-dependent and had to be performed manually. The methodology for determining the training set was user-defined, adapting to the specific characteristics of each system under investigation. In particular, for the two systems we considered, while in the case of NiO calculations changing the spin arrangement played a crucial role, in the case of hBN crystal the calculations were characterized by first-principles simulations where excitations were essential.

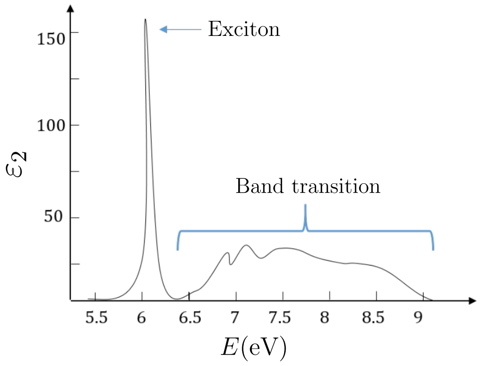


Figure 5.5: Absorption spectrum of the hBN crystal. A strong absorption is observed around 6.1 eV, indicating the presence of an exciton and an interband transition. Figure adapted from [180].

Electron-lattice training set

In the *previous methodology*, the training set for fitting the electron-lattice parameters was composed by first-principles calculations in which the atomic geometry was systematically distorted relative to the RAG while maintaining the system within the designated RED.

The calculations comprising the training set were characterized by the displacement of individual atoms along a single spatial component relative to their positions in the RAG, while the remaining atoms retained their original positions. Consequently, for each calculation in the training set related to the electron-lattice contribution, denoted as A, there existed a unique component displacement satisfying $u_{\lambda\alpha}(A) \neq 0$, where α designated a specific Cartesian direction.

Due to the absence of a symmetry implementation, the training set for the electron-lattice interaction was constructed using a large number of first-principles calculations to explicitly account for the system's symmetry.

5.6.1.2 Selection of one-electron Hamiltonian

Once the electron-electron and electron-lattice training sets were generated, the methodology applied a selection criterion of Hamiltonian matrix elements to reduce computational cost. Specifically, only pairs of Wannier functions $(\boldsymbol{a}, \boldsymbol{b})$ whose interaction term $h^{s}_{\boldsymbol{ab}}(A)$ obtained from a calculation A in the training set exceeded a predefined energy cutoff $\delta\varepsilon_{\mathrm{h}}$, were considered. This cutoff value was user-defined. Explicitly, the pair $(\boldsymbol{a}, \boldsymbol{b})$ was selected if

$$|h^{s}_{\boldsymbol{ab}}(A)| > \delta\varepsilon_{\mathrm{h}}, \quad \text{for at least one } A \text{ in the TS.} \tag{5.13}$$

For diagonal elements $h^{s}_{\boldsymbol{aa}}$, however, all terms were always included regardless of their magnitude in first-principles calculations.

The selection of a Wannier function pair $(\boldsymbol{a}, \boldsymbol{b})$ implied that, in constructing the parametric Hamiltonian $h^{s}_{\boldsymbol{ab}}$ [Eq. (4.37)], at least the parameters associated with the RAG, $\gamma^{\mathrm{sr,RAG}}_{\boldsymbol{ab}}$, were included in the Hamiltonian matrix elements.

This cutoff criterion has been consistently applied in the two novel methodologies proposed and analyzed in this work, *methodology I* and *methodology II*, in the form of a distance cutoff, as will be discussed in following sections.

Returning to the *previous methodology*, subsequently, the electron-electron and electron-lattice parameters were selected and generated for a subset of Wannier function pairs chosen from those retained in this step.

5.6.1.3 Selection of electron-electron parameters

In the *previous methodology*, the electron-electron parameters were determined by ensuring that the constructed model accurately reproduced the Hamiltonian matrix elements of the electron-electron training set, $h^{s}_{\boldsymbol{ab}}(A)$.

As described above, the training set used for fitting the electron-electron parameters consisted of configurations in which the RED of the system was modified while maintaining the atomic positions fixed in the RAG. Knowing the training set, the first step involved selecting the most significant electron-electron parameters. To accurately determine these parameters, it was necessary to isolate their contributions. For this purpose, the Hamiltonian matrix elements $h^U_{\boldsymbol{ab}}$ and $h^I_{\boldsymbol{ab}}$, depending on U and I respectively, were defined according to the following expressions.

Firstly, in order to eliminate dependence on I, the matrix element $h^U_{\boldsymbol{ab}}$ was defined as

$$h^U_{\boldsymbol{ab}}(A) = \frac{h^{\uparrow}_{\boldsymbol{ab}}(A) + h^{\downarrow}_{\boldsymbol{ab}}(A)}{2} = \gamma_{\boldsymbol{ab}} + \sum_{\boldsymbol{a'b'}} D^U_{\boldsymbol{a'b'}}(A) U_{\boldsymbol{ab},\boldsymbol{a'b'}}. \tag{5.14}$$

Although this expression removed the dependence on I, the one-electron parameters $\gamma_{\boldsymbol{ab}}$ remained present. To extract this dependence, the quantity $\bar{h}^U_{\boldsymbol{ab}}$ was introduced representing the average of the Hamiltonian matrix element $h^U_{\boldsymbol{ab}}$ over the electron-electron training set (i.e., over the RAG configurations in the training set, conforming a total of N_{RAG} simulations),

$$\bar{h}^U_{\boldsymbol{ab}} = \frac{1}{N_{\mathrm{RAG}}} \sum_A h^U_{\boldsymbol{ab}}(A) = \gamma_{\boldsymbol{ab}} + \sum_{\boldsymbol{a'b'}} \bar{D}^U_{\boldsymbol{a'b'}} U_{\boldsymbol{ab},\boldsymbol{a'b'}}. \tag{5.15}$$

The deviation of $h^U_{\boldsymbol{ab}}(A)$ from its average value was then quantified as

$$\left| h^U_{\boldsymbol{ab}}(A) - \bar{h}^U_{\boldsymbol{ab}} \right| = \left| \sum_{\boldsymbol{a'b'}} \left[D^U_{\boldsymbol{a'b'}}(A) - \bar{D}^U_{\boldsymbol{a'b'}} \right] U_{\boldsymbol{ab},\boldsymbol{a'b'}} \right|, \tag{5.16}$$

depending exclusively on the U parameter. Similarly to Eq. (5.15), the Hamiltonian matrix element $h^I_{\boldsymbol{ab}}$, incorporating only the I parameters was defined as

$$h^I_{\boldsymbol{ab}}(A) = \frac{h^{\uparrow}_{\boldsymbol{ab}}(A) - h^{\downarrow}_{\boldsymbol{ab}}(A)}{2} = \sum_{\boldsymbol{a'b'}} D^I_{\boldsymbol{a'b'}}(A) I_{\boldsymbol{ab},\boldsymbol{a'b'}}. \tag{5.17}$$

It was expected that the most significant U and I parameters corresponded to those associated with pairs of Wannier functions whose Hamiltonian matrix elements $h^U_{\boldsymbol{ab}}$ and $h^I_{\boldsymbol{ab}}$ exhibited the strongest dependence on the training set. Consequently, the Wannier function pairs contributing to electron-electron interactions were selected if they satisfied, for at least one training set configuration, either of the following conditions,

$$\left| h^U_{\boldsymbol{ab}}(A) - \bar{h}^U_{\boldsymbol{ab}} \right| = \left| \sum_{\boldsymbol{a'b'}} \left[D^U_{\boldsymbol{a'b'}}(A) - \bar{D}^U_{\boldsymbol{a'b'}} \right] U_{\boldsymbol{ab},\boldsymbol{a'b'}} \right| > \delta\varepsilon_{\mathrm{ee}}, \tag{5.18}$$

or

$$\left| h^I_{\boldsymbol{ab}}(A) \right| = \left| \sum_{\boldsymbol{a'b'}} D^I_{\boldsymbol{a'b'}}(A) I_{\boldsymbol{ab},\boldsymbol{a'b'}} \right| > \delta\varepsilon_{\mathrm{ee}}, \tag{5.19}$$

where $\delta\varepsilon_{\mathrm{ee}}$ represented an energy cutoff fixed by the user. Given that applying only $\delta\varepsilon_{\mathrm{ee}}$ did not sufficiently reduce the number of potential parameters in certain cases, a second cutoff was introduced to further refine the list of potentially significant U and I parameters for the fitting process. Then, once the pairs $\{(\boldsymbol{a},\boldsymbol{b})\}$ that met the criteria in Eqs. (5.18)-(5.19) were identified, this additional cutoff was applied to limit the selection. The new cutoff was enforced to the *difference density matrix*, ensuring that a variable $U_{\boldsymbol{ab},\boldsymbol{a}'\boldsymbol{b}'}$ or $I_{\boldsymbol{ab},\boldsymbol{a}'\boldsymbol{b}'}$ was included only if the pair of Wannier functions $(\boldsymbol{a}',\boldsymbol{b}')$ satisfied the criterion for one calculation in the electron-electron training set. Specifically, for the U variables,

$$|D^{U}_{\boldsymbol{a}'\boldsymbol{b}'}(A) - \bar{D}^{U}_{\boldsymbol{ab}}| > \delta D, \tag{5.20}$$

and, analogously, for the I variables,

$$|D^{I}_{\boldsymbol{a}'\boldsymbol{b}'}(A)| > \delta D. \tag{5.21}$$

5.6.1.4 Selection of electron-lattice parameters

In the theoretical framework of SCALE-UP, summarized in Chapter 4, electron-lattice interactions are considered in $h^{s}_{\boldsymbol{ab}}$ exclusively through the one-electron parameters $\gamma_{\boldsymbol{ab}}$ and do not influence the two-electron contribution $\delta h^{s,\mathrm{ee}}_{\boldsymbol{ab}}$ [Eq. (4.32)]. In the short-range contribution of those one-electron parameters, $\gamma^{\mathrm{sr}}_{\boldsymbol{ab}}$ [Eq. 4.24], the dependence of the Hamiltonian matrix elements with the geometry is included through the incorporation of the electron-lattice parameters $\vec{f}$ and $\overset{\leftrightarrow}{g}$.

Thus, in the *previous methodology*, in order to select the electron-lattice parameters which play a crucial role in the system, the electron-lattice coupling was quantified by measuring the deviations of the one-electron parameters $\gamma_{\boldsymbol{ab}}$ of distorted geometries belonging to the training set from their corresponding values in the RAG within the RED, denoted as $\gamma^{\mathrm{sr,RAG}}_{\boldsymbol{ab}}$. To systematically evaluate these deviations, for each calculation in the electron-lattice training set, labeled as A, and characterized by the displacement $u_{\boldsymbol{\lambda}\alpha}(A)$ of atom $\boldsymbol{\lambda}$ along the Cartesian direction α, each Hamiltonian matrix element $h_{\boldsymbol{ab}}(A)$ was compared against the Hamiltonian matrix element of the RAG, $h^{\mathrm{RAG}}_{\boldsymbol{ab}}$, which coincides with $\gamma^{\mathrm{sr,RAG}}_{\boldsymbol{ab}}$. The electron-lattice parameters $\vec{f}_{\boldsymbol{ab},\boldsymbol{\lambda v}}$ and $\overset{\leftrightarrow}{g}_{\boldsymbol{ab},\boldsymbol{\lambda v}}$ associated with a given atom $\boldsymbol{\lambda}$ were selected to be included in the electronic model if the following condition was met for at least one configuration in the training set,

$$\frac{1}{|u_{\boldsymbol{\lambda}\alpha}|}|h_{\boldsymbol{ab}}(A) - h^{\mathrm{RAG}}_{\boldsymbol{ab}}| > \delta f_{\mathrm{e-l}}. \tag{5.22}$$

Here, $\delta f_{\mathrm{e-l}}$ represented an energy cutoff expressed in eV/Å and A was the first-principles calculation belonging to the electron-lattice training set associated to the displacement $u_{\boldsymbol{\lambda}\alpha}$ of the $\boldsymbol{\lambda}$ atom.

If the condition in Eq. (5.22) was satisfied for the atom $\boldsymbol{\lambda}$, linear and quadratic variables concerning this atom were incorporated into the fitting procedure. It is important to highlight that the methodology described in this Section for the calculation of the model parameters, referred to as the *previous methodology* and outlined in Ref. [37], considered only the diagonal terms of the quadratic contributions, $\overset{\leftrightarrow}{g}_{\boldsymbol{ab},\boldsymbol{\lambda\lambda}}$, due to the fact that each training set calculation involved a single atomic displacement, as explained in Sec. 5.6.1.1.

5.6.1.5 Fitting procedure

The fitting procedure started with the selected set of parameters $\gamma_{\boldsymbol{ab}}$, $U_{\boldsymbol{ab},\boldsymbol{a}'\boldsymbol{b}'}$, and $I_{\boldsymbol{ab},\boldsymbol{a}'\boldsymbol{b}'}$. This procedure was carried out in a sequence of steps:

1. The $U_{ab,a'b'}$ parameters were fitted to reproduce the matrices $|h^U_{ab} - \bar{h}^U_{ab}|$, as described in Eq. (5.16).
2. The $I_{ab,a'b'}$ parameters were determined to reproduce the matrices h^I_{ab}, according to Eq. (5.17).
3. Once the values of the $U_{ab,a'b'}$ parameters were commputed, the γ_{ab} values were fitted by utilizing the h^U_{ab} matrix, as given in Eq. (5.15).
4. Knowing the values of γ_{ab}, it was possible to fit the electron-lattice parameters $\vec{f}_{ab,\lambda\upsilon}$ and $\overset{\leftrightarrow}{g}_{ab,\lambda\upsilon}$.

5.6.1.6 Limitations of *previous methodology*

The *previous methodology* presented several significant challenges, particularly in the selection of the training set, the incorporation of symmetry, and the treatment of second-order parameters. One of the primary issues was the lack of a formal, well-defined criterion for constructing the training set, especially for electron-electron interactions. The training set was composed of so-called *representative configurations* [37], which were inherently system-dependent. As a result, the training set selection was subjective, user-defined, and lacked a rigorous and reproducible framework.

Beyond this ambiguity, establishing a systematic and physically meaningful training set required a profound understanding of the underlying physics governing the system and the relevant computational processes. This introduced an additional preparatory step in which the system had to be extensively studied before generating the training set. Consequently, applying this methodology to previously unexplored systems became impractical, particularly in the context of MODELMAKER calculations. The dependence on prior expertise also meant that only users with deep knowledge of both MODELMAKER and the specific system under investigation could effectively apply the methodology. Furthermore, since each training set was uniquely tailored to a given system, direct comparisons among results obtained from different training sets became unfeasible, as each training set contained system-specific characteristics that were not standardized across studies.

Another major limitation of this approach was the absence of built-in symmetry enforcement within the computational framework. Since symmetry was not intrinsically accounted for in the code, it had to be imposed manually. For example, in the case of the electron-lattice training set, this was typically achieved by including additional calculations in the training set to enforce symmetry, but this workaround significantly increased computational costs. The need for extra calculations made the methodology inefficient, particularly for complex systems where symmetry considerations play a crucial role.

In addition to these issues, the methodology did not fully account for second-order electron-lattice coupling parameters. While it considered diagonal second-order parameters, $\overset{\leftrightarrow}{g}_{ab,\lambda\lambda}$, it neglected cross-term contributions, $\overset{\leftrightarrow}{g}_{ab,\lambda\upsilon}$ for $\lambda \neq \upsilon$. This omission limited the accuracy and completeness of the model, as cross-term interactions can be essential for capturing the full complexity of electron-lattice coupling.

Altogether, these limitations—ranging from the lack of a formalized training set selection process and the computational burden of manual symmetry enforcement to the incomplete treatment of second-order parameters—compromised both the robustness and reproducibility of the methodology. Furthermore, they restricted its applicability to new and complex systems, highlighting the need for a more systematic, automated, and physically grounded approach.

Thus, in this work, two alternative methodologies have been studied with the objective of addressing and overcoming the previously discussed issues related to parameter calculations. These methodologies aim to improve various aspects of the process, including the construction of the necessary training set, the criteria for parameter selection, and the computation of the parameters themselves. They are designed as *methodology I*, based on the fitting of the parameters and *methodology II* focused on the calculation of the electron-lattice variables.

5.6.2 Proposed methodology I: Fitting procedure

Methodology I is characterized by obtaining the model parameters through a fitting procedure using the available first-principles data. In this methodology, prior to the fitting process, all potential variables are identified (without computing its value) based on a set of predefined cutoff criteria, which will be discussed in detail in the following sections. Subsequently, the fitting procedure is applied to determine the values of one-electron parameters, electron-electron interaction terms, and electron-lattice coupling terms. It includes the symmetry of the different elements conforming the model according to Sec. 5.4. This methodology operates on the Hamiltonian matrix elements of the system, expressed within a Wannier function basis, $h_{\boldsymbol{ab}}$.

Starting from the DFT-calculated matrix elements, denoted as $h_{\boldsymbol{ab}}^{\mathrm{DFT}}$, the primary objective is to construct the second-principles Hamiltonian matrix elements, $h_{\boldsymbol{ab}}$, which depend on a set of parameters $\{p\}$. In the fitting procedure, the goal is to ensure that these matrix elements closely replicate the configuration and behavior observed in the training set calculations $\{A\}$, presented later in Secs. 5.6.2.3 and 5.6.2.5, which are characterized by the geometry and the electronic state of the system respectively. To achieve this, the parameters are obtained by minimizing the discrepancy between the first-principles results and the second-principles Hamiltonian, quantified through the quadratic error, known as the *goal function* and presented in Sec. 5.5.

In order to preserve the intrinsic symmetries of the physical system, as governed by Eqs. (5.5)-(5.9), these symmetry considerations are integrated into the optimization process as constraints. Specifically, the symmetry constraints are incorporated into the goal function through the use of Lagrange multipliers, ensuring their enforcement during parameter refinement. The final form of the goal function encapsulates both the quadratic error and the symmetry constraints, representing the framework for parameter extraction. The resulting goal-function to carry out the fitting of the parameters is

$$\Theta(\{A\},\{p\}) = \sum_{\boldsymbol{ab}} \sum_{A} \left[h_{\boldsymbol{ab}}(A,\{p\}) - h_{\boldsymbol{ab}}^{\mathrm{DFT}}(A) \right]^2 - \sum_{C} \lambda_C S_C(\{p\}), \tag{5.23}$$

where $\{C\}$ is the set of all symmetry constraints, λ_C the Lagrange multiplier associated to each of them and $S_C(\{p\})$ the Eqs. (5.5)-(5.9). Now, considering the splitting of the second-principles Hamiltonian matrix elements into short and long-range terms, regarding to Eq. (4.19),

$$h_{\boldsymbol{ab}}(A,\{p\}) = h^{\mathrm{sr}}_{\boldsymbol{ab}}(A,\{p\}) + h^{\mathrm{lr}}_{\boldsymbol{ab}}(A). \tag{5.24}$$

The long-range contribution is computed using a multipolar expansion based on charges and dipoles. Adding this division into Eq. (5.23),

$$\Theta(\{A\},\{p\}) = \sum_{\boldsymbol{ab}}\sum_{A}\left[h^{\mathrm{sr}}_{\boldsymbol{ab}}(A,\{p\}) + h^{\mathrm{lr}}_{\boldsymbol{ab}}(A) - h^{\mathrm{DFT}}_{\boldsymbol{ab}}(A)\right]^2 - \sum_{C}\lambda_C S_C(\{p\}). \tag{5.25}$$

The minimization of the goal function Θ with respect to the parameters $\{p\} = \gamma^{\mathrm{RAG}}_{\boldsymbol{ab}}, \vec{f}_{\boldsymbol{ab},\boldsymbol{\lambda v}}, \overset{\leftrightarrow}{g}_{\boldsymbol{ab},\boldsymbol{\lambda v}}, U_{\boldsymbol{ab},\boldsymbol{a'b'}}, I_{\boldsymbol{ab},\boldsymbol{a'b'}}$ and the Lagrange multipliers λ_C leads to the following linear equation system,

$$\sum_{\boldsymbol{ab}}\sum_{A} h^{\mathrm{sr}}_{\boldsymbol{ab}}(A,\{p\})\frac{\partial h^{\mathrm{sr}}_{\boldsymbol{ab}}}{\partial p_i}(A) - \frac{1}{2}\sum_{C}\lambda_C\frac{\partial S_C}{\partial p_i}(\{p\}) = \sum_{\boldsymbol{ab}}\sum_{A}\left[h^{\mathrm{DFT}}_{\boldsymbol{ab}}(A) - h^{\mathrm{lr}}_{\boldsymbol{ab}}(A)\right]\frac{\partial h^{\mathrm{sr}}_{\boldsymbol{ab}}}{\partial p_i}(A), \tag{5.26}$$

$$S_C(\{p\}) = 0. \tag{5.27}$$

In this methodology, the term “goal function” can be used differently depending on the nature of the training set. Specifically, we refer to the *electron-lattice goal function* when the training set consists of calculations involving geometric distortions of the system. Conversely, the *electron-electron goal function* term is employed when the training set exclusively captures variations in the electronic structure while maintaining the lattice configuration as in the RAG.

5.6.2.1 One-electron parameters: Reference state Hamiltonian

The values of the one-electron parameters $\gamma^{\mathrm{sr,RAG}}_{\boldsymbol{ab}}$ could be directly extracted from the output generated by the WANNIER90 code (used as a post-processing tool of DFT calculations) for a calculation corresponding to the system in its non-magnetic ground state (within the RED) at the RAG. However, fulfilling the symmetry constrains, Eq. (5.5), precisely requires performing the DFT calculation setting many of the usual computational variables, like the reciprocal space mesh, to very high accuracy, which is costly in computational power. A more efficient way to proceed, as implemented in MODELMAKER, is to use Eq. (5.5) to perform a weighted average of the Hamiltonian matrix elements obtained from WANNIER90 code inside a symmetry group. In this way, the symmetry constraints can be satisfied while maintaining computational efficiency.

Another important aspect to consider is that the training set is computed within a supercell of the *unit cell*, as previously stated. Therefore, it is essential that the Hamiltonian matrix elements obtained from WANNIER90 for the *supercell* in the RAG are consistent with those obtained for the *unit cell*. The *unit cell* consistency can be rigorously preserved by using explicit symmetry constraint on the unitary transformation matrices [70]. The procedure for achieving this property has already been discussed in Sec. 2.1.1. To verify the consistency between the elements obtained for the *unit cell* and those derived from the *supercell*, the proposed methodology incorporates a validation step in which the code compares the Hamiltonian matrix elements $\gamma_{\boldsymbol{ab}}^{\mathrm{sr,RAG}}$ obtained from the *unit cell* and the *supercell* for the RAG within the RED configuration. If the difference between these values falls below a predefined threshold, the corresponding Hamiltonian matrix element in the *supercell* is assigned to the equivalent interaction in the *unit cell*. However, if the discrepancy exceeds the established threshold, the code issues a warning message. In such cases, modifications to the Wannierization procedure are required. Possible adjustments include refining the $\vec{k}$-point mesh of the first Brillouin zone, modifying the energy windows of the manifolds, or selecting alternative initial projector orbitals to generate a different set of Wannier functions. For instance, in covalent systems, instead of projecting Wannier functions onto atomic orbitals, it may be beneficial in some cases to construct Wannier functions that represent molecular orbitals, depending on the specific characteristics of the system.

Once the parameters $\gamma_{\boldsymbol{ab}}^{\mathrm{sr,RAG}}$ are obtained, a Hamiltonian distance cutoff, δr_{h}, is imposed to exclude interactions between Wannier functions whose centroids are separated by a distance greater than δr_{h}. The purpose of applying this distance cutoff is analogous to that of the energy cutoff $\delta \varepsilon_{\mathrm{h}}$ introduced in the *previous methodology*. The criterion established by δr_{h} determines the selected pairs $(\boldsymbol{a}, \boldsymbol{b})$, including all symmetry-related interactions, for constructing the Hamiltonian matrix elements $h_{\boldsymbol{ab}}^{s}$ considered in the model. Consequently, when constructing the parametric Hamiltonian $h_{\boldsymbol{ab}}^{s}$ [Eq. (4.37)] for the pair $(\boldsymbol{a}, \boldsymbol{b})$, at least the RAG-associated term, $\gamma_{\boldsymbol{ab}}^{\mathrm{sr,RAG}}$, is included. Later, the electron-lattice coupling and the electron-electron interaction will be generated for a subset of Wannier function pairs chosen from those selected in this step.

The possibility to impose this cutoff δr_{h} stems from the localized nature of the Wannier functions in real space. Since the Hamiltonian matrix elements expressed in a WFs basis set decay rapidly with increasing distance, only a limited number of Hamiltonian matrix elements are typically significant for accurately describing the band structure of the RAG, as illustrated in Eq. (1.52). By implementing the cutoff distances δr_{h}, our Hamiltonian is explicitly range-limited and local, a critical property for ensuring that second-principles calculations scale linearly in computational power relative to system size.

A similar approach is applied to the position operator matrix elements, $\vec{r}_{\boldsymbol{ab}}$. Here, we include all diagonal elements that define the position of the Wannier function centroid while discarding off-diagonal elements $(\boldsymbol{a} \neq \boldsymbol{b})$ whose $|\vec{r}_{\boldsymbol{ab}}|$ magnitude is below the distance cutoff δr_{r}.

5.6.2.2 Electron-lattice parameters

In the previous step, we selected the pairs of Wannier functions for which the one-electron Hamiltonian matrix elements, $h_{\boldsymbol{ab}}^{s}$, are included in the model. As previously discussed, and following Eq. (4.37), these elements incorporate the contribution from the undistorted system in the RED, $\gamma_{\boldsymbol{ab}}^{\mathrm{sr,RAG}}$. From this set of Hamiltonian matrix elements, a subset will be selected to include the electron-lattice interaction, thereby incorporating the electron-lattice coupling correction in relation to Eq. (4.24),

$$\delta \gamma_{\boldsymbol{ab}}^{\mathrm{el}} = \sum_{\boldsymbol{\lambda v}} \left[-\vec{f}_{\boldsymbol{ab},\boldsymbol{\lambda v}}^{T} \delta \vec{\tau}_{\boldsymbol{\lambda v}} + \sum_{\boldsymbol{\lambda' v'}} \delta \vec{\tau}_{\boldsymbol{\lambda v}}^{T} \overset{\leftrightarrow}{g}_{\boldsymbol{ab},\boldsymbol{\lambda v}} \delta \vec{\tau}_{\boldsymbol{\lambda v}} \right]. \tag{5.28}$$

To determine which Wannier function pairs require an electron-lattice correction, it is necessary to construct a training set that identifies the most relevant electron-lattice variables, prioritizing those for which the goal function (which measures the quadratic error) exhibits the highest contributions.

Once the electron-lattice parameters to be included in the Hamiltonian matrix elements are selected, their values are determined by solving the system of equations defined by Eqs. (5.26)–(5.27). This process employs a modified version of the stepwise forward-selection procedure outlined in Ref. [51] (see Sec. 5.6.2.7), adapted to the variables of the electronic model.

Thus, in order to select the electron-lattice parameters applying *methodology I*, the first step is to construct the training set.

5.6.2.3 Electron-lattice training set

To select and fit the electron-lattice parameters, a training set is constructed from a collection of configurations obtained through first-principles calculations in the chosen *supercell*. The first-principles calculations within this training set involve atomic distortions consisting of random displacements around their reference positions (i.e., their positions in the RAG) within a sphere of a predefined radius, following a uniform distribution, as illustrated in Fig. 5.6 for the $SrTiO_3$ system. Notably, in these calculations, the system's RED remains unchanged.

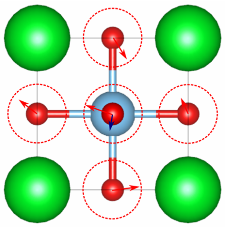


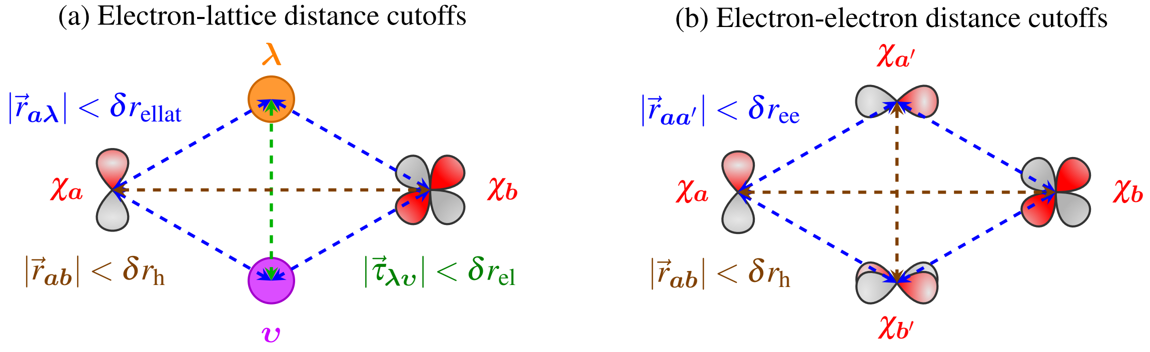
Figure 5.6: Random displacements of O and Ti atoms around a sphere to generate calculations for the training set of the $SrTiO_3$ system.

This training set construction methodology is intended to activate all possible vibrational modes of the system. Relying solely on predefined phases with high symmetry could leave certain modes unrepresented, resulting in an incomplete characterization of electron-lattice interactions.

Cutoffs

Before the fitting procedure, MODELMAKER code generates a list of the potentially relevant $\vec{f}$ and $\overleftrightarrow{g}$ variables to be considered in the model. This list can become excessively long due to the large number of electron-lattice variables involved: these variables depend on the position of two Wannier functions and two atoms. Although the number of free parameters is often reduced by the symmetry constraints, the list may still require further trimming to manage the computational complexity of the second-principles model.

As a result, to enhance the efficiency of the model, while maintaining a reasonable level of accuracy, specific cutoffs are introduced for the selection of the most relevant electron-lattice variables. On the one hand, the cutoff-distance δr_{el} defines the maximum distance, measured in the RAG, between the two atoms $\boldsymbol{\lambda}$ and $\boldsymbol{\upsilon}$ to form a variable [represented in green in Fig. 5.7(a)]. To this cutoff distance, we have to add $\delta r_{\mathrm{ellat}}$ which establishes the maximum distance between the center of the Wannier functions $\chi_{\boldsymbol{a}}$ and $\chi_{\boldsymbol{b}}$ and the position of the atoms $\boldsymbol{\lambda}$ and $\boldsymbol{\upsilon}$, measured in the RAG [represented in blue in Fig. 5.7(a)]. These distance thresholds are applied to refine the list of variables considered for the model. The application of distance cutoff $\delta r_{\mathrm{ellat}}$ is not restricted, given the localized nature of the Wannier functions, it is reasonable to assume that atoms separated by large distances from Wannier functions are effectively uncorrelated. In those cases, the electron-lattice coupling between them is sufficiently weak to be considered negligible. This approach allows for a significant reduction in computational complexity without severely compromising the model's accuracy.


(a) Electron-lattice distance cutoffs

(b) Electron-electron distance cutoffs

Figure 5.7: Distance cutoffs defined in MODELMAKER and controlled by the user. All cutoffs are measured with respect to the positions of the Wannier functions and the atoms in the RAG. In panel (a), distance cutoffs used in the selection of electron-lattice parameters $\{\vec{f}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}}, \overleftrightarrow{g}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}}\}$. The cutoff δr_{h} establishes the maximum distance between pairs of Wannier functions $(\boldsymbol{a},\boldsymbol{b})$. The cutoff δr_{el} defines the maximum distance between a pair of atoms $(\boldsymbol{\lambda},\boldsymbol{\upsilon})$ for the formation of a variable. The cutoff $\delta r_{\mathrm{ellat}}$ determines the maximum distance between Wannier functions $(\boldsymbol{a},\boldsymbol{b})$ and atoms $(\boldsymbol{\lambda},\boldsymbol{\upsilon})$ for the selection of electron-lattice parameters. In panel (b), distance cutoffs used in the selection of electron-electron parameters $\{U_{\boldsymbol{ab},\boldsymbol{a'b'}}, I_{\boldsymbol{ab},\boldsymbol{a'b'}}\}$. The cutoff δr_{h} establishes the maximum distance between the two Wannier functions forming a pair: $(\boldsymbol{a},\boldsymbol{b})$ and $(\boldsymbol{a}',\boldsymbol{b}')$. The cutoff δr_{ee} defines the maximum distance between the pairs $(\boldsymbol{a},\boldsymbol{a}')$, $(\boldsymbol{a},\boldsymbol{b}')$, $(\boldsymbol{b},\boldsymbol{a}')$, and $(\boldsymbol{b},\boldsymbol{b}')$ for considering the pairs $(\boldsymbol{a},\boldsymbol{b})$ and $(\boldsymbol{a}',\boldsymbol{b}')$ in the selection of an electron-electron variable.

The impact of electron-lattice interactions is inherently tied to the nature and symmetry of the Wannier functions involved. As such, the Hamiltonian matrix elements $h_{\boldsymbol{ab}}$ will exhibit varying degrees of significance for electron-lattice interactions. In some cases, the interaction may be negligible, while in others, it plays a pivotal role. To streamline the calculations and reduce the number of electron-lattice parameters, a novel energy cutoff, $\delta\varepsilon_{\mathrm{el}}$, is introduced. This cutoff serves to identify and select symmetry groups for which the electron-lattice interaction is most relevant.

For a given symmetry group G, its contribution to the electron-lattice goal function Θ_{el}, which is obtained from Eq. (5.23) when considering the electron-lattice training set calculations, is determined by assessing the error that arises when the second-principles models are simply described by $\gamma_{\boldsymbol{ab}}$ calculated in the RAG for RED. The contribution is thus quantified as

$$\Theta_{\text{el}}^{G} = \sum_{A^{\text{el}}} \sum_{\boldsymbol{ab} \in G} \left[h_{\boldsymbol{ab}}^{\text{DFT}}(A) - \left(h_{\boldsymbol{ab}}^{\text{RAG,sr}} + h_{\boldsymbol{ab}}^{\text{lr}} \right) \right]^2 . \tag{5.29}$$

Thus, if the contribution of the symmetry group G to Θ_{el} over the hole electron-lattice training set is larger than $\delta\varepsilon_{\text{el}}$ (measured in eV^2), $\Theta_{\text{el}}^{G} > \delta\varepsilon_{\text{el}}$, then the electron-lattice parameters associated with pairs of WFs $(\boldsymbol{a}, \boldsymbol{b})$, belonging to symmetry group G, are considered to form electron-lattice parameters. See the results chapter, Chapter 6, for some examples.

5.6.2.4 Electron-electron parameters

To calculate the electron-electron parameters, we compute the derivative of the one-electron Hamiltonian, Eq. (4.37), with respect to $p_i \in \{U_{\boldsymbol{ab},\boldsymbol{a}'\boldsymbol{b}'}, I_{\boldsymbol{ab},\boldsymbol{a}'\boldsymbol{b}'}\}$ to derive the expression in Eq. (5.26) with respect to the electron-electron parameters,

$$\frac{\partial h_{\boldsymbol{ab}}^{\text{s, sr}}}{\partial U_{\boldsymbol{ab},\boldsymbol{a}'\boldsymbol{b}'}} = D_{\boldsymbol{a}'\boldsymbol{b}'}^{s} + D_{\boldsymbol{a}'\boldsymbol{b}'}^{-s}, \tag{5.30}$$

$$\frac{\partial h_{\boldsymbol{ab}}^{\text{s, sr}}}{\partial I_{\boldsymbol{ab},\boldsymbol{a}'\boldsymbol{b}'}} = -\left(D_{\boldsymbol{a}'\boldsymbol{b}'}^{s} - D_{\boldsymbol{a}'\boldsymbol{b}'}^{-s} \right). \tag{5.31}$$

With the equations introduced above, it becomes evident that (i) the parameters $U_{\boldsymbol{ab},\boldsymbol{a}'\boldsymbol{b}'}$ and $I_{\boldsymbol{ab},\boldsymbol{a}'\boldsymbol{b}'}$ quantify the sensitivity of the one-electron Hamiltonian, $h_{\boldsymbol{ab}}^{s}$, to variations in the charge and spin polarization encoded in the density matrix elements $D_{\boldsymbol{ab}}$. It is important to emphasize that these quantities do not represent pure electron-electron interactions. Instead, they describe the effective response of the electronic structure to changes in the density matrix, incorporating not only electron-electron interactions but also contributions from the electron kinetic energy and the electron-nuclear potential. (ii) Accurate evaluation of these parameters requires explicit control over the electronic density—a task that, as previously discussed, is challenging to implement within standard first-principles frameworks. For this reason, $U_{\boldsymbol{ab},\boldsymbol{a}'\boldsymbol{b}'}$ and $I_{\boldsymbol{ab},\boldsymbol{a}'\boldsymbol{b}'}$ are obtained in *methodology I*, as in the case of electron-lattice parameters, by solving the fitting equations, Eqs. (5.26)–(5.27), using a modified stepwise regression algorithm with forward selection, adapted from the method described in Ref. [51] to accommodate the variables of the present electronic model.

After establishing the equations necessary for the fit, we need to determine which calculations can be used to create an adequate training set to obtain $U_{\boldsymbol{ab},\boldsymbol{a}'\boldsymbol{b}'}$ and $I_{\boldsymbol{ab},\boldsymbol{a}'\boldsymbol{b}'}$.

5.6.2.5 Electron-electron training set

According to previous equations it is possible to examine the fundamental characteristics of the electron-electron variables. This preliminary analysis provides a deeper understanding of their nature and behavior, setting the stage for a more systematic exploration of the required computational data and delving into the collection of first-principles calculations necessary to derive the electron-electron parameters.

From the preceding equations, it becomes possible to examine the fundamental properties of the electron-electron variables. This initial analysis not only deepens our understanding of their inherent characteristics but also lays the groundwork for a more systematic approach to analyzing the computational data and selecting the appropriate set of first-principles calculations required to determine the electron-electron parameters. Attending to Eqs.(5.30)-(5.31), the first-principles calculations employed to fit the electron-electron parameters must be characterized by changes in density with respect to the RED, including both changes in orbital occupations and spin polarization. Thus, the first-principles calculations included in the training set can be classified into four distinct categories:

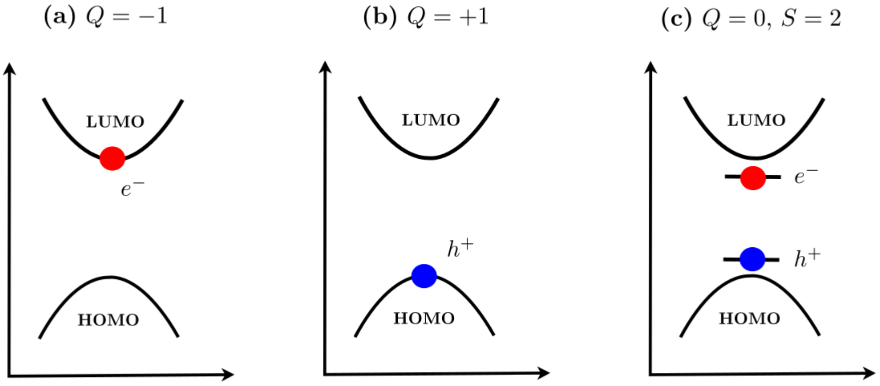


Figure 5.8: Training set employed for the calculation of electron–electron interaction parameters, based on variations in electron and hole populations, which in turn alter the total charge (Q) and total spin (S) of the system. This approach enables the evaluation of electron–hole coupling, a key component for simulating excitonic effects. Panel (a) illustrates electron doping of the system, panel (b) shows hole doping, and panel (c) depicts the excitation of a valence electron to the conduction band, resulting in the formation of an electron–hole pair. Figure extracted from Ref. [111].

Doping with electrons and holes: The variables $U_{ab,a'b'}$ are specifically linked to the occupations of the system, without accounting for spin polarization. Consequently, the first-principles calculations performed within the selected *supercell* and included in the training set are characterized by electron and hole doping relative to the chosen RED, distributed equally across the two spin channels.

For instance, in the case of an insulating system where the selected RED corresponds to the ground state, some first-principles simulations involve doping the conduction band with electrons Fig. 5.8(a) and other calculations focus on modifying the electronic configuration by introducing holes in the valence band, Fig. 5.8(b). This approach ensures the accurate representation of the system's response to variations in occupation. The magnitude of doping is systematically varied, ranging from -0.3 *e* to 0.3 *e* in steps of 0.1 *e*. Here, negative values denote electron doping, while positive values correspond to the introduction of holes. This controlled doping scheme captures the subtle effects of occupation changes on the system's electronic properties.

Spin-polarization. The $I_{ab,a'b'}$ variables are closely associated with the system's behavior arising from spin polarization. To account for this, the first-principles simulations used to capture the differences in spin channel occupations are designed to include both electron and hole doping, with the key constraint that the system is forced into a fully spin-polarized state. In other words, the charge introduced for doping is entirely represented by the difference in occupations between the spin-up and spin-down channels.

This approach ensures that the calculations effectively capture the system's behavior under spin polarization conditions, providing critical information about its magnetic and electronic properties in scenarios where spin polarization plays a dominant role.

Excitations in the RAG: Furthermore, the training set for the electron-electron variables incorporates first-principles simulations where the system remains neutral in charge, but where the spin momenta is enforced to be different from zero, reflecting distinct occupations of the spin channels. This group of calculations is specifically designed to explore optical properties, such as the formation of excitons. For example, in the case of the hBN insulating system previously discussed, these simulations model the excitation of electrons from the valence band to the conduction band, which results in the creation of holes in the valence band, Fig. 5.8(c). As a consequence, the resulting model is capable of capturing the excitonic properties of the system, providing a deeper understanding of its optical characteristics.

Excitations in distorted geometries: The study of polarons and excitons from second-principles represents one of the key objectives in understanding optical properties in solid-state physics. Therefore, the inclusion of first-principles calculations that describe excitations into the training set is particularly significant. However, when these calculations are integrated while maintaining the system in the RAG, the excited charge tends to remain delocalized within the crystal. This localization prevents the development of a second-principles model capable of simulating polaronic effects, as the parameters fail to capture the response of Wannier functions to effectively localize the charge. Consequently, in order to address this limitation and incorporate both excitonic and polaronic effects, the current approach involves modifying both the electron density and the atomic geometry with respect to the RED and the RAG, respectively, during the training set calculations. This fact deviates from the initial goal of producing a fully automated code. It is expected that, over time, this type of problem can be resolved.

Cutoffs

The $U_{ab,a'b'}$ and $I_{ab,a'b'}$ parameters represent integrals which depend on the position of four WFs. As a result, the number of electron-electron parameters is typically too large to be able to determine all of them, a problem that becomes more severe due to the inability to perturb specific density matrix elements through the training set described above. Thus, akin to the selection of electron-lattice parameters, we establish two criteria to determine which integrals will be fitted to the data. The first one is based on the distance between Wannier functions, δr_{ee}, that sets the maximum separation between the centroids of any pair of WFs involved in the four-center integral, measured in the RAG [see Fig. 5.7(b)]. The second is an energy cutoff, $\delta\Theta$, which is based on the contribution of a given symmetry group G to the electron-electron goal function, Θ_{ee}^{G}, built from Eq. (5.25) considering the electron-electron training set calculations, when the second-principles models are simply described by the $\gamma_{\boldsymbol{ab}}$ calculated at the RAG for the RED,

$$\Theta_{\text{ee}}^{G} = \sum_{A^{\text{ee}}} \sum_{\boldsymbol{ab} \in G} \left[h_{\boldsymbol{ab}}^{\text{DFT}}(A) - \left(h_{\boldsymbol{ab}}^{\text{RAG,sr}} + h_{\boldsymbol{ab}}^{\text{lr}} \right) \right]^{2}. \tag{5.32}$$

Thus, if the contribution to Θ_{ee}^{G} over the whole electron-electron TS of any group G is larger than $\delta\Theta$ (measured in eV^2) then electron-electron parameters associated with $(\boldsymbol{a}, \boldsymbol{b})$ pairs of WFs of that group are considered to form four-center integrals that will be included in the model.

5.6.2.6 Determining the variables: linear problem

Considering the final Hamiltonian formulated for our model, it depends linearly with the parameters aimed to fit, $\{p_i\}_{1 \leq i \leq P}$. So that, for each configuration we can write the Hamiltonian as

$$h_{\boldsymbol{ab}}(A) = \sum_{i} \theta_{Ai} p_i \tag{5.33}$$

where θ_{Ai} is a configuration-dependent constant that contains the correlation between the parameter p_i and DFT Hamiltonian element for a particular configuration A. The goal function can be then expressed in terms of the Eq. (5.33),

$$\Theta(\{A\}, \{p_i\}) = \sum_{A} \left[h_{\boldsymbol{ab}}^{\text{DFT}}(A) - \sum_{i} \theta_{Ai} p_i \right]^{2}. \tag{5.34}$$

Regarding the geometrical representation of Θ, it is a P-dimensional parabola which always satisfies $\Theta \geq 0$. Since the goal function is positive semi-define, the eigenvalue of the associated Hessian will be always positive or zero. As a result, a critical point of Θ must be a minimum.

The extrema of the goal function satisfy $\partial\Theta/\partial p_i = 0\ \forall i$. Applying this equation for the P parameters, we obtain the linear system for the calculation of the $\{p_i\}$ parameters,

$$\sum_{k} A_{ik} p_k = b_i, \tag{5.35}$$

where

$$A_{ik} = \sum_{A} \theta_{iA} \theta_{kA}, \tag{5.36}$$

and

$$b_i = \sum_{A} \theta_{iA} h^{\text{DFT}}(A). \tag{5.37}$$

However, there may be linear dependencies in our system of equations leasing to a overdetermined model. Then we need to find which are the significant variables of our model. This problem is addressed in the following section.

5.6.2.7 Finding the best model of p variables

In the preceding sections, we introduced the electronic degrees of freedom within the second-principles framework through the electron-lattice and electron-electron interaction parameters. We have detailed the methodology for generating these parameters, emphasizing the pivotal role of the selected training sets and the various cutoffs in their determination. As a consequence, it becomes evident that even with a realistic cutoff, the number of independent parameters remains orders of magnitude too large to be fitted in full. Achieving a sufficient number of data points from *ab initio* calculations to constrain all parameters would be computationally prohibitive. Moreover, an excessive number of parameters in the Hamiltonian matrix elements would undermine our objective of constructing a computationally efficient model. Additionally, fitting the entire set of parameters would result in an overfitted model, which would fail to generalize beyond the specific training set.

To address this issue, a systematic reduction of the parameter space is necessary, selecting only the most relevant parameters from the full set. Our objective is to develop a model that captures the essential physical trends rather than one that overfits the training set, thereby incorporating noise and specific fluctuations. An algorithmic approach for this selective parameterization was introduced by Escorihuela-Sayalero et al. [51] and has been reimplemented in MODELMAKER.

The advanced fitting algorithm implemented in MODELMAKER consists of two key components: first, the selection of relevant model parameters from the complete set of possibilities, and second, the determination of these parameters through a least-squares fitting procedure. The fitting is performed by minimizing the goal function, as defined in Eq. (5.25). In this approach, the fitted quantities correspond to the *ab initio* data from the training set.

In the following, we outline the step-by-step fitting procedure used to obtain the model parameters.

We consider $\mathscr{X}_P$ as the set of all the parameters generated by the code, a collection of P parameters (considering the training set and the different cutoffs) which determine *all* the possible interactions in the system of interest. Now, from the complete set of P parameters, we define a p-model as a model of p parameters: $\mathscr{X}_p \subset \mathscr{X}_P$. We need a procedure to build the best p-model, i.e., a criterion to choose the p parameters that minimizes the goal function. The first idea to solve this problem consists on computing all the possible combinations of p parameters of a set of P parameters and compare the resultant goal functions. However, this method is quite inefficient. For example, if we have a model with $P = 500$ parameters and we want to built a $p-$model such that $p = 20$, we have about 10^{35} different combinations of parameters, i.e, 10^{35} different models. As a result, although we have selected an efficient strategy to calculate the parameters of a p-model, the analysis of so many combinations is unfeasible from the computational point of view.

In order to overcome this problem, the model search is restricted by a *step wise procedure with forward selection*. The method starts considering $p = 1$ and building all the P possible 1-model candidates with its respective goal functions. Let $\mathscr{X}_1$ represents all the possible 1-models associated with the goal function

$\Theta_1(p_i)$, which considers only one parameter. The next step is focused on selecting the best 1-model by minimizing $\Theta_1(p_i)$ according to Eq. (5.34). Comparing the $\Theta_1(p_i)$ $\forall i$, we choose the smallest one. At this point, we can identify the first parameter of our model, p_1^* which composes the set $\mathscr{X}_1^*$. Then, we move to $p+1$ and we consider all the possible $(p+1)$-models which contains the variables of the best $p-$model, i.e, we study all the $p+1$-models with the constraint $\mathscr{X}_p^* \subset \mathscr{X}_{p+1}$. Keeping the p parameters of the best p-model, the $p+1$-models are built adding one by one the $P-p$ remaining variables. Again, we build the goal function $\Theta_p(\{\{p_j^*\}, p_i\})$ $\forall j \leq p$, $\forall p_i \neq p_j^*$ and it is minimized according to Eq. (5.34). Then, we select the parameter p_i which makes minimum $\Theta_p(\{\{p_j^*\}, p_i\})$ $\forall j \leq p$, $\forall p_i \neq p_j^*$.

Below, the parameters with the star in $p+1$-model come from the best p-model. It should be noted that, in the $p+1$ model, only the variables from the p-model are introduced, while their values are determined through optimization at each iteration.

- Build best model 1 variable : $\{p_1\}$
- Build best model 2 variables : $\{p_1^*, p_2\}$
- Build best model 3 variables : $\{p_1^*, p_2^*, p_3\}$
- ...

Different parameters will be included in the model until the objective function falls below a predefined threshold.

5.6.2.8 Problems of the fitting procedure

The *methodology I*, based on the fitting procedure, presents several intrinsic challenges that must be addressed to ensure the robustness and reliability of the model. The various challenges are associated with the electron-lattice variables. These difficulties stem from the characteristics of the training set, the selection of relevant variables, and the underdetermined nature of the fitting equations.

One of the main issues arises in the construction of the electron-lattice training set. As previously noted, this set is generated by introducing random atomic displacements within a predefined radius. However, this stochastic approach prevents the exact reproduction of model calculations. Furthermore, once a model has been generated, incorporating additional calculations into the electron-lattice training set—even while maintaining the original dataset—modifies the selection of the most relevant variables identified in the initial fitting procedure, affecting subsequent minimization levels. In other words, the mere addition of a single calculation to the electron-lattice training set leads to the emergence of new variables in the fitting process, thereby introducing inconsistency. This shift in variable selection complicates the verification process.

Another related issue with the electron-lattice fitting procedure encountered during model calculations pertains to the indeterminacy in selecting fitting variables. Specifically, when using the same dataset of first-principles calculations, the variables chosen during the fitting process were not always identical. This inconsistency arises because the fitting procedure involves solving an underdetermined system of equations, where the number of model parameters (or degrees of freedom) exceeds the range of the system. As previously stated, this issue primarily concerns the electron-lattice parameters, particularly those associated with the linear terms. According to the definition in Eq. (3.119), the linear terms $\vec{f}_{ab,\lambda}$ depend on a single atom. However, the methodology introduces an artificial dependence on two atoms, resulting in modified linear variables $\vec{f}_{ab,\lambda\upsilon}$ in Eq. (4.24). This modification is implemented to automatically enforce the Acoustic Sum Rule (ASR). However, it results in an increased number of parameters relative to the available equations, thereby exacerbating the underdetermined nature of the system. This issue will be explicitly addressed in Sec. 5.6.3.

To solve these challenges, it is required to develop an improved methodological framework for the determination of parameters.

As discussed in the introduction of the Sec. 5.6, the parameters within a second-principles framework can either be explicitly calculated or fitted [36; 37; 51]. At first glance, the fitting procedure has been selected due to the inherent complexity of directly implementing a rigorous computational methodology for determining electron-electron parameters. For instance, the direct calculation of electron-electron parameters, $U_{ab,a'b'}$ and $I_{ab,a'b'}$, is particularly challenging, as employing a finite-difference method would necessitate precise control over the density matrix within a DFT framework [see Eq. (4.36) and Eqs. (5.30)-(5.31)]. While this could, in principle, be achieved using constrained DFT [181], its direct implementation falls beyond the scope of this work and is not readily transferable between different DFT codes.

However, due to the presented inconsistencies in the obtained results related to the electron-lattice contribution, direct parameter calculation may be preferable in the case of these parameters, as it often provides clearer physical interpretations. In contrast to electron-electron variables, the computation of electron-lattice parameters via finite-difference methods employing their definition in Eqs. (3.119)-(3.120) is feasible. This is because the derivatives of the Hamiltonian with respect to these parameters depend solely on atomic positions rather than on the density matrix. Consequently, these parameters capture variations in h_{ab} induced by atomic displacements within the RAG—a perturbation that is trivial to implement in first-principles calculations.

Therefore, we adopt a hybrid approach: electron-lattice parameters, $\vec{f}_{ab,\lambda\upsilon}$, $\overset{\leftrightarrow}{g}_{ab,\lambda\upsilon}$, are extracted from WANNIER90 and they are are explicitly calculated considering the symmetry constraints, while those associated with electron-electron interactions— $U_{ab,a'b'}$ and $I_{ab,a'b'}$—are determined via fitting. The parameter $\gamma_{ab}^{\mathrm{sr,RAG}}$ are obtained according to Sec. 5.6.2.1.

In the following Section, we establish the new methodology for computing electron-lattice parameters through direct calculation. This methodology is referred in this work as *methodology II*. In this approach, the electron-lattice parameters are determined directly by accounting for atomic displacements, $\vec{u}_{\lambda}$. However, the impact of strain has not been thoroughly investigated within this framework. Future research will address this aspect in detail.

5.6.3 Proposed methodology II: Calculation of electron-lattice parameters

As discussed in Chapter 3, a fundamental aspect of many physical phenomena is the dependence of material properties on the geometric configuration of the system. In the framework of second-principles density functional theory (SPDFT), this effect is accounted for by expanding the one-electron parameters, $\gamma^{\text{sr}}_{\boldsymbol{ab}}$, as a function of atomic coordinates. Accordingly, the variation of $\gamma^{\text{sr}}_{\boldsymbol{ab}}$ with respect to geometry was initially expressed as Eq. (4.24),

$$\gamma^{\text{sr}}_{\boldsymbol{ab}} = \gamma^{\text{sr,RAG}}_{\boldsymbol{ab}} + \sum_{\boldsymbol{\lambda\upsilon}} \left[-\vec{f}^{T}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda\upsilon}} + \delta\vec{\tau}^{T}_{\boldsymbol{\lambda\upsilon}} \cdot \overleftrightarrow{g}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda\upsilon}} \right], \tag{5.38}$$

where $\delta\vec{\tau}_{\boldsymbol{\lambda\upsilon}}$ measures the difference of displacements of atoms $\boldsymbol{\lambda}$ and $\boldsymbol{\upsilon}$,

$$\delta\vec{\tau}_{\boldsymbol{\lambda\upsilon}} = \overleftrightarrow{\eta} \left(\vec{R}_{\Upsilon} - \vec{R}_{\Lambda} + \vec{\tau}^{(0)}_{\upsilon} - \vec{\tau}^{(0)}_{\lambda} \right) + \vec{u}_{\boldsymbol{\upsilon}} - \vec{u}_{\boldsymbol{\lambda}}. \tag{5.39}$$

Due to the dependence of Eq. (5.38) on the difference between the displacements of atoms $\boldsymbol{\lambda}$ and $\boldsymbol{\upsilon}$, the acoustic sum rule is automatically satisfied, as previously discussed in Sec. 4.2.2. Moreover, attending to Eq. (5.38), the linear variables exhibit a dependence on two atoms rather than a single atom, as defined in Eq. (3.119).

While trying to systematically fit all the electron-lattice parameters according to Eq. (5.38) we found large inconsistencies in the values of $\vec{f}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}}$ when two different sets of reference calculations were used, as already commented. The issue becomes clear when comparing Eq. (5.38) and the direct expansion of the one-electron parameter of the Hamiltonian matrix elements in terms of the atomic coordinates,

$$\gamma^{\text{sr}}_{\boldsymbol{ab}} = \gamma^{\text{RAG,sr}}_{\boldsymbol{ab}} - \sum_{\boldsymbol{\lambda}} \vec{f}^{T}_{\boldsymbol{ab},\boldsymbol{\lambda}} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} + \sum_{\boldsymbol{\lambda\upsilon}} \delta\vec{\tau}^{T}_{\boldsymbol{\lambda}} \cdot \overleftrightarrow{g}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}} \cdot \delta\vec{\tau}_{\boldsymbol{\upsilon}}, \tag{5.40}$$

where

$$\delta\vec{\tau}_{\boldsymbol{\lambda}} = \overleftrightarrow{\eta} \left(\vec{R}_{\Lambda} + \vec{\tau}^{(0)}_{\lambda} \right) + \vec{u}_{\boldsymbol{\lambda}}. \tag{5.41}$$

At this point we compare both expressions: Eq. (5.38) and Eq. (5.40). It is clear that $\delta\vec{\tau}_{\boldsymbol{\lambda\upsilon}}$, $\vec{\tau}_{\boldsymbol{\lambda}}$ and $\delta\vec{\tau}_{\boldsymbol{\upsilon}}$ are related through

$$\delta\vec{\tau}_{\boldsymbol{\lambda\upsilon}} = \delta\vec{\tau}_{\boldsymbol{\upsilon}} - \delta\vec{\tau}_{\boldsymbol{\lambda}}. \tag{5.42}$$

Examining the linear terms, we observe that if the system contains N_{at} atoms, Eq. (5.40) requires fitting $3N_{\text{at}}$ independent parameters, accounting for the three Cartesian coordinates. In contrast, analyzing the previously employed expression, Eq. (5.38), we find that the number of parameters to be fitted is significantly larger, specifically $(3N_{\text{at}})^2 - 3N_{\text{at}}$, where the subtraction accounts for the diagonal terms. Notably, the linear terms in both formulations are related through the following expression,

$$\vec{f}_{\boldsymbol{ab},\boldsymbol{\lambda}} = -2 \sum_{\boldsymbol{\upsilon}} \vec{f}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}}. \tag{5.43}$$

Relation between one-atom linear terms, $\vec{f}_{ab,\lambda}$, and two-atoms linear terms, $\vec{f}_{ab,\lambda\upsilon}$.

Proof. To prove the relation in Eq. (5.43), we consider the linear contribution to the one-electron parameters in Eq. (5.38),

$$-\sum_{\lambda}\sum_{\upsilon}\vec{f}^{T}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\lambda\upsilon}=\sum_{\lambda}\sum_{\upsilon}\vec{f}^{T}_{ab,\lambda\upsilon}\cdot(\delta\vec{\tau}_{\lambda}-\delta\vec{\tau}_{\upsilon})=\sum_{\lambda}\sum_{\upsilon}\vec{f}^{T}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\lambda}-\vec{f}^{T}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\upsilon}. \tag{5.44}$$

A reasonable constraint is,

$$\vec{f}_{ab,\lambda\upsilon}=-\vec{f}_{ab,\upsilon\lambda}. \tag{5.45}$$

By incorporating this constraint, Eq. (5.45), into Eq. (5.44),

$$\sum_{\lambda}\sum_{\upsilon}\vec{f}^{T}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\lambda}-\vec{f}^{T}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\upsilon}=\sum_{\lambda}\sum_{\upsilon}\vec{f}^{T}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\lambda}+\vec{f}^{T}_{ab,\upsilon\lambda}\cdot\delta\vec{\tau}_{\upsilon}=2\sum_{\lambda}\sum_{\upsilon}\vec{f}^{T}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\lambda}. \tag{5.46}$$

Now, if we equalize this expression with the corresponding one-atom notation in Eq. (5.40),

$$-\sum_{\lambda}\vec{f}^{T}_{ab,\lambda}\cdot\delta\vec{\tau}_{\lambda}=2\sum_{\lambda}\sum_{\upsilon}\vec{f}^{T}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\lambda}, \tag{5.47}$$

we arrive to Eq. (5.43). □

As a result, the two-atoms linear constants $\vec{f}_{ab,\lambda\upsilon}$ are not well-defined, since there are many more of these variables than $\vec{f}_{ab,\lambda}$. Thus, it seems that using Eq. (5.40) is a better starting point to create second-principles models than Eq. (5.38).

A key challenge when employing Eq. (5.40) is ensuring that the acoustic sum rule is properly satisfied. By enforcing the physical requirement that the total energy remains invariant under a rigid translation of the entire lattice, one obtains for the linear parameters the constraint

$$\sum_{\lambda}\vec{f}_{ab,\lambda}=0. \tag{5.48}$$

The individual coefficients $\vec{f}_{ab,\lambda}$ can, in principle, be computed from first-principles by interpreting them as the negative first derivative of γ_{ab} with respect to atomic displacements [Eq. (3.119)] and evaluating them using finite differences. A straightforward application of this approach—displacing a single atom λ by a small distance d along a chosen direction and extracting the derivative—results in noticeable violations of Eq. (5.48) due to the neglect of global translational invariance.

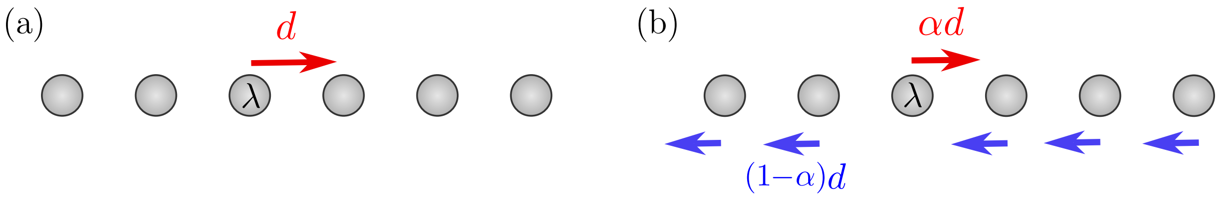


Figure 5.9: (a) Atoms of an equispaced one-dimensional linear chain, where the atom λ has been displaced by a distance d from its equilibrium position. (b) The geometry in [a] is equivalent to a linear chain in which the atom λ has been displaced by a distance αd, while the remaining atoms have been moved by a distance $(1-\alpha)d$ in the opposite direction.

A more robust finite-difference strategy involves displacing atom $\boldsymbol{\lambda}$ by a distance αd in a given direction while simultaneously shifting all other atoms in the simulation cell by a distance $(1-\alpha)d$ in the opposite direction. Here, α is a real parameter between 0 and 1. Since the relative displacements among atoms are preserved, the computed value of $\vec{f}_{\boldsymbol{ab},\boldsymbol{\lambda}}$ remains unchanged. In the following, we outline the reasoning behind the development of our methodology and we extract the value of α. For simplicity, we perform the derivation in a single spatial direction—the same reasoning applies equivalently to all three Cartesian directions.

Consider an atom $\boldsymbol{\lambda}$, displaced by a distance d in the x-direction, as illustrated in Fig. 5.9(a). This atomic configuration is equivalent to an atomic geometry in which the atom $\boldsymbol{\lambda}$ is displaced by a quantity αd, while all remaining atoms in the system are collectively shifted by $(1-\alpha)d$ in the opposite direction, as show in Fig. 5.9(b).

In order to compute the value of α, we consider in the system of Fig. 5.9(b) does not move the center of mass. At this point, the center of mass of the system satisfies

$$N_{\mathrm{at}} m \frac{[-(N_{\mathrm{at}}-1)(1-\alpha)d + \alpha d]}{N_{\mathrm{at}} m} = 0, \tag{5.49}$$

leading to

$$\alpha = \frac{N_{\mathrm{at}}-1}{N_{\mathrm{at}}}. \tag{5.50}$$

Once we have extracted the value of α, we return to the initial system, Fig. 5.9(a), to describe how this methodology is going to be applied. We can extract the electron-lattice linear coefficients from finite differences, denoted as $\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}}$. The linear term $\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}}$ represents the force exerted by atom $\boldsymbol{\lambda}$ when it is displaced by a distance d from equilibrium position. The equivalent scheme in which the ASR is strictly enforced while still extracting the electron-lattice linear coefficients from finite differences, consists on the following steps:

1. Consider the parameter $\alpha = (N_{\mathrm{at}}-1)/N_{\mathrm{at}}$, where N_{at} is the total number of atoms in the simulation box.
2. For each atom $\boldsymbol{\lambda}$, we assign a linear electron-lattice coupling constant $\vec{f}_{\boldsymbol{ab},\boldsymbol{\lambda}} = \alpha \vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}}$.
3. The remaining atoms $\boldsymbol{\upsilon} \neq \boldsymbol{\lambda}$ are each assigned a value $\vec{f}_{\boldsymbol{ab},\boldsymbol{\upsilon}} = -(1-\alpha)\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}}$.

This procedure ensures that the electron-lattice parameters are determined consistently while preserving the physical constraints imposed by the system. Indeed, to demonstrate the equivalence between the original formulation, Eq.(5.40), and the results obtained for the intermediate step, consider a scenario in which each atom $\boldsymbol{\lambda}$ is displaced by the amount $\delta\tau_{\boldsymbol{\lambda}}$. Then, in the original formulation, the contribution of the electron-lattice linear correction term to the Hamiltonian matrix elements $h_{\boldsymbol{ab}}$ is

$$\sum_{\boldsymbol{\lambda}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^{T} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}}. \tag{5.51}$$

Now, for each atom $\boldsymbol{\lambda}$ we introduce the equivalent situation described previously. Then, rather than directly considering the expression in Eq. (5.51) we analyze the formulation with the assignment,

$$\vec{f}_{\boldsymbol{ab},\boldsymbol{\lambda}} = \alpha \vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}}, \quad \text{and} \quad \vec{f}_{\boldsymbol{ab},\boldsymbol{v}} = -(1-\alpha)\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}}. \tag{5.52}$$

The resulting expression is

$$\sum_{\boldsymbol{\lambda}} \left[\vec{f}^{\,T}_{\boldsymbol{ab},\boldsymbol{\lambda}} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} + \sum_{\boldsymbol{v}\neq\boldsymbol{\lambda}} \vec{f}^{\,T}_{\boldsymbol{ab},\boldsymbol{v}} \cdot \delta\vec{\tau}_{\boldsymbol{v}} \right] = \sum_{\boldsymbol{\lambda}} \left[\alpha(\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^{T} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} + \sum_{\boldsymbol{v}\neq\boldsymbol{\lambda}} -(1-\alpha)(\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^{T} \cdot \delta\vec{\tau}_{\boldsymbol{v}} \right]. \tag{5.53}$$

We will now establish that both expressions, Eq. (5.51) and Eq. (5.53), are equivalent from a theoretical perspective. The second formulation explicitly incorporates the acoustic sum rule, as we will see later, which justifies its use in the numerical implementation within the SCALE-UP code instead of the original theoretical expression.

Mathematical correspondence between linear contributions in two equivalent geometries: single-atom displacement and collective atomic displacement.

Proof. To verify both expressions Eq. (5.51) and Eq. (5.53) are equivalent, we must prove the following equality,

$$\sum_{\boldsymbol{\lambda}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^{T} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} = \sum_{\boldsymbol{\lambda}} \left[\alpha(\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^{T} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} + \sum_{\boldsymbol{v}\neq\boldsymbol{\lambda}} -(1-\alpha)(\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^{T} \cdot \delta\vec{\tau}_{\boldsymbol{v}} \right]. \tag{5.54}$$

We start from the right-hand side and we simplify it step by step,

$$\sum_{\boldsymbol{\lambda}} \left[\alpha(\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^{T} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} + \sum_{\boldsymbol{v}\neq\boldsymbol{\lambda}} -(1-\alpha)(\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^{T} \cdot \delta\vec{\tau}_{\boldsymbol{v}} \right] = \sum_{\boldsymbol{\lambda}} \left[\frac{N_{\mathrm{at}}-1}{N_{\mathrm{at}}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^{T} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} - \frac{1}{N_{\mathrm{at}}} \sum_{\boldsymbol{v}\neq\boldsymbol{\lambda}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^{T} \cdot \delta\vec{\tau}_{\boldsymbol{v}} \right]. \tag{5.55}$$

By expanding and reordering the terms, we obtain,

$$\sum_{\boldsymbol{\lambda}} \left[\frac{N_{\mathrm{at}}-1}{N_{\mathrm{at}}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^{T} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} - \frac{1}{N_{\mathrm{at}}} \sum_{\boldsymbol{v}\neq\boldsymbol{\lambda}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{v}})^{T} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} \right] = \sum_{\boldsymbol{\lambda}} \left[(\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^{T} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} - \frac{1}{N_{\mathrm{at}}} \sum_{\boldsymbol{v}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{v}})^{T} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} \right]. \tag{5.56}$$

The sum $\sum_{\boldsymbol{v}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{v}})^{T}$ corresponds to the ASR condition. From a theoretical standpoint, it must be zero according to Eq. (5.48), leading to

$$\sum_{\boldsymbol{\lambda}} \left[(\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^{T} - \frac{1}{N_{\mathrm{at}}} \sum_{\boldsymbol{v}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{v}})^{T} \right] \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} = \sum_{\boldsymbol{\lambda}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^{T} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}}, \tag{5.57}$$

as we wanted to prove. ☐

Once we have shown both expressions are equivalent, the following step consists on proving that the ASR is satisfied automatically. Specifically, we aim to demonstrate that, by considering the previously described intermediate step when taking into account the electron-lattice correction, we obtain an expression that depends on displacement differences, $\delta\vec{\tau}_{\boldsymbol{\lambda v}}$, rather than absolute displacements, $\delta\vec{\tau}_{\boldsymbol{\lambda}}$. If the contribution from the linear terms can be rewritten as a function of displacement differences, the expression will inherently satisfy the acoustic sum rule. Then, we want to prove the equality

$$\sum_{\boldsymbol{\lambda}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^T \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} = \frac{1}{N_{\mathrm{at}}} \sum_{\boldsymbol{\lambda}} \sum_{\boldsymbol{\upsilon}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^T \cdot \delta\vec{\tau}_{\boldsymbol{\lambda\upsilon}}. \tag{5.58}$$

On the left-hand side, we have the contribution from the linear terms in Eq. (5.40), while on the right-hand side, we obtain the target expression, which explicitly depends on $\delta\vec{\tau}_{\boldsymbol{\lambda\upsilon}}$.

Equivalence of linear coupling contribution expressions using one-atom displacement $\delta\vec{\tau}_{\boldsymbol{\lambda}}$ and two-atom displacement $\delta\vec{\tau}_{\boldsymbol{\lambda\upsilon}}$.

Proof. At this point, we are going to prove Eq. (5.58). Employing Eq. (5.54), proving the equality in Eq. (5.58) is equivalent to prove

$$\sum_{\boldsymbol{\lambda}} \left[\alpha(\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^T \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} + \sum_{\boldsymbol{\upsilon}\neq\boldsymbol{\lambda}} -(1-\alpha)(\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^T \cdot \delta\vec{\tau}_{\boldsymbol{\upsilon}} \right] = \frac{1}{N_{\mathrm{at}}} \sum_{\boldsymbol{\lambda}} \sum_{\boldsymbol{\upsilon}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^T \cdot \delta\vec{\tau}_{\boldsymbol{\lambda\upsilon}}. \tag{5.59}$$

Considering Eq. (5.55),

$$\sum_{\boldsymbol{\lambda}} \left[\alpha(\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^T \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} + \sum_{\boldsymbol{\upsilon}\neq\boldsymbol{\lambda}} -(1-\alpha)(\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^T \cdot \delta\vec{\tau}_{\boldsymbol{\upsilon}} \right] = \sum_{\boldsymbol{\lambda}} \left[\frac{N_{\mathrm{at}}-1}{N_{\mathrm{at}}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^T \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} - \frac{1}{N_{\mathrm{at}}} \sum_{\boldsymbol{\upsilon}\neq\boldsymbol{\lambda}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^T \cdot \delta\vec{\tau}_{\boldsymbol{\upsilon}} \right]. \tag{5.60}$$

Grouping the terms with the prefactor $1/N_{\mathrm{at}}$, we obtain

$$\begin{aligned} \sum_{\boldsymbol{\lambda}} \left[\frac{N_{\mathrm{at}}-1}{N_{\mathrm{at}}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^T \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} - \frac{1}{N_{\mathrm{at}}} \sum_{\boldsymbol{\upsilon}\neq\boldsymbol{\lambda}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^T \cdot \delta\vec{\tau}_{\boldsymbol{\upsilon}} \right] &= \sum_{\boldsymbol{\lambda}} \left[(\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^T \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} - \frac{1}{N_{\mathrm{at}}} \sum_{\boldsymbol{\upsilon}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^T \cdot \delta\vec{\tau}_{\boldsymbol{\upsilon}} \right] = \\ = \frac{1}{N_{\mathrm{at}}} \sum_{\boldsymbol{\lambda}} \sum_{\boldsymbol{\upsilon}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^T \cdot \delta\vec{\tau}_{\boldsymbol{\lambda}} - \frac{1}{N_{\mathrm{at}}} \sum_{\boldsymbol{\lambda}} \sum_{\boldsymbol{\upsilon}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^T \cdot \delta\vec{\tau}_{\boldsymbol{\upsilon}} &= \frac{1}{N_{\mathrm{at}}} \sum_{\boldsymbol{\lambda}} \sum_{\boldsymbol{\upsilon}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^T \cdot (\delta\vec{\tau}_{\boldsymbol{\lambda}} - \delta\vec{\tau}_{\boldsymbol{\upsilon}}) = \\ = \frac{1}{N_{\mathrm{at}}} \sum_{\boldsymbol{\lambda}} \sum_{\boldsymbol{\upsilon}} (\vec{f}^{\,\mathrm{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^T \cdot \delta\vec{\tau}_{\boldsymbol{\lambda\upsilon}}. \end{aligned} \tag{5.61}$$

From the first to the second line, we have added a sum over $\boldsymbol{\upsilon}$ and then, the factor $1/N_{\mathrm{at}}$. The final expression gives the dependence of the linear contribution with the difference of displacements, satisfying inherently the ASR. □

Examining the second-order terms, we first observe that the number of $\overleftrightarrow{g}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}}$ variables in Eq. (5.38) and Eq.(5.40) is identical. In fact, it can be easily verified, as we will show here, that the value of $\overleftrightarrow{g}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}}$ in Eq. (5.38) is simply the corresponding value in Eq. (5.40) multiplied by -1.

To establish the relationship between the $\overleftrightarrow{g}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}}$ variables in Eq. (5.38) and Eq. (5.40), consider the displacements of two atoms, $\boldsymbol{\lambda}$ and $\boldsymbol{\upsilon}$. Furthermore, attending to Eq. (5.40), we find that in order to satisfy the ASR (which ensures no change in energy under a rigid shift of the atoms), the following condition must hold for the quadratic terms,

$$\sum_{\boldsymbol{\lambda}\neq\boldsymbol{\upsilon}} \overleftrightarrow{g}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}} = -\overleftrightarrow{g}_{\boldsymbol{ab},\boldsymbol{\upsilon\upsilon}}. \tag{5.62}$$

Since the values of $\overleftrightarrow{g}_{ab,\lambda\upsilon}$ for $\lambda \neq \upsilon$ can be easily computed using finite differences, imposing the ASR in the second-principles framework via Eq. (5.62) is straightforward.

According to Eq. (5.40), the values of the second order terms $\overleftrightarrow{g}_{ab,\lambda\upsilon}$ correspond to the obtained value from finite differences, $\overleftrightarrow{g}^{\text{fd}}_{ab,\lambda\upsilon}$. Thus, the correction to the Hamiltonian matrix element considering exclusively the second-order terms can be expressed as

$$\sum_{\lambda}\sum_{\upsilon}\frac{1}{2}\left[\delta\vec{\tau}_{\lambda}^{T}\cdot\overleftrightarrow{g}^{\text{fd}}_{ab,\lambda\lambda}\cdot\delta\vec{\tau}_{\lambda}+\delta\vec{\tau}_{\upsilon}^{T}\cdot\overleftrightarrow{g}^{\text{fd}}_{ab,\upsilon\upsilon}\cdot\delta\vec{\tau}_{\upsilon}+\delta\vec{\tau}_{\lambda}^{T}\cdot\overleftrightarrow{g}^{\text{fd}}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\upsilon}+\delta\vec{\tau}_{\upsilon}^{T}\cdot\overleftrightarrow{g}^{\text{fd}}_{ab,\upsilon\lambda}\cdot\delta\vec{\tau}_{\lambda}\right]. \tag{5.63}$$

Now, our objective is to relate the quadratic contribution as a function of absolute displacements, Eq. (5.38) and the expression associated to differences of displacements, Eq. (5.40).

Equivalence of quadratic coupling contribution expressions using one-atom displacement $\delta\vec{\tau}_{\lambda}$ and two-atom displacement $\delta\vec{\tau}_{\lambda\upsilon}$.

Proof. In order to compute the correction to the Hamiltonian matrix elements as a function of the differences between atomic displacements, it is necessary to establish the relationship between the second-order terms associated to absolute displacements $\overleftrightarrow{g}^{\text{fd}}_{ab,\lambda\upsilon}$, and the values considering differences of displacements, $\overleftrightarrow{g}_{ab,\lambda\upsilon}$. At this point, we compute the correction to the Hamiltonian matrix elements in terms of Eq. (5.38), i.e, considering differences of displacements,

$$\begin{aligned}
&\sum_{\lambda}\sum_{\upsilon\neq\lambda}\delta\vec{\tau}_{\lambda\upsilon}^{T}\cdot\overleftrightarrow{g}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\lambda\upsilon}=\\
&=\sum_{\lambda}\sum_{\upsilon\neq\lambda}\left(\delta\vec{\tau}_{\upsilon}-\delta\vec{\tau}_{\lambda}\right)^{T}\cdot\overleftrightarrow{g}_{ab,\lambda\upsilon}\cdot\left(\delta\vec{\tau}_{\upsilon}-\delta\vec{\tau}_{\lambda}\right)=\sum_{\lambda}\sum_{\upsilon\neq\lambda}\left(\delta\vec{\tau}_{\lambda}-\delta\vec{\tau}_{\upsilon}\right)^{T}\cdot\overleftrightarrow{g}_{ab,\lambda\upsilon}\cdot\left(\delta\vec{\tau}_{\lambda}-\delta\vec{\tau}_{\upsilon}\right)=\\
&=\sum_{\lambda}\sum_{\upsilon\neq\lambda}\delta\vec{\tau}_{\lambda}^{T}\cdot\overleftrightarrow{g}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\lambda}+\sum_{\lambda}\sum_{\upsilon\neq\lambda}\delta\vec{\tau}_{\upsilon}^{T}\cdot\overleftrightarrow{g}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\upsilon}-\\
&-\sum_{\lambda}\sum_{\upsilon\neq\lambda}\delta\vec{\tau}_{\lambda}^{T}\cdot\overleftrightarrow{g}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\upsilon}-\sum_{\lambda}\sum_{\upsilon\neq\lambda}\delta\vec{\tau}_{\upsilon}^{T}\cdot\overleftrightarrow{g}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\lambda}=\\
&=\sum_{\lambda}\delta\vec{\tau}_{\lambda}^{T}\cdot\left(\sum_{\upsilon\neq\lambda}\overleftrightarrow{g}_{ab,\lambda\upsilon}\right)\cdot\delta\vec{\tau}_{\lambda}+\sum_{\upsilon}\delta\vec{\tau}_{\upsilon}^{T}\cdot\left(\sum_{\lambda\neq\upsilon}\overleftrightarrow{g}_{ab,\lambda\upsilon}\right)\cdot\delta\vec{\tau}_{\upsilon}+\\
&+\sum_{\lambda}\sum_{\upsilon\neq\lambda}\delta\vec{\tau}_{\lambda}^{T}\cdot\overleftrightarrow{g}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\upsilon}-\sum_{\lambda}\sum_{\upsilon\neq\lambda}\delta\vec{\tau}_{\upsilon}^{T}\cdot\overleftrightarrow{g}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\lambda}=\\
&=2\sum_{\lambda}\delta\vec{\tau}_{\lambda}^{T}\cdot\left(\sum_{\upsilon\neq\lambda}\overleftrightarrow{g}_{ab,\lambda\upsilon}\right)\cdot\delta\vec{\tau}_{\lambda}-\sum_{\lambda}\sum_{\upsilon\neq\lambda}\delta\vec{\tau}_{\lambda}^{T}\cdot\overleftrightarrow{g}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\upsilon}-\sum_{\lambda}\sum_{\upsilon\neq\lambda}\delta\vec{\tau}_{\upsilon}^{T}\cdot\overleftrightarrow{g}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\lambda}=\\
&=\sum_{\lambda}\delta\vec{\tau}_{\lambda}^{T}\cdot\left(2\sum_{\upsilon\neq\lambda}\overleftrightarrow{g}_{ab,\lambda\upsilon}\right)\cdot\delta\vec{\tau}_{\lambda}-\sum_{\lambda}\sum_{\upsilon\neq\lambda}\delta\vec{\tau}_{\lambda}^{T}\cdot\overleftrightarrow{g}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\upsilon}-\sum_{\lambda}\sum_{\upsilon\neq\lambda}\delta\vec{\tau}_{\upsilon}^{T}\cdot\overleftrightarrow{g}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\lambda}=\\
&=-2\sum_{\lambda}\delta\vec{\tau}_{\lambda}^{T}\cdot\overleftrightarrow{g}_{ab,\lambda\lambda}\cdot\delta\vec{\tau}_{\lambda}-\sum_{\lambda}\sum_{\upsilon\neq\lambda}\delta\vec{\tau}_{\lambda}^{T}\cdot\overleftrightarrow{g}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\upsilon}-\sum_{\lambda}\sum_{\upsilon\neq\lambda}\delta\vec{\tau}_{\upsilon}^{T}\cdot\overleftrightarrow{g}_{ab,\lambda\upsilon}\cdot\delta\vec{\tau}_{\lambda}.
\end{aligned} \tag{5.64}$$

In the last step, the relation in Eq. (5.62) has been applied.

Comparing expression Eq. (5.63) and Eq. (5.64), they are equivalent when the values of the second expression satisfy

$$\overset{\leftrightarrow}{g}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}} = -\overset{\leftrightarrow}{g}{}^{\text{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}}. \tag{5.65}$$

□

The final expression for the short-range one-electron parameters is

$$\gamma^{\text{sr}}_{\boldsymbol{ab}} = \gamma^{\text{RAG,sr}}_{\boldsymbol{ab}} - \frac{1}{N_{\text{at}}} \sum_{\boldsymbol{\lambda}} \sum_{\boldsymbol{\upsilon}\neq\boldsymbol{\lambda}} (\vec{f}^{\text{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}})^{T} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda\upsilon}} - \sum_{\boldsymbol{\lambda}} \sum_{\boldsymbol{\upsilon}\neq\boldsymbol{\lambda}} \delta\vec{\tau}^{T}_{\boldsymbol{\lambda\upsilon}} \cdot \overset{\leftrightarrow}{g}{}^{\text{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}} \cdot \delta\vec{\tau}_{\boldsymbol{\lambda\upsilon}}. \tag{5.66}$$

The symmetry relations established in Sec. 5.4 remain valid for this redefinition of the electron-lattice parameters.

Building upon the previous reasoning, we have developed a methodology in which the electron-lattice parameters directly correspond to the terms in the Taylor series expansion. Moreover, this approach inherently incorporates the Acoustic Sum Rule (ASR) into the electron-lattice corrections. The methodology is based on the theoretical linear and quadratic terms, which can be systematically computed using the finite difference method. At this stage, we present the procedure for determining these parameters and describe the first-principles data required to construct the required training set.

5.6.3.1 Finite difference formulas for the calculation of parameters

Electron-lattice coupling terms are computed using finite-difference approximations applied to the one-electron Hamiltonian obtained from density functional theory (DFT). Specifically, the x-component of the linear electron-lattice coupling vector, $\vec{f}^{\text{fd}}_{\boldsymbol{ab},\boldsymbol{\lambda}}$, is estimated using the central difference formula

$$f^{\text{fd},x}_{\boldsymbol{ab},\boldsymbol{\lambda}} \approx -\frac{h^{\text{DFT}}_{\boldsymbol{ab},\boldsymbol{\lambda}+x} - h^{\text{DFT}}_{\boldsymbol{ab},\boldsymbol{\lambda}-x}}{2\delta_x}, \tag{5.67}$$

where the terms $h^{\text{DFT}}_{\boldsymbol{ab},\boldsymbol{\lambda}\pm x}$ denote the DFT one-electron Hamiltonian matrix elements associated with the Wannier functions $(\boldsymbol{a},\boldsymbol{b})$. These elements correspond to geometries in which the atom $\boldsymbol{\lambda}$ is shifted by a distance $\pm\delta_x$ along the x-direction, relative to its reference geometry positions, remaining the rest of the atoms fixed at their reference geometry positions. The displacement distance δ_x serves as the step size for the finite difference formulas.

Similarly, the second-order parameters are computed following

$$g^{xy}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}} \approx \frac{h^{\text{DFT}}_{\boldsymbol{ab},\boldsymbol{\lambda}+x\ \boldsymbol{\upsilon}+y} - h^{\text{DFT}}_{\boldsymbol{ab},\boldsymbol{\lambda}+x\ \boldsymbol{\upsilon}-y} - h^{\text{DFT}}_{\boldsymbol{ab},\boldsymbol{\lambda}-x\ \boldsymbol{\upsilon}+y} + h^{\text{DFT}}_{\boldsymbol{ab},\boldsymbol{\lambda}-x\ \boldsymbol{\upsilon}-y}}{4\delta_x\delta_y}. \tag{5.68}$$

To compute the second-order terms, the first-principles calculations involve the simultaneous displacement of two atoms, $\boldsymbol{\lambda}$ and $\boldsymbol{\upsilon}$. In this context, the DFT one-electron Hamiltonian matrix element for the pair of Wannier functions $(\boldsymbol{a},\boldsymbol{b})$, $h_{\boldsymbol{ab},\boldsymbol{\lambda}\pm x\ \boldsymbol{\upsilon}\pm y}$, corresponds to the geometry in which the atom $\boldsymbol{\lambda}$ is displaced by a quantity $\pm\delta_x$ along the x coordinate, while the atom $\boldsymbol{\upsilon}$ is moved a value of $\pm\delta_y$ along the y-direction.

During the development of the computational framework for the quadratic variables, an alternative formulation was implemented, wherein the calculations performed for the evaluation of the linear terms were reutilized. This approach resulted in a reduction of the number of new first-principles calculations required to compute $\overset{\leftrightarrow}{g}$ values. Specifically, the expression

$$g^{xy}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}} \approx \frac{h^{\mathrm{DFT}}_{\boldsymbol{ab},\boldsymbol{\lambda}+x\ \boldsymbol{\upsilon}+y} - h^{\mathrm{DFT}}_{\boldsymbol{ab},\boldsymbol{\lambda}+x\ \boldsymbol{\upsilon}} - h^{\mathrm{DFT}}_{\boldsymbol{ab},\boldsymbol{\lambda}\ \boldsymbol{\upsilon}+y} + 2h^{\mathrm{DFT}}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}} - h^{\mathrm{DFT}}_{\boldsymbol{ab},\boldsymbol{\lambda}-x\ \boldsymbol{\upsilon}} - h^{\mathrm{DFT}}_{\boldsymbol{ab},\boldsymbol{\lambda}\ \boldsymbol{\upsilon}-y} + h^{\mathrm{DFT}}_{\boldsymbol{ab},\boldsymbol{\lambda}-x\ \boldsymbol{\upsilon}-y}}{2\delta_x\delta_y} \tag{5.69}$$

was employed. Nevertheless, this alternative method led to higher errors in the calculation of the second-order electron-lattice coupling parameters.

5.6.3.2 Training set and distance cutoff

At this stage, it is required to build the training set for the electron-lattice parameters in *methodology II*. As previously explained, the elements $h^{\mathrm{DFT}}_{\boldsymbol{ab},\boldsymbol{\lambda}\pm x}$ and $h^{\mathrm{DFT}}_{\boldsymbol{ab},\boldsymbol{\lambda}\pm x\,\boldsymbol{\upsilon}\pm y}$ necessary to compute the linear and quadratic electron-lattice interactions are obtained from density function theory calculations in the selected *supercell* (see Sec. 5.6). However, direct applications of formulas Eq. (5.67) and, particularly, Eq. (5.68) would lead to a very large number of first-principles calculations. To generate these geometries, several configurations are constructed starting from the reference atomic configuration, by displacing one (linear terms) or two atoms (quadratic terms). These geometries are systematically created by considering all possible combinations of displacements: taking into account the Cartesian axes and the positive and negative senses of the displacement. The total number of geometry combinations (derived from the number of atom pairs, Cartesian directions, and displacement senses) can result in a substantial number first-principles calculations. For example, in the $\mathrm{SrTiO_3}$ system, using a $2\times2\times2$ *supercell* with five atoms per unit cell, the number of required first-principles calculations for the linear terms is

$$5 \text{ atoms/cell} \times 8 \text{ cells} \times 3 \text{ directions} \times 2 \text{ senses} = 2400 \text{ geometries.}$$

Regarding to the quadratic terms, the amount of geometries is the square of the linear terms calculations.

We apply a strategy to reduce the number of first-principles calculations:

Firstly, given the geometry of a specific configuration required in the training set, the symmetry operations of the space group associated with the RAG are applied to identify all geometrically equivalent configurations, avoiding then their explicit calculation (see Sec. 5.4).

Secondly, translational symmetry can be leveraged in the linear terms to further minimize the computational effort. The chosen approach for handling the linear variables involves generating, exclusively, the specific geometries to compute the linear terms associated with the atoms in the reference *unit cell*. Once these parameters are obtained, the corresponding values for atoms in the remaining unit cells of the *supercell* are assigned through translational mapping of the selected *unit cell*.

The third approach addresses the quadratic terms, which represent the most computationally demanding components, even applying symmetry constraints, due to their scaling with the square of the number of atoms. To mitigate this complexity, it is assumed that the displacements of two atoms separated by a sufficiently large distance are uncorrelated, and their contributions can be effectively captured by the linear terms. Based on this assumption, the computational framework introduces a cutoff distance that defines the maximum separation (at the RAG) allowed between atom pairs in the *supercell* for evaluating second-order terms. If the interatomic distance exceeds this cutoff , denoted as δr_{el} [represented in green in Fig. 5.7(a)],

$$\delta r_{\mathrm{el}} > \left| \vec{R}_{\Upsilon} - \vec{R}_{\Lambda} + \vec{\tau}_{\upsilon}^{(0)} - \vec{\tau}_{\lambda}^{(0)} \right| , \tag{5.70}$$

the corresponding quadratic term is omitted and treated as negligible. Collectively, these three strategies significantly reduce the number of first-principles calculations required to determine the electron-lattice parameters while preserving the accuracy.

5.6.3.3 Energy cutoffs

Despite the previously implemented strategies to reduce the number of directly computed parameters, the amount of linear and quadratic constants calculated in the procedure is still very large, with many of them having a very small influence on the value of the $h_{\boldsymbol{ab}}$ elements. This sheer volume of parameters can make second-principles calculations computationally demanding, primarily due to the extensive number of linear and quadratic terms involved. Thus, when building a model, we only include vectors $\vec{f}_{\boldsymbol{ab},\boldsymbol{\lambda}}$ and matrices $\overset{\leftrightarrow}{g}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}}$ where *any* of its components are above the cutoffs δf and δg measured, respectively, in eV/Å and eV/Å^2. In general we find that good results can be found when these cutoffs take values around 0.1 eV/Å and 0.1 eV/Å^2 (see the results chapter, Chapter 6, for more details).

Metals are more difficult to treat than insulators as their density matrix in the ground state changes with the geometry. Thus, in these cases we first obtain the value of $U_{\boldsymbol{ab},\boldsymbol{a'b'}}$ (and also $I_{\boldsymbol{ab},\boldsymbol{a'b'}}$, although this is not always necessary) and correct the value of $h_{\boldsymbol{ab}}^{\mathrm{DFT}}$ as to avoid double-counting the effect of the change of geometry and the density-change when obtaining $\vec{f}_{\boldsymbol{ab},\boldsymbol{\lambda}}$ and $\overset{\leftrightarrow}{g}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}}$.

5.6.4 Verification of the models

To evaluate the performance of our models, we employ an error function similar to Θ, defined in Eq. (5.25), applied to a *test-set* of density functional theory (DFT) calculations, denoted as $\{A\}$. These test sets consist of structures not included in the training set, selected to probe the model's ability to generalize to configurations representative of physically relevant distortions. Specifically, test geometries are generated by displacing atoms randomly within a cube of side length $2d$, centered at the relaxed atomic geometry (RAG), where d is a parameter controlling the displacement amplitude. In Sec. 6.3 and Sec. 6.4, the error function Θ is evaluated by averaging over ten such randomly distorted geometries for each value of d. Additionally, when available, we assess model performance on alternative stable phases not included in the training. For example, in the case of $SrTiO_3$, we include both the displacements present on the tetragonal $I4/mcm$ phase and a ferroelectric phase stabilized under uniaxial strain, keeping in the calculations the lattice in its cubic structure. In other systems, we explore the formation of polarons by introducing lattice distortions and comparing the resulting electronic and structural responses—governed by both electron-lattice and electron-electron interactions—with those obtained from DFT calculations.

The methodology presented in this Chapter is the result of a significant amount of work and sustained effort, grounded in extensive testing and iterative refinement. Throughout its development, various strategies were explored to identify the most suitable training set for the reliable extraction of model parameters. Considerable attention was also devoted to designing and optimizing the methodology itself, with a focus on improving both its accuracy and computational efficiency. Moreover, the approach was carefully constructed to incorporate key physical constraints, such as the acoustic sum rule and the symmetry properties of the system, in order to preserve the intrinsic characteristics of the material under study.

In summary, this Chapter represents the culmination of a long process of trial and error, guided by both physical insight and practical considerations. A selection of the tests performed and the results obtained will be discussed in Chapter 6.

Chapter 6

Results

To demonstrate the application of the methodologies described above, we consider two representative systems. The first is the semicovalent transition-metal perovskite $SrTiO_3$, a material that exhibits strong electron-lattice coupling arising from two key structural instabilities: the incipient ferroelectric distortion [43; 182], and the antiferrodistortive octahedra rotations associated with the cubic-to-tetragonal phase transition. The second example is LiF, a wide-band-gap insulator with the rock-salt structure. In this system, the strong electron-hole interaction leads to prominent optical effects, including the formation of tightly bound excitons [183; 184; 185].

In this discussion, we present an analysis of the results obtained for the determination of the model parameters, focusing on the methodologies outlined in Chapter 5. For the electron-lattice coupling parameters, two approaches are considered: *methodology I*, described in Sec. 5.6.2, which involves fitting the parameters to reference data; and the *methodology II*, detailed in Sec. 5.6.3, in which the electron-lattice parameters are computed directly from first-principles calculations. We will demonstrate how *methodology I* fails to accurately reproduce certain relevant phases in the analysis of electron-lattice coupling, thereby justifying the introduction of *methodology II*.

In contrast, for the electron-electron interaction parameters we apply *methodology I*.

The procedure to obtain the model parameters includes the selection of the relevant Wannier manifolds; the choice of the Hamiltonian interaction distance cutoff, δr_{h}; the electron-electron interaction cutoff, δr_{ee}; and the corresponding goal-function convergence threshold, $\delta\Theta$. For the electron-lattice interaction, we additionally stablish the spatial and energy cutoffs according to the applied methodology.

6.1 Computational details

Before presenting the obtained results, the computational details of the calculations are described in this Section.

6.1.1 SIESTA

First-principles calculations were performed using SIESTA [38; 174], which is based on the standard Kohn-Sham self-consistent formulation of density functional theory (DFT) and employs a numerical atomic orbital (NAO) basis set.

In order to test the method we employed different exchange-correlation functionals for each of the studied systems. For $SrTiO_3$ the exchange-correlation functional was approximated within the local density approximation (LDA), adopting the usual Ceperley and Alder [186; 187] parametrization. In the case of LiF, the exchange and correlation was treated using the Perdew-Burke-Ernzerhof (PBE) approximation [188].

For $SrTiO_3$, the parameters that define the shape and spatial extent of the basis functions were determined by computing the eigenfunctions of isolated atoms confined within the soft-confinement potential proposed in Ref. [189]. We used a single-ζ basis set for the semicore states of Ti and Sr, and double-ζ plus polarization for the valence states of all the atoms. For Sr an extra shell of $4d$ orbitals was added. All the parameters that define the shape and the range of the basis functions for Sr, Ti, and O were obtained by a variational optimization following the recipe given in Ref. [189].

For the LiF system, a double-ζ basis set was employed for the $2s$ and $2p$ orbitals of Li, while a double-ζ polarized basis set was selected for F. In both cases, the spatial extent of the orbitals was defined using the default parameters.

The core electrons were replaced by *ab initio* norm-conserving pseudopotentials [190]. The pseudopotentials have been generated using the Troullier-Martins scheme [191] in the Kleinman-Bylander fully non-local separable representation [192].

The electronic density, Hartree potential, and exchange-correlation potential, along with the corresponding matrix elements between the basis orbitals, were computed on a uniform real-space grid [38]. To accurately represent the charge density in both systems, an equivalent plane-wave cutoff of 600 Ry was employed. For the reciprocal-space integration, considering the conventional cell geometries of the systems, the integrals were ensured to be well converged. In all cases, a Monkhorst-Pack mesh of equivalent quality was utilized, specifically a $9 \times 9 \times 9$ grid for the $SrTiO_3$ system and a $8 \times 8 \times 8$ grid for the LiF system [71]. The RAG structure was relaxed for both systems until the maximum component of the stress tensor was below 0.003 eV/Å^3.

The second-principles code SCALE-UP and, by extension, the program generating the models, MOD-

ELMAKER, employ a basis of Wannier functions [65; 68], as this provides a compact and short-range basis that is ideal to create large-scale methods. These functions are constructed by projecting Bloch states onto SIESTA's atomic basis orbitals, which serve as initial guess functions, as implemented in the WANNIER90 code [39; 40]. To avoid problems with the symmetry arising from the minimization of the spread, the WFs were directly taken from these projections according to Sec. 1.4.5. A uniform $4 \times 4 \times 4$ reciprocal-space mesh was used to compute the Wannier functions and evaluate relevant quantities in real space.

6.1.2 MODELMAKER

MODELMAKER is a Python software package designed as an integrated tool for generating and utilizing second-principles models—specifically, polynomial expansions—that incorporate various electronic degrees of freedom, enabling large-scale simulations of up to several hundred thousand atoms. It is released as part of the SCALE-UP code [37] and employs the SIESTA code [38; 174], which includes an interface with WANNIER90 [39], to perform the first-principles simulations in a basis of Wannier functions. The first-principles calculations conform the training-set for both electron-lattice and electron-electron parameters and these simulations are stored in a database (`mm_database`). Since MODELMAKER employs the results of *ab initio* DFT calculations conducted with SIESTA to generate second-principles models, the code leverages the precision of SIESTA in electronic structure properties. While this initial implementation of MODELMAKER is linked to SIESTA, the code could be extended to any first-principles software that integrates WANNIER90.

The output of MODELMAKER is a model which includes information on bands, and electron-lattice and electron-electron interactions. The model is stored in human-readable `.xml` format for its use by the second-principles code, SCALE-UP [37].

6.1.2.1 Input data and workflow of the code

The initial ingredients required to build an electronic model using the MODELMAKER software are a Python script, `run_mm.py`, which controls the workflow of the code through its various stages, and a folder, `files`, that stores the physical information characterizing the system under study. Below, we provide an explanation of each of these components.

The Python script `run_mm.py` serves as (i) the interface where users specify SIESTA and SCALE-UP executable files; (ii) the entry point where the user sets the cutoffs that regulate the model's accuracy; (iii) a mechanism for invoking MODELMAKER's functions to build the second-principles model and define the code workflow of the code. In this Section, we will focus on outlining the flowchart for building second-principles models, as described in point (iii), while the selection and description of cutoffs are addressed below.

The execution of the MODELMAKER code begins with running the SIESTA calculation using the WANNIER90 interface for the *unit cell* system in the RAG within the RED. This calculation is employed to compute the reference Hamiltonian, expressed in the Wannier function basis set, which will be used in several steps of the model generation process.

The second step consists of recalculating the reference atomic geometry (RAG) in the RED using SIESTA, this time considering the selected *supercell* that will be employed to construct the training set.

Next, the collection of calculations composing the training set are computed employing the SIESTA software. The configurations are run in the *supercell*. On one hand, the training set used to compute the electron-lattice parameters is created by modifying the system's geometry relative to the RAG, considering random atomic displacements if we want to fit the parameters, *methodology I*, or considering the geometries required for the finite difference formulas, *methodology II*. These calculations are performed while keeping the system in the RED. For *methodology I*, we establish the radii of the random displacements as 0.1 Å and we vary the values of the δr_{el} and $\delta r_{\mathrm{ellat}}$ cutoffs. In the case of *methodology II*, the only variables that need to be selected are the finite difference step size, 0.04 Borhs, for both systems, and the cutoff distance δr_{el}, which defines the maximum distance between atoms $\boldsymbol{\lambda}$ and $\boldsymbol{\upsilon}$ for which a $\overleftrightarrow{g}_{\boldsymbol{ab},\boldsymbol{\lambda\upsilon}}$ variable is computed.

On the other hand, the calculations that make up the training set for fitting the electron-electron parameters are defined. These calculations are characterized by the density matrix, which is modified from the RED. In the `run_mm.py` file, the net charge (doping with holes or electrons) in units of e^{-} and the spin of the system, which leads to a spin-polarized calculation, are specified.

Once the training set is built, the electron-lattice parameters are selected following *methodology I* (see Sec. 5.6.2) according to the established electron-lattice cutoffs; or computed employing the finite differences procedure, *methodology II*, explained in Sec. 5.6.3. In addition, the symmetry of the system (see Sec. 5.4) is considered. Regarding the electron-electron cutoffs, they determine the possible electron-electron terms which will later be computed and fitted to the training set according to *methodology I*.

The next step involves calling the SCALE-UP code, using a model for the RAG and the RED to calculate the long-range electrostatic corrections and the density matrix for the different calculations in the training set.

Then, in the application of the *methodology I*, the electron-lattice and electron-electron parameters are fitted to their training sets to identify the optimal model. If we are working on the *methodology II* approach, the electron-lattice parameters are recalculated, incorporating the electrostatic corrections.

The final step involves generating the `.xml` file, in which the model stores various quantities that characterize the physical properties of the system, extracted from the folder `files`. Additionally, it includes the one-electron Hamiltonian expressed in the Wannier function basis set for the RAG and RED, the system's density matrix, and the electron-lattice and electron-electron parameters.

Once the MODELMAKER workflow has been explained, we will proceed to describe the data stored in folder `files`.

The folder `files` contains the following files:

- RAG of the system, stored in `SystemLabel.XV`. The reference geometry provided to MODELMAKER should correspond to the smallest conventional cell considered. The selection of the conventional cell (which sometimes coincides with the primitive cell) instead of the primitive cell is crucial, as the symmetry relations are established based on the axes of the conventional cell. The designation for this geometry is *unit cell* (see Sec. 5.1), abbreviated as *uc*. For example, considering the $SrTiO_3$ system, the chosen RAG is its cubic centrosymmetric paraelectric phase, characterized by a lattice parameter of 3.874 Å. For the LiF system, the selected RAG is the conventional cell, which corresponds to a face-centered cubic (FCC) structure with 8 atoms per unit cell. The lattice parameter is 4.026 Å.
- SIESTA input file. A template of the SIESTA input file is defined for the system for which the electronic model is going to be constructed. Specifically, for the crystals under study, the input file includes the features outlined in Sec. 6.1.1. Additionally, this file specifies the instructions for Wannierizing the band manifolds of interest. In the particular cases of $SrTiO_3$ and LiF, the details of these calculations are provided in Secs. 6.3 and 6.4 respectively.
- The required pseudopotentials for all the chemical species (`.psf` or `.psml` files) for the corresponding functional and the selected basis set. According to the solid state systems under study, the selected pseudopotentials are described in Sec. 6.1.1.
- The symmetry operations associated to the point group of the selected RAG. These symmetry operations are essential for establishing the symmetry relations between the Hamiltonian matrix elements and the model parameters, according to Sec. 5.4. To get the symmetry operations, we shall make use of the Bilbao Crystallographic Server [175] selecting the point group of the desired structure. In particular, the point group of the $SrTiO_3$ system is Pm-3m while the point group of LiF is Fm-3m.
- The Born effective charges of the atoms in the system, `SystemLabel.born`. The Born effective charges quantify the coupling between atomic displacements and the macroscopic electric field in a crystal. For each atom in the crystal, they can be computed employing the SIESTA code.
- The electronic dielectric tensor, `SystemLabel.diel`. The electronic dielectric tensor is required for a proper treatment of the electrostatics of the system. It can be accessed from a DFPT calculation, as implemented in the ABINIT code.
- A file pointing to all the previous files that MODELMAKER uses to generate the model, `setup.yml` file. In addition, this file includes the size of the *supercell* where the training set calculations will be performed, the type of the Bravais lattice and the number of electrons in the WFs.
  For the $SrTiO_3$ system, a $2 \times 2 \times 2$ *supercell* is used to perform the training set calculations. A supercell is necessary to accommodate certain phases that characterize this complex system, such as

the oxygen octahedra rotations in the perovskite structure (antiferrodistortive phase), which cannot be accurately modeled with a simulation box containing only five atoms per unit cell. In the case of LiF, the selected *unit cell* serves as the *supercell*. As previously explained, in this system the selected *unit cell* is the conventional cell which in this case represents a supercell $2 \times 2 \times 2$ of the primitive cell.

To establish the number of electrons in the Wannier functions, for both systems the WFs have been derived from the Bloch states at the top of the valence band (VB) and the bottom of the conduction band (CB). Since both crystals are insulating, the top of the valence band is fully occupied, with two electrons per band, while the bottom of the conduction band remains unoccupied. As a result, the total number of valence electrons in $SrTiO_3$ is 18 electrons, as the top of the valence band consists of 9 bands. Similarly, for LiF, 12 valence bands correspond to a total of 24 electrons.

6.2 Symmetry

The incorporation of symmetry in the MODELMAKER code, as compared to the implementation in Ref. [37], represents a significant advancement in the efficiency and accuracy of second-principles models. Specifically, the inclusion of symmetry substantially reduces the computational cost associated with model generation across multiple sections of the code.

Firstly, the application of the symmetry constraints explained in Sec. 5.4 allows us reducing the number of calculations to obtain the electron-lattice matrix elements. In $SrTiO_3$ for *methodology II* the number of single-point calculations necessary to create the electron-lattice model is 57840 without the use of symmetry. This is reduced to 15 when using symmetry and an electron-lattice cutoff $\delta r_{\text{el}} = 2.0$ Å, or 209 if $\delta r_{\text{el}} = 5.6$ Å (see Sec. 6.3). In LiF, a similar reduction is achieved going from 2353 single-point calculations without symmetry to 21 considering the symmetry of the system and $\delta r_{\text{el}} = 3.0$ Å.

Secondly, the incorporation of symmetry enables the direct calculation of a subset of parameters, with the remaining parameters being determined by applying the corresponding symmetry relations.

In addition to enhancing computational efficiency, the inclusion of symmetry also improves the accuracy of the results, such as the reproduction of degeneracies observed in first-principles electronic band structures as we will show in the following sections.

6.3 $SrTiO_3$

The first system we consider is $SrTiO_3$, a widely studied transition-metal perovskite that serves as a common substrate for the epitaxial growth of other perovskites, such as ferroelectric $PbTiO_3$ [193; 194]. $SrTiO_3$ is also a key component in the design of complex oxide superlattices exhibiting intricate polarization patterns [195] and has been associated with emergent metallic behavior at interfaces [196].

The reference atomic geometry (RAG) adopted for this study corresponds to the high-symmetry cubic paraelectric phase (space group Pm-3m), which contains five atoms per unit cell. Using the SIESTA code within the local density approximation (LDA), the equilibrium lattice parameter is found to be 3.874 Å.

The electronic structure of SrTiO$_3$ is characterized by an insulating gap separating the O($2p$)-derived valence bands and Ti($3d$)-derived conduction bands. As shown in Fig. 6.1, the top of the valence band is dominated by O($2p$) states, while the bottom of the conduction band consists primarily of Ti($3d_{xy}$), Ti($3d_{xz}$), and Ti($3d_{yz}$) orbitals forming a t_{2g} manifold, as expected for isolated TiO$_6$ octahedra. While the experimental indirect gap is approximately 3.25 eV [197], the DFT-LDA result yields a smaller value near 1.7 eV (see Fig. 6.1). To construct a second-principles model capable of capturing essential physical phenomena in SrTiO$_3$, including polaron formation, we build a Hamiltonian incorporating Wannier functions that describe both the valence band and the t_{2g} conduction manifold. In this context, as outlined in Sec. 5.3, second-principles simulations require the construction of Wannier functions derived from the Bloch states associated to two distinct manifolds. In the computational framework, the *first* manifold corresponds to the top of the VB, associated to the $2p$ oxygen states (shown in blue in the PDOS representation, Fig.6.1). It is reasonable to use the nine $2p$ orbitals associated with the three oxygen atoms in the unit cell (three $2p$ orbitals per atom) as the basis for projection to obtain the Wannier functions. The *second* manifold encompasses the energy range associated with the bottom of the CB, dominated by a t_{2g} character associated to Ti $3d$ (represented by green in the PDOS plot, Fig. 6.1). Then, d_{xy}, d_{yz}, d_{xz} orbitals are selected as seed to express the states of the conduction band in Wannier functions basis set. Additionally, a *third* manifold is introduced, which integrates both the first and second manifolds. This third manifold is essential for investigating the electronic transitions between the first and second manifolds. As mentioned in Sec. 5.3 we employ SIESTA to wannierize the valence and conduction bands individually and then combine both to calculate the position operator elements $\vec{r}_{ab}$.

In defining the second manifold, a challenge arises due to the necessity of applying a *disentanglement* procedure (see Sec. 1.6) [69]. As shown in Fig. 6.1, where the band structure of the RAG for the unit cell is depicted, the t_{2g} bands—belonging to the second manifold—overlap in energy within the Brillouin zone with the e_g bands, corresponding to Ti($3d_{z^2}$), Ti($3d_{x^2-y^2}$) atomic orbitals. This overlap is particularly evident in Fig. 1.8, which presents the RAG bands for a $2 \times 2 \times 2$ supercell. At the Γ point, this crossing is clearly visible.

To effectively separate the three t_{2g} bands from the e_g bands, two energy windows are defined: (i) the *inner energy window* (frozen manifold), highlighted in Fig. 1.7, encompasses the three t_{2g} bands within the energy range of -6.0 eV to -3.0 eV. (ii) The *outer energy window* (disentanglement window) is selected to include all five $3d$ bands, ensuring that the t_{2g} bands remain frozen within the inner window. This outer window spans an energy range from -6.0 eV to 2.0 eV. Since the third manifold contains the second, a similar disentanglement procedure must be applied, represented in Fig. 6.1(a). In this case, both the inner and outer energy windows must include both the valence band (VB) and conduction band (CB). Consequently, the inner energy window extends from -15.0 eV to -3.0 eV, while the outer window covers the range from -15.0 eV to 2.0 eV.

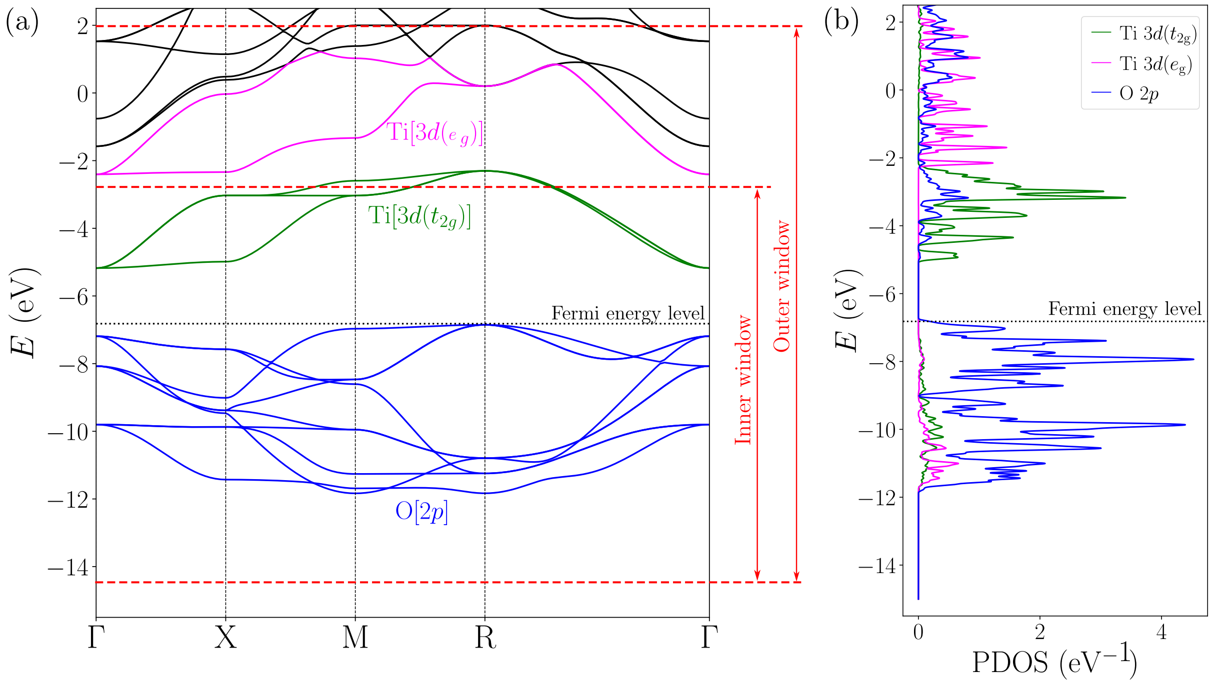


Figure 6.1: (a) Electronic band structure of $SrTiO_3$ at the cubic RAG where the valence band manifold with O($2p$) character is represented in blue and the conduction Ti[$3d(t_{2g})$] and Ti[$3d(e_g)$] manifolds are displayed in green and magenta, respectively. The global windows used during the Wannierization procedure are indicated with dashed red lines. (b) Projected density of states.

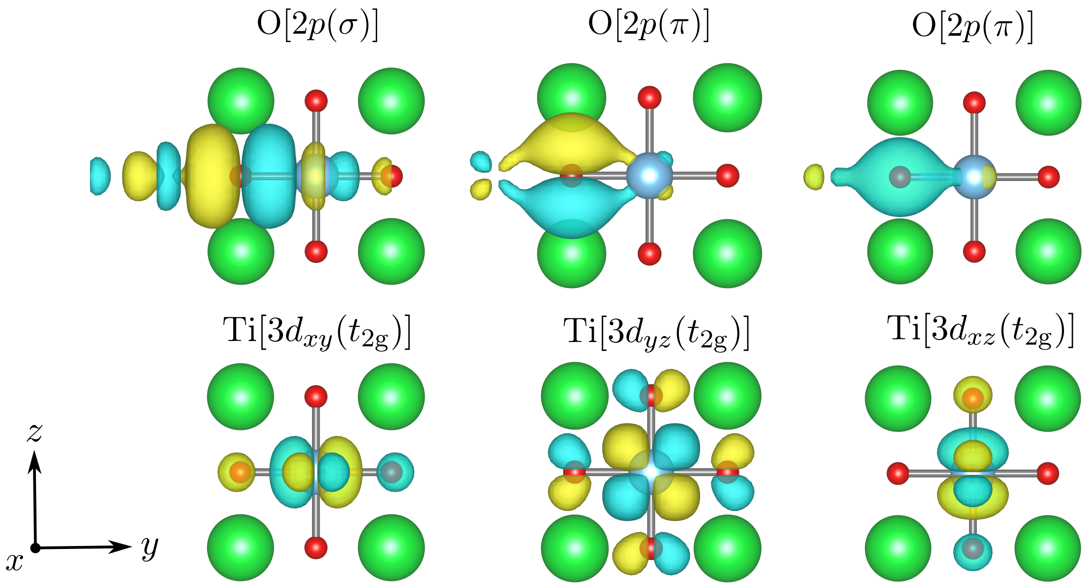


Figure 6.2: Wannier functions of $SrTiO_3$ used as the basis for second-principles models that accurately reproduce the top of the valence band with predominant O($2p$) character, as well as the lower conduction bands associated with Ti(t_{2g}) orbitals. Green, blue, and red spheres represent Sr, Ti, and O atoms, respectively.

For Wannier function calculations and the evaluation of the Hamiltonian and position matrix elements in real space, a $\vec{k}$-point sampling $4 \times 4 \times 4$ of the Brillouin zone is employed for a unit cell. The Wannier functions in this quasi-atomic representation, represented in Fig. 6.2, exhibit compact localization, with spreads of 1.17 Å^2 for O[$2p(\sigma)$] Wannier functions and 1.28 Å^2 for O[$2p(\pi)$] Wannier functions. In contrast, the t_{2g} Wannier functions are less localized, with spreads of 1.74 Å^2.

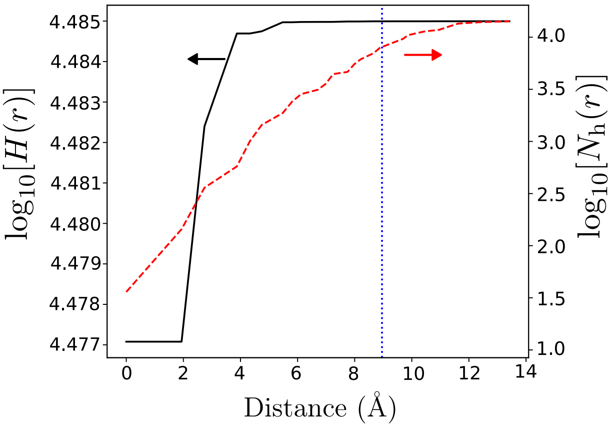


Figure 6.3: Black solid line (left axis): Decimal logarithm of the cumulative function $H(r)$, quantifying the contribution of one-electron Hamiltonian elements as a function of the inter-Wannier distance r. Red dashed line (right axis): Decimal logarithm of the number of Hamiltonian matrix elements with centroid separation less than r. The vertical dotted line marks the chosen one-electron Hamiltonian cutoff distance, δr_h = 9.0 Å, used in the SrTiO$_3$ model to reproduce the band structure shown with dashed red lines in Fig. 6.1(a).

To determine an appropriate value for the one-electron interaction cutoff distance δr_h, we compute the function

$$H(r) = \sum_{|\vec{r}_b - \vec{r}_a| < r} h^2_{\boldsymbol{ab}}(r), \tag{6.1}$$

which sums the squared Hamiltonian matrix elements between pairs of Wannier functions $\boldsymbol{a}$ and $\boldsymbol{b}$ with centroids separated by less than distance r. We can observe in Fig. 6.3 that the analysis of the black curve $H(r)$ can be divided in four regions that we will define over the rest lines. In support of this explanation, we shall employ Fig. 6.4. This picture illustrates a scheme of the SrTiO$_3$ system projected in the z-plane. By selecting the Ti atom of the home unit cell placed in the center of the drawing, we color different shells in the z-plane classifying the interactions according to the neighbors' distances (we neglect the Sr atom which does not present Wannier functions in the active set). We expect that the Hamiltonian matrix elements should decay fast with the distances between the Wannier functions. This effect is represented in the figure by decreasing the intensity of the orange color with the distance.

RANGE 0–1.9 Å. This region encompasses distances from 0 Å to 1.9 Å. Within this range, the initial goal function remains approximately flat over the entire interval. This behavior arises from the crystal structure of SrTiO$_3$. In the RAG, the lattice constant is $a = 3.874$ Å [198], resulting in a nearest-neighbor distance of $a/2 = 1.973$ Å. Given the ionic character of the SrTiO$_3$ crystal, the Wannier functions are strongly localized and centered on individual atoms. Consequently, for distances below $a/2$, interactions between WFs centered on different atoms are negligible. In this short-range regime, only on-site interactions—those occurring among WFs centered on the same atom—are relevant. This on-site terms range is depicted in Fig. 6.4 as intense orange area, corresponding to the Ti atom in the reference unit cell.

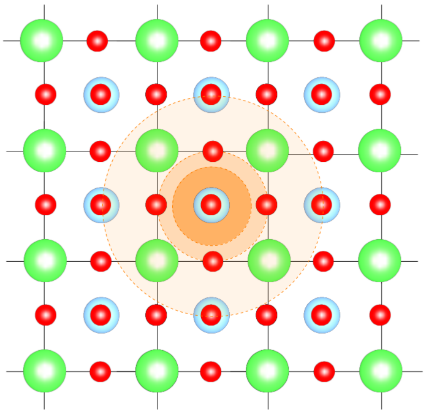


Figure 6.4: Different shells classifying interaction of the Wannier functions of Ti atom in the home unit cell with the Wannier functions of its neighbors. The more intense the orange color, the stronger the interaction between the orbitals.

RANGE 1.9–4 Å. The second region spans from 1.9 Å to 4 Å, where a marked increase in the initial goal function is observed. This sharp rise reflects the emergence of non-zero Hamiltonian matrix elements. In particular, for the studied WFs [O $(2p)$ and Ti (t_{2g}) orbitals], the following interactions become relevant:

- FIRST NEIGHBOR SHELL: Interactions between WFs located on Ti and O atoms within the TiO_6 complex, separated by $a/2 = 1.973$ Å.
- SECOND NEIGHBOR SHELL: Interactions between WFs on O atoms within the same unit cell, with a separation of $a/\sqrt{2} = 2.739$ Å.
- THIRD NEIGHBOR SHELL: Ti–Ti interactions between atoms located at $[\pm1,0,0]a$, $[0,\pm1,0]a$, and $[0,0,\pm1]a$ positions, corresponding to a distance of $a = 3.874$ Å.
- THIRD NEIGHBOR SHELL: Analogous O–O interactions at the same displacement vectors, also at $a = 3.874$ Å.

In summary, within this distance range, first-, second-, and third-neighbor Hamiltonian elements become non-zero, and their magnitudes are significant due to the spatial proximity of the involved WFs.

RANGE 4–9 Å. In the interval between 4 Å and 9 Å, additional non-zero interactions continue to emerge, yet their contribution to the initial goal function remains comparatively small. This behavior can be attributed to the strong localization of the WFs. Although the number of interaction terms increases beyond 4 Å, the magnitude of the corresponding matrix elements h_{ab} decreases significantly due to the rapid decay of the WFs and reduced spatial overlap.

Based on the behavior of the goal function, a cutoff value of $\delta r_{\mathrm{h}} = 6$ Å may appear sufficient. However, a detailed analysis of the resulting band structure reveals residual discrepancies in the valence band near the Γ point [see Fig. 6.5(b)].

RANGE > 9 Å. For interatomic distances beyond 9 Å, the Hamiltonian elements become negligibly small, and the goal function resumes a nearly constant value, indicating the absence of further significant interactions.

Based on this analysis, a cutoff radius $\delta r_{\mathrm{h}} = 9.0$ Å is identified as a reasonable compromise. This value ensures that all relevant short-range interactions are incorporated into the model. In Fig. 6.3, this cutoff is represented by the vertical dotted blue line. It is important to note that increasing δr_{h} also leads to a rapid growth in the number of Hamiltonian terms (see red curve in Fig. 6.3), which in turn increases the computational cost associated to running second-principles calculations.

To study the influence of the cutoff parameter δr_{h} on the electronic band structure and to demonstrate the ability of Hamiltonian matrix elements expressed in a Wannier function basis to accurately reproduce first-principles results as δr_{h} increases, Fig. 6.5 presents the calculated bands for the RAG geometry and in the RED for a range of cutoff values δr_{h}. In Fig. 6.5, the band structure for the RAG within the RED obtained from full first-principles calculations using density functional theory (DFT), as implemented in the SIESTA package, is shown in blue. Superimposed on this, we present the band structure computed from the Hamiltonian matrix elements in the Wannier function basis (represented in red). These bands are derived as described in Sec. 2.1.6, where the summation over lattice vectors $\vec{R}$ is truncated for pairs of Wannier functions whose centers are separated by more than the specified cutoff distance δr_{h}.

From a model-building perspective, the cutoff distance δr_{h} directly determines which Hamiltonian matrix elements $h_{\boldsymbol{ab}}$ are retained, effectively selecting the interacting Wannier function pairs $(\boldsymbol{a}, \boldsymbol{b})$ to be included in the model. In this way, we construct four models characterized by cutoff values $\delta r_{\mathrm{h}} \in \{4$ Å, 6 Å, 8 Å, 9 Å$\}$. In addition, since the system we aim to represent is not distorted from the reference geometry and its electronic configuration remains in the ground state, the exclusive parameter which is going to contribute to the hamiltonian matrix elements in Eq. (4.37) is $\gamma_{\boldsymbol{ab}}^{\mathrm{sh,RAG}}$.

Attending to the band diagram representation in Fig. 6.5, the corresponding cutoff values and the number of retained one-electron interactions in each model are indicated at the top of the respective panels. These values provide a basis to evaluate the size of the model and to estimate the computational cost required to achieve an accurate representation of the electronic structure of the system.

By increasing δr_{h}, the accuracy of the band structure calculation improves systematically. As illustrated in Fig. 6.5, the agreement between the model-derived and first-principles bands becomes progressively better with larger values of the cutoff. A satisfactory compromise is achieved at $\delta r_{\mathrm{h}} = 9$Å. Therefore, this value is selected as the cutoff for the subsequent analysis, as previously anticipated.

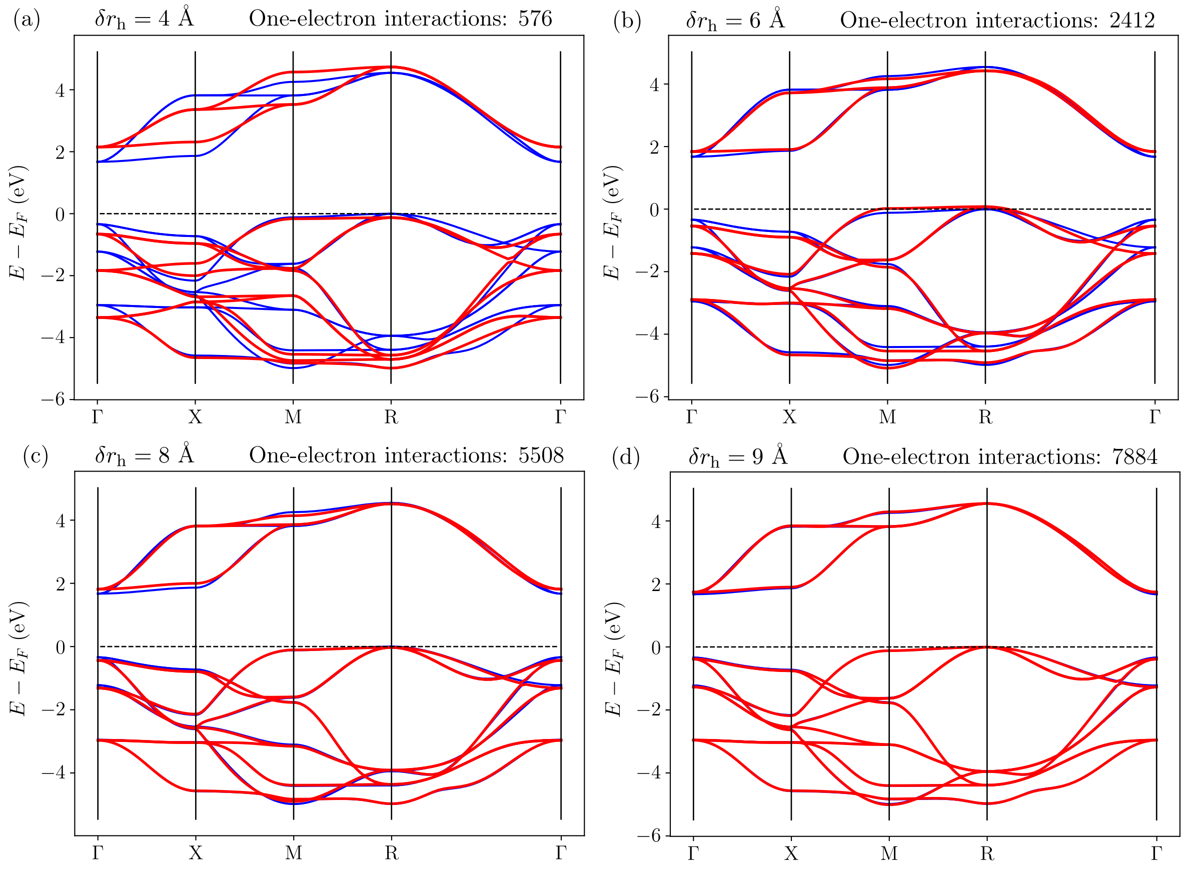


Figure 6.5: Representation of the bands of $SrTiO_3$ in the centrosymmetric paraelectric cubic structure for various values of the distance cutoff δr_{h}. Blue lines represent the bands obtained from first-principles, while the results from the second-principles model are shown in red. On top of each plot we indicate the corresponding number of $h_{\boldsymbol{ab}}$ matrix elements per unit cell that are included.

6.3.1 Electron-lattice parameters: methodology I

After establishing the adequate description of the RAG in the previous steps, we now move to determine the electron-lattice coupling in the form of linear, $\vec{f}_{\boldsymbol{ab},\boldsymbol{\lambda v}}$, and quadratic, $\overleftrightarrow{g}_{\boldsymbol{ab},\boldsymbol{\lambda v}}$ terms, according to the *methodology I*.

The *methodology I* was further evaluated for the $SrTiO_3$ system. In this Section, we present a collection of selected results along with a discussion of the challenges encountered during its development.

As described in Sect. 5.6.2, for this methodology, the training set employed consists of a collection of calculations characterized by random atomic displacements. Each atom is randomly displaced from its position in the RAG within a sphere of radius 0.1 Å.

Given the training set, the analysis of electron-lattice coupling begins with the identification of appropriate cutoff parameters, a crucial aspect controlled by the user. Cutoffs determine which variables are included in the fitting procedure based on two primary criteria: proximity in real space (useful due to the localized nature of the Wannier functions) and the quadratic error generated by the electron-lattice training set associated with each symmetry group—a set of pairs of Wannier functions which are related by symmetry.

Then, following the generation of the training set, the next step involves identifying the symmetry groups that introduce the largest errors. Groups with larger errors significantly influence the overall model accuracy and are therefore prioritized, as the most critical parameters involve Wannier function pairs dependent on the training set state. Specifically, a group G is selected for the generation of electron-lattice variables if it satisfies the condition

$$\Theta_{\mathrm{el}}^{G} > \delta\varepsilon_{\mathrm{el}}. \tag{6.2}$$

where Θ_{el}^{G} represents the goal function per symmetry group. Notably, in this step, the goal function is independent of any electron-lattice or electron-electron model parameters.

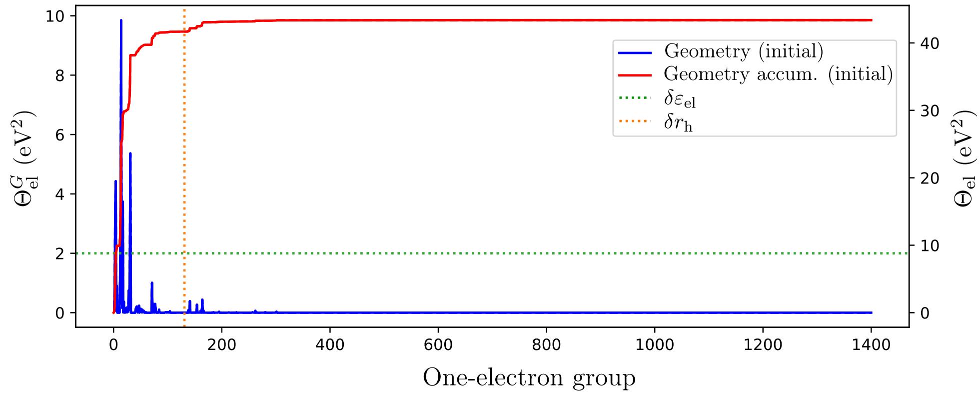


Figure 6.6: Quadratic error generated by the electron-lattice training set is shown per group, Θ_{el}^{G} (left y-axis), as solid vertical blue lines. This error is computed in the absence of electron-lattice parameters. The groups are ordered by distance. The cumulative sum of these errors is represented by the red curve, corresponding to the right y-axis. The cutoff distance $\delta r_{\mathrm{h}} = 9\,\text{Å}$ is indicated by a vertical orange dotted line, while the cutoff energy $\delta\varepsilon_{\mathrm{el}} = 2.0\,\mathrm{eV}^2$ is represented by a horizontal green dotted line.

Figure 6.6 presents the obtained values of Θ_{el}^{G} (blue vertical lines), illustrating the error contribution of different groups. Since all interactions within a given group share the same interaction distance, the groups can be ordered by increasing interaction range, as represented on the x-axis. The first groups are associated with shorter distances while the last ones will be characterized by large interaction distances. Additionally, the red curve in Fig. 6.6 represents the cumulative sum of errors (right y-axis), where its final value corresponds to the total error contribution from all groups, given by Θ_{el}. In the figure, the cutoff distance δr_{h} is depicted as a dotted vertical orange line, while the cutoff energy $\delta\varepsilon_{\mathrm{el}}$ is shown as a dotted horizontal green line. The groups associated with error bars exceeding $\delta\varepsilon_{\mathrm{el}}$ are selected [Eq. (6.2)] for the generation of electron-lattice variables. This visualization provides a clear understanding of the primary groups sources of error and their dependence on interaction distance, guiding the selection of relevant variables for the model.

Thus, the selection of groups for electron-lattice variable generation depends on the value of $\delta\varepsilon_{\mathrm{el}}$. For instance, considering two models with a fixed cutoff distance $\delta r_{\mathrm{h}} = 9.0$ Å. A cutoff energy of $\delta\varepsilon_{\mathrm{el}} = 2.0\ \mathrm{eV}^2$ results in the selection of six groups. As $\delta\varepsilon_{\mathrm{el}}$ decreases, the number of selected groups increases. When the cutoff energy is reduced to $\delta\varepsilon_{\mathrm{el}} = 0.5\ \mathrm{eV}^2$, the number of selected groups for electron-lattice variable generation rises to twelve.

At this stage, we have identified the groups, specifically the pairs of Wannier functions $(\boldsymbol{a},\boldsymbol{b})$, that will be considered in the selection of electron-lattice parameters.

After selecting the pairs of Wannier functions, the next step is to pick the atoms involved in the formation of electron-lattice variables $\vec{f}_{\boldsymbol{ab},\boldsymbol{\lambda v}}$ and $\overset{\leftrightarrow}{g}_{\boldsymbol{ab},\boldsymbol{\lambda v}}$. In this methodology (*methodology I*) both variables depend on two atoms as defined in Eq. (4.24). The list of atomic pairs consists of atoms whose distance measured in the RAG is smaller than the chosen cutoff δr_{el}. For $\mathrm{SrTiO_3}$, the considered cutoff values are 2.5 Å and 4.0 Å. In the cubic paraelectric geometry of $\mathrm{SrTiO_3}$ the lattice parameter is $a = 3.874$ Å. Then, the cutoff $\delta r_{\mathrm{el}} = 2.5$ Å includes first-neighbor atomic pairs, such as Ti-O, while $\delta r_{\mathrm{el}} = 4.0$ Å adds atomic pairs that are one periodic replica away, such as Ti-Ti and O-O. These longer-range interactions are particularly significant in the context of octahedra rotations.

Considering the pairs of Wannier functions selected by $\delta\varepsilon_{\mathrm{el}}$ and the pairs of atoms defined by δr_{el}, the final step involves the selection of electron-lattice variables. Specifically, if a pair of Wannier functions $(\boldsymbol{a},\boldsymbol{b})$ and a pair of atoms $(\boldsymbol{\lambda},\boldsymbol{v})$ satisfy the condition imposed by the cutoff $\delta r_{\mathrm{ellat}}$, then the variables $\vec{f}_{\boldsymbol{ab},\boldsymbol{\lambda v}}$ and $\overset{\leftrightarrow}{g}_{\boldsymbol{ab},\boldsymbol{\lambda v}}$ are consider in the fitting procedure. In this study, the selected values for the electron-lattice parameters are $\delta r_{\mathrm{ellat}} = 4.0$ Å and $\delta r_{\mathrm{ellat}} = 6.0$ Å. For larger cutoff distances δr_{el} and $\delta r_{\mathrm{ellat}}$, the calculations become impractical.

After the selection of electron-lattice variables for a given model, the fitting procedure is initiated. In this step, the most relevant electron-lattice parameters are selected from the generated set, and their corresponding values are computed. This procedure aims to identify the optimal model by fitting the electron-lattice variables to the training set (see Sec. 5.6.2.7). Figure 6.7 illustrates the goal function (the error produced by the training set of the system) as a function of the number of variables included in each step of the minimization in the different studied models.

The models under investigation are characterized by the following cutoffs:

a) $\delta\varepsilon_{\mathrm{el}} = 2.0\ \mathrm{eV}^2$, $\delta r_{\mathrm{el}} = 2.5$ Å, $\delta r_{\mathrm{ellat}} = 4.0$ Å

b) $\delta\varepsilon_{\mathrm{el}} = 0.5\ \mathrm{eV}^2$, $\delta r_{\mathrm{el}} = 2.5$ Å, $\delta r_{\mathrm{ellat}} = 4.0$ Å

c) $\delta\varepsilon_{\mathrm{el}} = 2.0\ \mathrm{eV}^2$, $\delta r_{\mathrm{el}} = 4.0$ Å, $\delta r_{\mathrm{ellat}} = 4.0$ Å

d) $\delta\varepsilon_{\mathrm{el}} = 2.0\ \mathrm{eV}^2$, $\delta r_{\mathrm{el}} = 2.5$ Å, $\delta r_{\mathrm{ellat}} = 6.0$ Å

In Fig. 6.7, the dots represent electron-lattice variables, while the crosses correspond to electron-electron variables, which will be addressed in the following section.

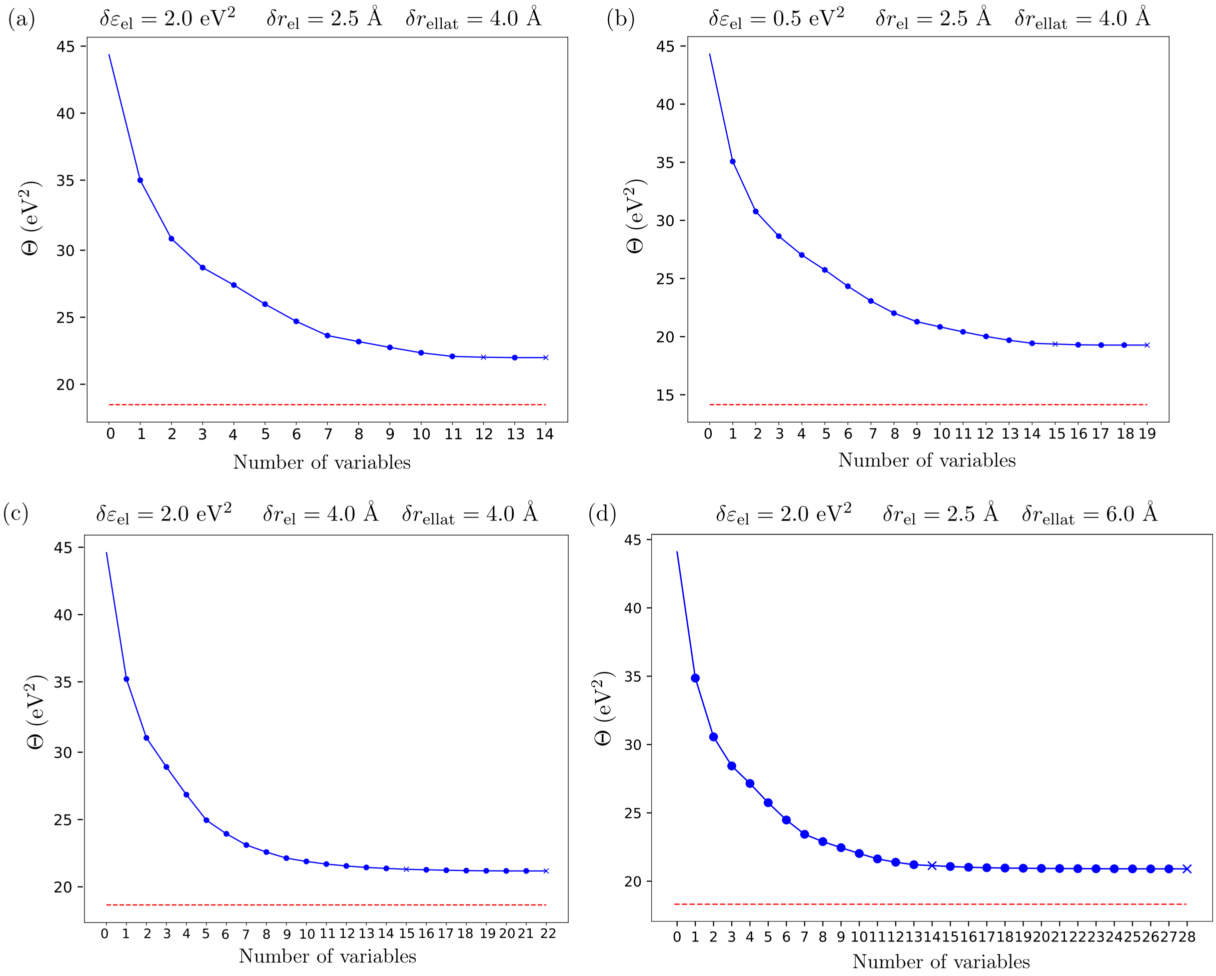


Figure 6.7: Total goal function, Θ, as a function of the number of variables included in the model during the fitting procedure represented by in blue. Blue dots indicate that the corresponding selected variable is an electron-lattice variable. The red dashed line represents the error associated with the optimal model that can be constructed given the chosen cutoffs δr_{h} and $\delta\varepsilon_{\mathrm{el}}$. Each panel corresponds to a different set of electron-lattice cutoff parameters, which are specified at the top of the respective plots leading to different models.

As shown in Fig. 6.7, the total goal function quickly decreases with the inclusion of variables, with the most significant reductions occurring with the incorporation of the first (most relevant) variables. However, as more variables are added, the reduction in error becomes less pronounced, particularly for the last variables selected.

When adjusting the cutoffs to improve the model, it is evident that the number of considered electron-lattice variables increases. For instance, when changing the energy cutoff from $\delta\varepsilon_{\mathrm{el}} = 2.0\,\mathrm{eV}^2$ [Fig. 6.7(a)] to a smaller cutoff, $\delta\varepsilon_{\mathrm{el}} = 0.5\,\mathrm{eV}^2$, which allows for the formation of electron-lattice variables for groups associated to smaller Θ^{G}_{el} error, [Fig. 6.7(b)], the number of variables increases from 14 to 19. When adding atomic pairs at larger distances, the number of selected variables increases further. Comparing

Fig. 6.7(a), characterized by $\delta r_{\text{el}} = 2.5$ Å, with Fig. 6.7(c), characterized by $\delta r_{\text{el}} = 4.0$ Å, the number of selected variables increases to 22. Finally, when the cutoff δr_{ellat} is increased from 4.0 Å to 6.0 Å, the number of selected variables rises to 28.

While adjusting the cutoffs improves the model, it does not lead to a significant reduction in the quadratic error. It is apparent that beyond the 14th variable, the curve flattens. This indicates that the decreasing in the goal function is minimal, and the final selected variables do not exhibit a clear priority over others. Indeed, we observe that with the inclusion of additional variables when "improving" the cutoffs, the curve becomes flat, showing little enhancement in the reproduction of the training set. Therefore, although we may improve the cutoffs and attempt to create a less computationally intensive model, these adjustments do not significantly reduce the error. We can just compare the different plots, all of them reduce the error until 20 eV^2 approximately.

By employing the previously generated models—which integrate the one-electron parameters along with the electron-lattice coupling corrections—it becomes possible to analyze the improvements introduced by the electron-lattice interaction in representing the training set used to fit the model. For the different models, the quadratic error associated with each symmetry group can be compared before and after the inclusion of the electron-lattice coupling terms. Figure 6.8 presents a plot analogous to that shown in Fig. 6.6 for model (a) characterized by $\delta\varepsilon_{\text{el}} = 2.0$ eV^2, $\delta r_{\text{el}} = 2.5$ Å, $\delta r_{\text{ellat}} = 4.0$ Å. In this case, the vertical dashed blue lines indicate the quadratic error per group obtained without incorporating electron-lattice coupling (i.e., the values shown in Fig. 6.6), whereas the solid vertical blue lines represent the corresponding errors after the inclusion of electron-lattice terms, that is, with the final model. The four models yield similar plots in which the changes are negligible. Therefore, in this work, we have presented only one representative example.

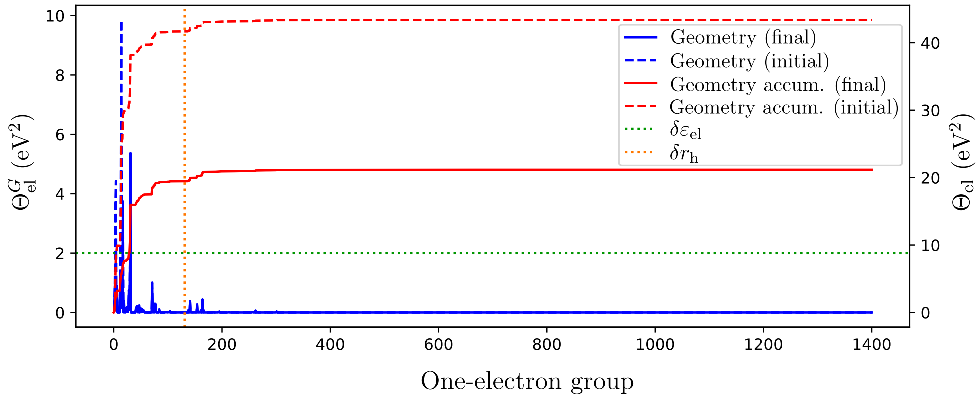


Figure 6.8: Quadratic error generated by the electron-lattice training set is shown per group, Θ_{el}^G (left y-axis), as solid vertical blue lines. This error is computed after incorporating the electron-lattice parameters, which account for geometry distortions. The groups are ordered by distance. The cumulative sum of these errors is represented by the solid red curve, corresponding to the right y-axis. Dashed blue vertical lines represent the initial quadratic error (the ones shown in Fig. 6.6). The cumulative sum of the initial errors is represented by the dashed red curve, corresponding to the right y-axis. The cutoff distance $\delta r_{\text{h}} = 9$ Å is indicated by a vertical orange dotted line, while the cutoff energy $\delta\varepsilon_{\text{el}} = 2.0\,\text{eV}^2$ is represented by a horizontal green dotted line.

At this point, we are going to validate the accuracy of our proposed strategy. In order to check the method, we will discuss its application to different configurations of the SrTiO$_3$ system out of the training set. Moreover, for the purpose of building the best model regarding a balance between accuracy and efficiency, we explore different values for the cutoff variables. Here, the validation of the models is based on the comparison of the band diagram representations obtained from first-principles calculations versus second-principles calculations. The latter take advantage of the models. The first-principles bands data have been obtained from calculations with the SIESTA code based on density-functional theory by local density approximation (LDA). In the representation, they are shown in blue while the band diagrams from second-principles calculations are colored in red. Let us now discuss the predictions for the electronic band structure that our obtained models yield for SrTiO$_3$'s ferroelectric and antiferrodistortive phases. It is worth noting that in the study of electron-lattice parameters, the effect of strain has not yet been incorporated. Therefore, when referring to the different phases, we are actually referring to the atomic displacements that characterize these phases within the cubic cell.

In Fig. 6.9 we compare the full first-principles bands for a SrTiO$_3$ unit cell in a ferroelectric phase with those obtained by second-principles from models corresponding to $\delta\varepsilon_{\text{el}} = 2.0$ eV^2, $\delta r_{\text{el}} = 2.5$ Å and $\delta r_{\text{ellat}} = 6.0$ Å.

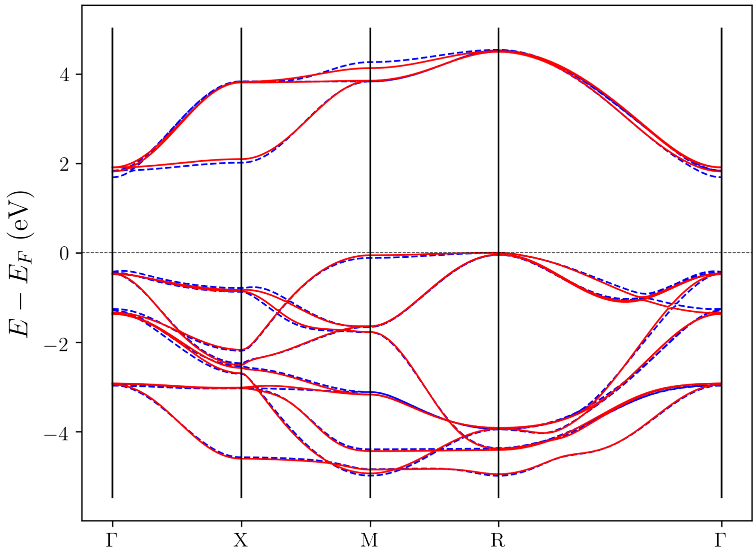


Figure 6.9: Representation of the bands of ferroelectric SrTiO$_3$. Blue lines show the bands obtained from first-principles and the red ones, the results from the second-principles models.

The incipient ferroelectricity observed in SrTiO$_3$ along the z-axis induces a spontaneous polarization and drives the system into a tetragonal phase, accompanied by atomic displacements that break the inversion symmetry. While the x and y directions remain equivalent, the z-axis becomes distinct, lowering the overall symmetry of the crystal. In particular, in Fig. 6.9 we have considered exclusively the atomic displacements, without including the effect of the strain. This structural distortion is reflected in the electronic band structure. In the high-symmetry paraelectric cubic phase, the three lowest conduction bands, primarily characterized by t_{2g} states, are degenerate at the Γ point. However, upon transitioning to the ferroelectric phase, this degeneracy is lifted: two of the bands, corresponding to the x and y directions, remain degenerate, while the third band shifts downward in energy. This results in an energy splitting of $\Delta E_{\mathrm{t_{2g}}} = 0.15$ eV between the non-degenerate z-oriented state and the x-y doublet.

The symmetry breaking and the associated band splitting arise from a strong coupling between the electronic states and lattice distortions—an interaction that is not only central to ferroelectricity but also plays a crucial role in the formation of polarons. In this context, the lowered symmetry and anisotropic band structure enhance the localization tendencies of charge carriers. Indeed, electron-doped $\mathrm{SrTiO_3}$, with its high lattice polarizability and low-energy conduction bands, provides a prototypical setting for examining electron-phonon interactions and polaron formation in perovskites. As an example, the formation of small polarons has been observed experimentally through a substantial enhancement of the electron's effective mass [199]. As a result, reproducing the observed splitting through electron-lattice correction terms constitutes one of the primary objectives of our models. However, when applying *methodology I*, achieving an accurate value of this splitting became a significant challenge. The models constructed, some of them presented above, tended to underestimate the splitting, yielding values $\Delta E_{\mathrm{t_{2g}}}$ between 0.05 and 0.07 eV, as shown in the band structure representation, Fig. 6.9.

In an attempt to reproduce the splitting, several models were generated in which both the training set and the cutoffs were varied. However, due to the inability to accurately reproduce this splitting, the analysis shifted to studying the interactions between Wannier functions in the ferroelectric phase, aiming to investigate the quantitative origin of the splitting. Specifically, the focus was on determining the range of distances that needed to be considered between two Wannier functions generated for the ferroelectric phase in order to accurately reproduce the observed splitting. The ferroelectric phase was wannierized, resulting in Wannier functions that were perturbed from those obtained for the cubic centrosymmetric phase due to changes in hybridization caused by the geometric distortions.

Then, a model was built where the selected RAG was the ferroelectric phase (without electron-lattice parameters). Subsequently, a study was performed to assess the variation of the splitting as a function of the cutoff δr_{h}, as summarized in Tab. 6.1. Since the study focuses on the conduction bands associated with the t_{2g} states of the titanium atoms, the choice of δr_{h} in this table was based on the interactions between a titanium atom and its periodic replicas. According to the table, the first column specifies the number of neighboring cells considered by the model for constructing the bands. The second column represents the hopping (expressed in terms of the lattice vectors of a unit cell) between the titanium atom in the home unit cell and the titanium atoms in a neighbor cell $\vec{R}$, determining the selected cutoff δr_{h}. The third column provides the corresponding value of the cutoff δr_{h} needed to include the neighboring cell, while the fourth column shows the resulting splitting of the conduction band.

From the results, it is evident that accurately reproducing the splitting requires incorporating the effects of Wannier functions up to the fourth nearest neighbors.

If we return to a model where the selected reference atomic geometry (RAG) is the cubic centrosymmetric phase, the changes in the Wannier functions of the ferroelectric phase, compared to those of the RAG, should be captured by the electron-lattice coupling correction. To accurately capture the interaction between two Wannier functions at the fifth nearest neighbors the cutoff δr_{h} must be at least 9 Å. Additionally, this requires considering electron-lattice parameters associated with pairs of Wannier functions $(\boldsymbol{a}, \boldsymbol{b})$ and pairs of atoms $(\boldsymbol{\lambda}, \boldsymbol{v})$ that are 9 Å apart. As a result, $\delta r_{\mathrm{ellat}}$ should be set to a minimum of 9 Å, making the use of *methodology I* computationally prohibitive.

The introduction of *methodology II* is motivated by these challenges, particularly the difficulty in reproducing the ferroelectric phase at a manageable computational cost. Moreover, *methodology II* aims to address the need for a methodology in which the fitting process of electron-lattice variables is fully determined. To this end, the new methodology removes the cutoff $\delta r_{\mathrm{ellat}}$, instead introducing energy-based cutoffs that filter electron-lattice parameters according to their energy values, rather than the distance between Wannier functions and the involved atoms.

Ti-Ti neighbor cell	R_x	R_y	R_z	δr_{h} (Å)	$\Delta E_{t_{2g}}$ (eV)
on-site	0	0	0	1.0	0.016
1st-neigh	0	1	0	4.0	0.073
2nd-neigh	1	1	0	5.5	0.104
3rd-neigh	1	-1	-1	7.0	0.101
4th-neigh	2	0	0	8.0	0.120
5th-neigh	0	2	1	9.0	**0.146**
6th-neigh	1	-2	1	9.5	0.144
7th-neigh	2	2	0	11.0	0.149
8th-neigh	2	-2	1	12.0	0.149
Complete	2	2	2	50.0	0.150

Table 6.1: Table displaying hopping parameters, cutoff distances, and corresponding splitting values.

Once the atomic displacements on the ferroelectric phase have been studied with *methodology I*, we consider the atomic distortions in the antiferrodistortive geometry (see Fig. 6.10). As previously mentioned, the strain is not included in the phase.

Electronic band structure of a SrTiO$_3$ antiferrodistortive geometry is obtained from first-principles and second-principles calculations. This phase is simulated in a supercell $2 \times 2 \times 2$ since the antiphase rotation of the octahedra can't be reproduced by translations of the unit cell with 5 atoms. For this geometry, the electronic band structures corresponding to the previously defined models were calculated using the second-principles framework, with the results shown as red curves in Figs. 6.11(a)–6.11(d). For comparison, the corresponding first-principles band structure is shown in blue.

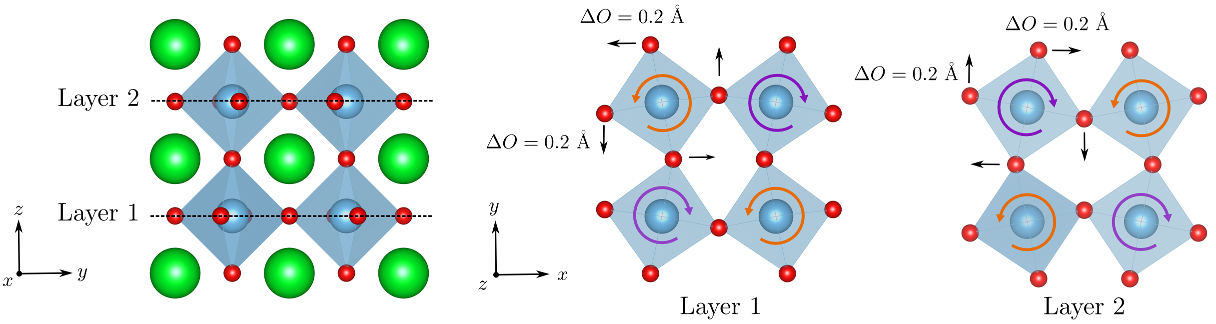


Figure 6.10: Representation of the rotations of oxygen octahedra in the $SrTiO_3$ antiferrodistortive phase, characterized by octahedra rotations about the z-axis. The octahedra rotate in anti-phase along the three Cartesian directions. The figure shows the rotational pattern of the four octahedra in each titanium atomic layer, denoted as "Layer 1" and "Layer 2". The octahedra rotations alternate between clockwise (indicated by purple arrows) and anti-clockwise (indicated by orange arrows). These rotations result from displacements of oxygen atoms: by $\Delta O = 0.2\,\text{Å}$ along the x-direction within the Y-plane, and by $\Delta O = 0.2\,\text{Å}$ along the y-direction within the X-plane, with the senses represented in the scheme.

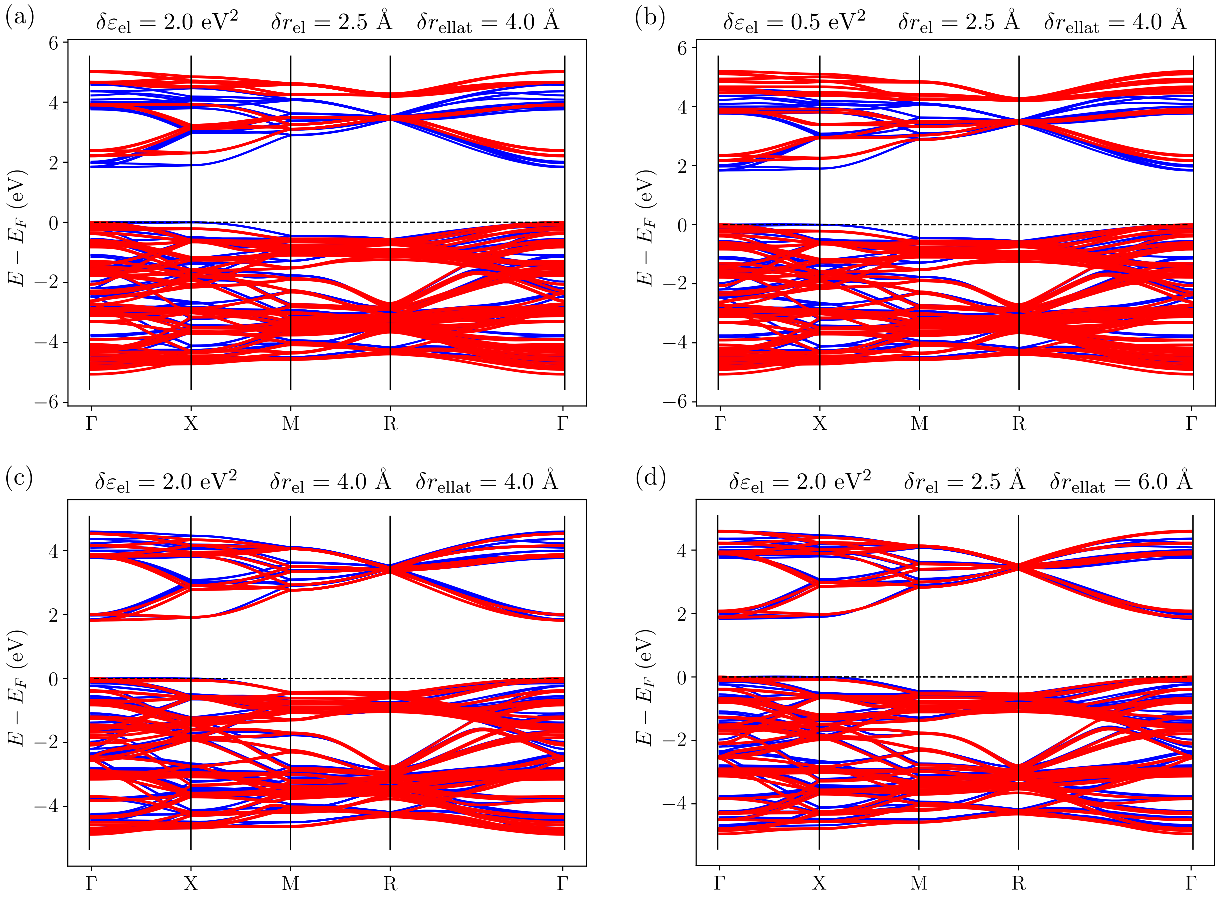


Figure 6.11: Electronic band structure of $SrTiO_3$ system characterized by a geometry with the atomic displacements in the antiferrodistortive $SrTiO_3$ phase. Blue lines represent the bands obtained from first-principles, while the results from the second-principles models, characterized by a collection of cutoffs, are shown in red. At the top of each plot, we indicate the corresponding cutoff values.

Analyzing the different electronic band structures, we observe neither the model associated to Fig. 6.11(a) nor the model improving $\delta\varepsilon_{\mathrm{el}}$, Fig. 6.11(b), are able to reproduce the band diagram of this geometry. The main and obvious discrepancies with the first-principles calculations are the band gap and the energies obtained for the conduction band in the *R* symmetry point. On the one hand, the band gap from second-principles is larger that the one obtained from first-principles. This fact shifts the conduction band with respect the blue bands. On the other hand, from first-principles calculations the energies at *R* point are degenerate for the conduction band. Looking at second-principles calculations, this degeneration is splitted into two energies. Now, attending to models where the distance cutoffs δr_{el} and $\delta r_{\mathrm{ellat}}$ have been increased leading to Fig. 6.11(c) and Fig. 6.11(d), respectively, we observe a clear improvement of the matching between red and blue bands. In such cases, the band gap and the degeneration of the *R* point in the conduction band are perfectly reproduced. The large amount of bands in the valence band makes it difficult to compare the results for the two calculations. By analyzing the conduction band, we can see how the bands from the different calculations fit quite well. As consequence, the cutoffs $\delta r_{\mathrm{ellat}}$ and δr_{el} play a decisive role in the construction of the SrTiO$_3$ model, since when these cutoffs are increased, the model incorporates variables associated with two oxygen atoms that are periodic replicas, representing a critical interaction in the rotational dynamics of the octahedra.

6.3.2 Electron-lattice parameters: methodology II

We now move to compute the electron-lattice coupling applying the *methodology II*. Firstly, we proceed to determine the electron-lattice coupling terms—specifically, the linear couplings $\vec{f}_{\boldsymbol{ab},\boldsymbol{\lambda}}$ and the quadratic couplings $\overleftrightarrow{g}_{\boldsymbol{ab},\boldsymbol{\lambda\nu}}$. We begin by defining the electron-lattice interaction cutoff distance, δr_{el}, which sets the maximum separation (at the RAG) between atoms participating in the quadratic electron-lattice coupling. It is important to emphasize that increasing δr_{el} significantly impacts on the computational cost, both in terms of the number of DFT calculations required to generate the training set and the resulting size of the second-principles model, which in turn affects its efficiency in simulations.

To evaluate the impact of different δr_{el} values, we construct a test set composed of 10 randomly distorted geometries, as described in Sec. 5.6.4, characterized by an atomic displacement parameter d. The resulting second-principles energies are compared against their DFT counterparts, and the average error per calculation in testing set is plotted in Fig. 6.12(a) as a function of d. The black curve corresponds to a model without any electron-lattice correction, while the olive green curve includes only linear electron-lattice terms. The remaining curves (purple, blue, and orange) incorporate quadratic corrections for increasing values of δr_{el}: 2.0 Å which includes only nearest-neighbor Ti–O interactions; 4.0 Å capturing interactions between atoms spanning a full five-atom unit cell; and 5.6 Å which incorporates longer-range contributions beyond a single unit cell.

To clearly identify the errors introduced by the models with the inclusion of electron-lattice coupling, Fig. 6.12(b) presents the same set of curves as in the previous plot (a), excluding the errors associated with the model in the absence of electron-lattice interaction. As shown in Fig. 6.12(b), the model with

δr_{el} = 2.0 Å exhibits a more rapid accumulation of error compared to the models with δr_{el} = 4.0 Å and 5.6 Å which yield nearly indistinguishable results. This behavior underscores the importance of including quadratic oxygen–oxygen interactions within a single TiO_6 octahedron in order to accurately capture the modifications to the Hamiltonian induced by strong electron-lattice coupling in $SrTiO_3$. To reduce the number of electron-lattice coupling terms and thereby simplify the model, we introduce cutoff thresholds δf and δg for the inclusion of linear and quadratic coupling constants, respectively. The effect of these thresholds on model accuracy is presented in Fig. 6.13. For large thresholds ($\delta f = 0.5$ eV/Å, $\delta g = 0.5$ eV/Å^2), the error increases rapidly with the displacement d. Reducing both δf and δg by a factor of five (δf = 0.1 eV/Å, δg = 0.1 eV/Å^2) leads to a fourfold reduction in error for $d = 0.17$ Å. These errors can be further reduced until we reach thresholds that are one order of magnitude smaller, $\delta f = 0.01$ eV/Å, $\delta g = 0.01$ eV/Å^2, where we seem to find the limit for the quadratic coupling approximation employed here.

As before, to clearly identify the errors introduced by the models with the inclusion of electron-lattice coupling, Fig. 6.13(b) presents the same set of curves as in the previous plot (a), removing the curve associated with the model in the absence of electron-lattice interaction.

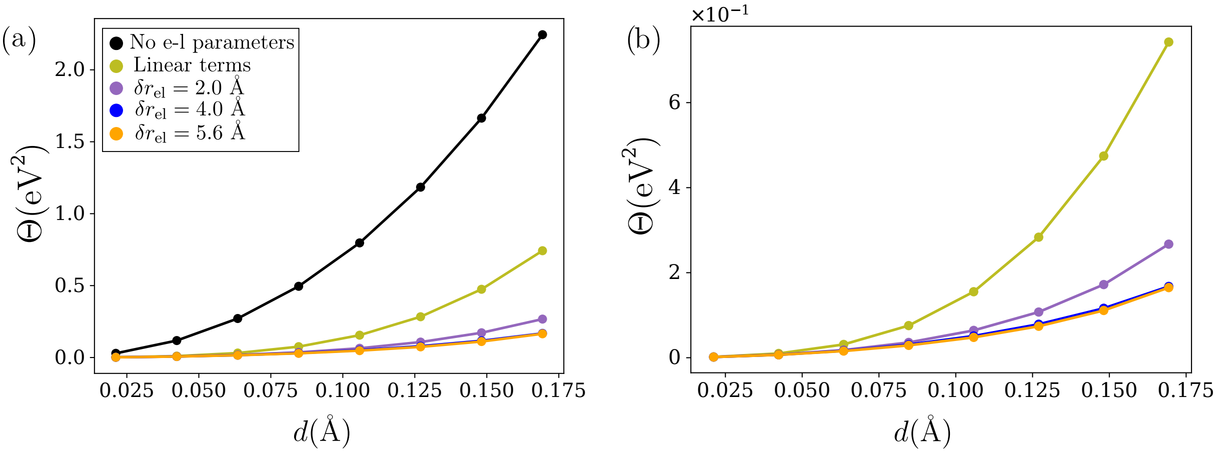


Figure 6.12: (a) Goal function Θ for $SrTiO_3$ measuring the average error per calculation in test set consisting of ten configurations with increasing random displacements d from the RAG. The cutoff values for the electron-lattice force constants are fixed at $\delta f = 0.1$ eV/Å and $\delta g = 0.1$ eV/Å^2. The black curve corresponds to a model without electron-lattice coupling; the olive green curve includes only linear terms. The purple, blue and orange curves include both linear and quadratic terms, with electron-lattice cutoffs $\delta r_{\mathrm{el}} = 2.0$ Å, $\delta r_{\mathrm{el}} = 4.0$ Å, and $\delta r_{\mathrm{el}} = 5.6$ Å, respectively. The blue and orange curves are indistinguishable. In panel (b), the results corresponding to the model without the electron-lattice coupling correction have been excluded to enable a focused view of the remaining data.

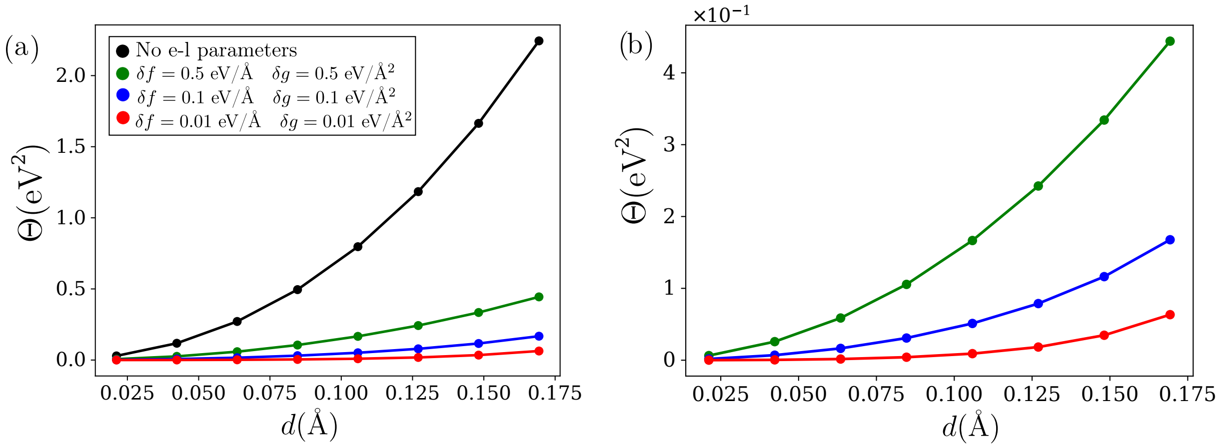


Figure 6.13: (a) Goal function Θ for $SrTiO_3$ measuring the average error per calculation in test set consisting of ten configurations with increasing random displacements d from the RAG. The curves measure the effect of varying the cutoff thresholds δf and δg on Θ, for fixed $\delta r_{\mathrm{el}} = 4.0$ Å. The green, blue, and red curves correspond to (δf, δg) values of (0.5 eV/Å, 0.5 eV/Å^2), (0.1 eV/Å, 0.1 eV/Å^2), and (0.01 eV/Å, 0.01 eV/Å^2), respectively. In panel (b), the results corresponding to the model without the electron-lattice coupling correction have been excluded to enable a focused view of the remaining data.

To assess how these errors affect the electronic structure, Fig. 6.14 presents band structures for $SrTiO_3$ including *error bars*. These error bars represent the statistical deviation of the second-principles model predictions from the DFT results for a test set with $d = 0.17$ Å. Similar trends are observed for smaller displacements, as presented in Fig. 6.15 for $d = 0.1$Å, although in those cases the errors are much smaller than the one presented in Fig. 6.14. To examine the influence of the electron-lattice cutoff δr_{el}, we compare the results of the test set for different cutoff values. Figure 6.14(a) shows the band structure in the absence of any electron-lattice coupling, while Figs. 6.14(b)–6.14(d) present results for $\delta r_{\mathrm{el}} = 2.0$, 4.0, and 5.6 Å respectively. A key result of this analysis is the essential role of electron-lattice coupling parameters in accurately reproducing the electronic band structure across different atomic configurations. Both the linear electron-lattice contributions and the quadratic terms associated with first-neighbor Ti–O interactions are found to be critical. This is demonstrated by the substantial reduction in the error bars of the electronic structure when progressing from a model without any electron-lattice coupling [Fig. 6.14(a)] to those incorporating linear and quadratic terms for increasing values of δr_{el} [Figs. 6.14(b)–6.14(d)]. Although the inclusion of linear terms and first-neighbor Ti–O interactions significantly improves the agreement with the reference calculations, the remaining discrepancies in Fig. 6.14(b) indicate that this level of approximation is insufficient for full accuracy.

A more detailed comparison between Figs. 6.14(b) and 6.14(c) underscores the necessity of incorporating longer-range electron-lattice interactions. In particular, extending the interaction range to include atoms across an entire five-atom unit cell yields a marked reduction in error bars, highlighting the importance of non-local effects. Notably, interactions between oxygen atoms and their periodic images within a single unit cell are especially relevant, as they are closely tied to octahedra rotation modes. For instance, when analyzing the geometry of the antiferrodistortive phase represented in Fig. 6.10, including octahedra rotations characterized by in-plane oxygens displacements $\Delta O = 0.2\,\text{Å}$, the choice of a cutoff parameter δr_{el} smaller than the cubic lattice parameter becomes a critical consideration, as evidenced in the band structures shown in Fig. 6.16. The model defined by $\delta r_{\text{el}} = 2.0\,\text{Å}$ produces band structures with significant deviations along the $\Gamma - X$ path, which are rectified by incorporating oxygen-oxygen interactions. The remaining discrepancies will be examined in further analysis.

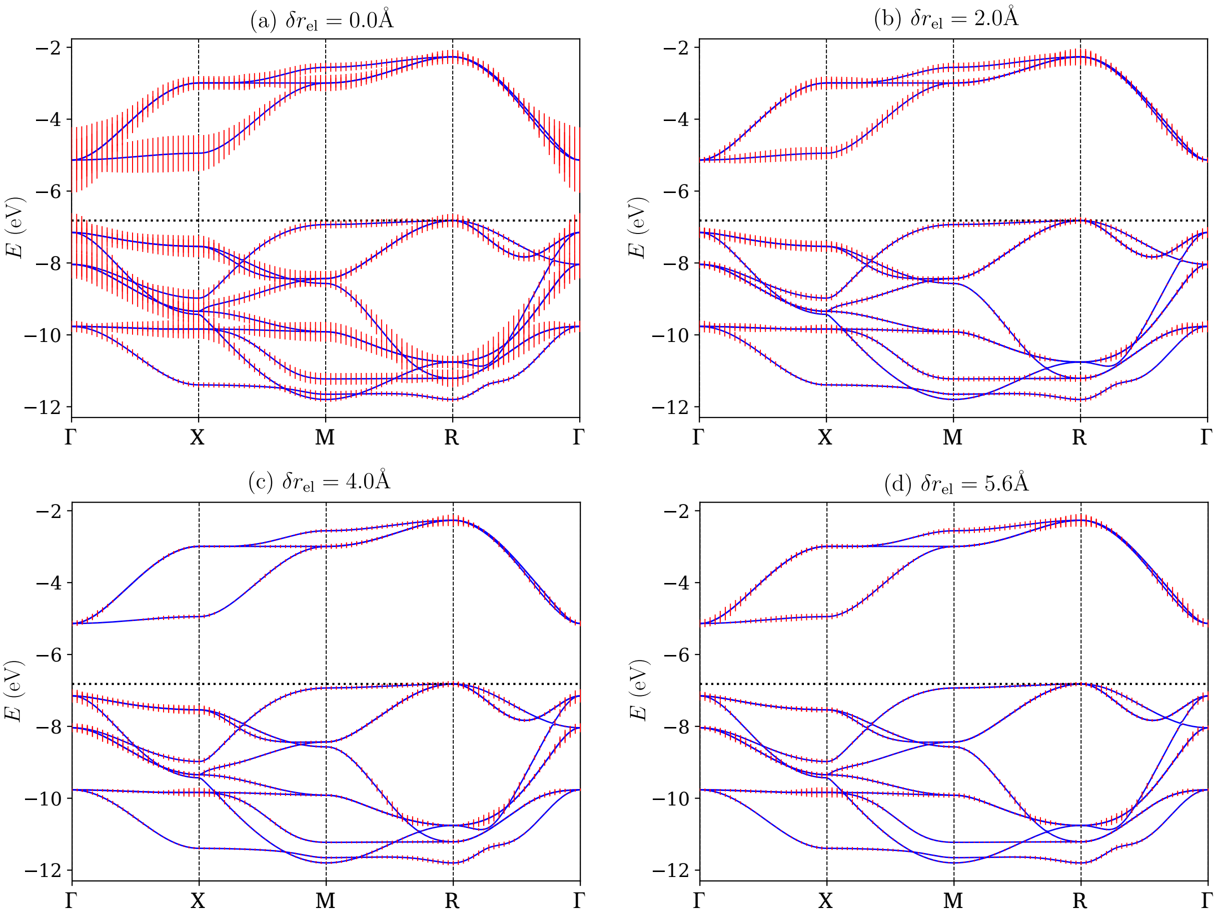


Figure 6.14: Electronic band structures of SrTiO_3 with error bars (red) indicating the statistical deviation of second-principles simulations from DFT results over a test set of ten geometries characterized by $d = 0.17$ Å. Panel (a) corresponds to a model without electron-lattice coupling. Panels (b)–(d) include both linear and quadratic electron-lattice interactions with increasing interaction range characterized by $\delta r_{\text{el}} = 2.0$, 4.0, and 5.6 Å, respectively. All models employ threshold parameters $\delta f = 0.1$ eV/Å and $\delta g = 0.1$ eV/Å^2.

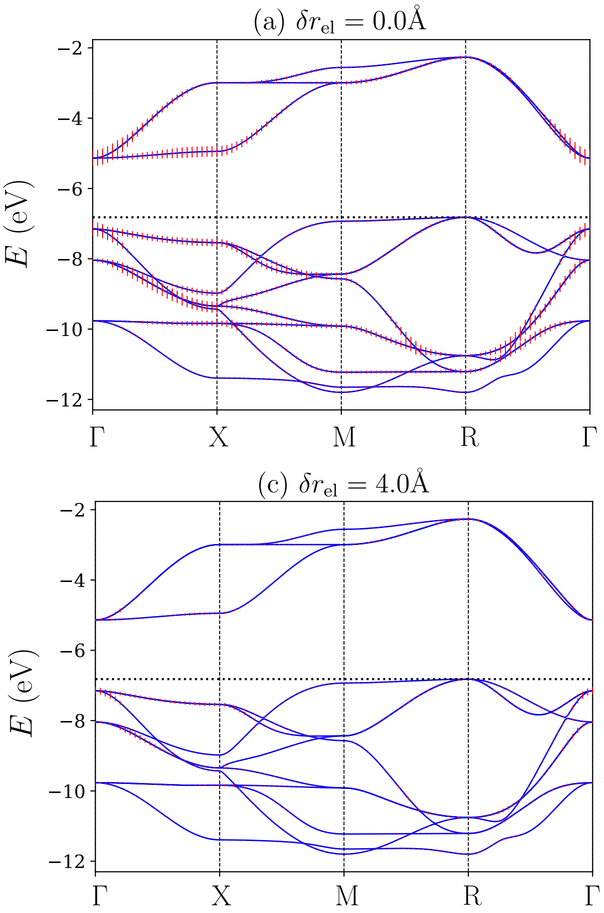


Figure 6.15: Electronic band structures of $SrTiO_3$ with error bars (red) indicating the statistical deviation of second-principles simulations from DFT results over a test set of ten geometries characterized by $d = 0.1$ Å. Panel (a) corresponds to a model without electron-lattice coupling. Panels (b)–(d) include both linear and quadratic electron-lattice interactions with increasing interaction range characterized by $\delta r_{\mathrm{el}} = 2.0$, 4.0, and 5.6 Å, respectively. All models employ threshold parameters $\delta f = 0.1$ eV/Å and $\delta g = 0.1$ eV/Å^2.

At this stage, by adopting larger electron-lattice cutoffs, we observe that quadratic interatomic interactions beyond $\delta r_{\mathrm{el}} = 4$ Å have a negligible impact on the electronic structure. This is seen in Fig. 6.14(d) and in the overlap of orange and blue curves in Fig. 6.12. Therefore, we find that $\delta r_{\mathrm{el}} = 4.0$ Å strikes a balanced trade-off between computational efficiency and accuracy.

Regarding the accuracy of the model, it is noteworthy that the errors in the valence and conduction bands remain small, even when atomic displacements reach the relatively large magnitude of 0.17 Å, as represented in Fig. 6.14. These errors are significantly smaller than the typical separation between band lines, further validating the robustness of the model.

Once the electron-lattice distance cutoff has been determined and analyzed, it is possible to ensure the convergence of the energy cutoffs. Specifically, for a fixed $\delta r_{\mathrm{el}} = 4$ Å, we can reduce the error present in Fig. 6.15(b) to become almost negligible by setting $\delta f = 0.01$ eV/Å and $\delta g = 0.01$ eV/Å^2. These results demonstrate that systematic improvements in model accuracy can be achieved by tightening the cutoffs on the electron-lattice coupling terms.

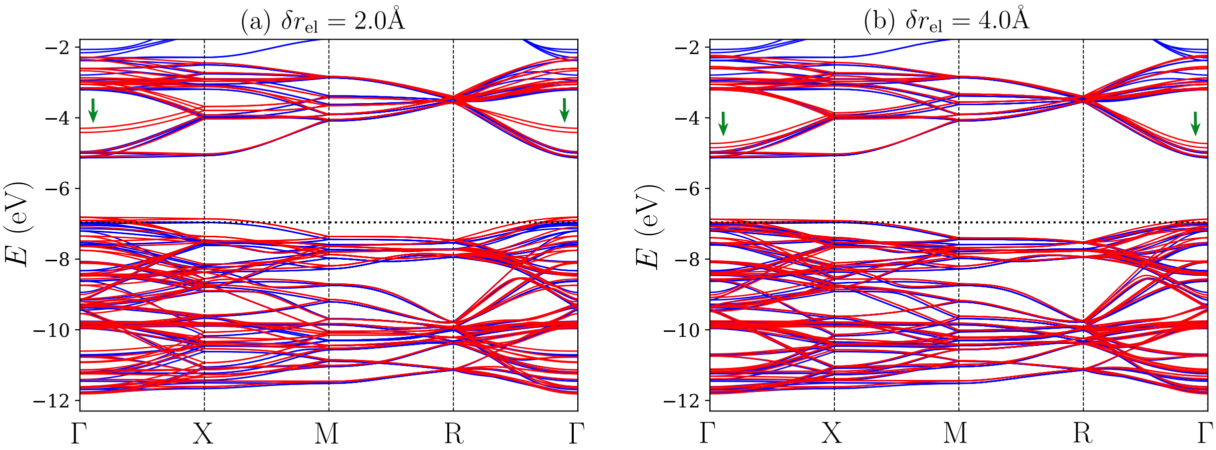


Figure 6.16: Electronic band structure of SrTiO$_3$ a phase characterized by in-plane octahedra rotations, with lattice parameters fixed to their cubic values. First-principles (DFT) results are shown in blue, while second-principles results appear in red. Panels (a) and (b) correspond to models constructed with electron-lattice interaction cutoff distances $\delta r_{\mathrm{el}} = 2.0$ and 4.0 Å, respectively. Green arrows highlight band features whose energies exhibit strong sensitivity to δr_{el}, underscoring the importance of including extended quadratic interactions.

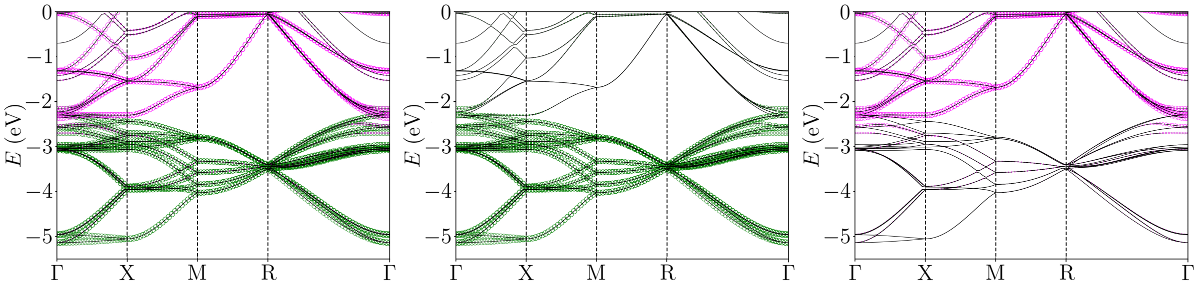
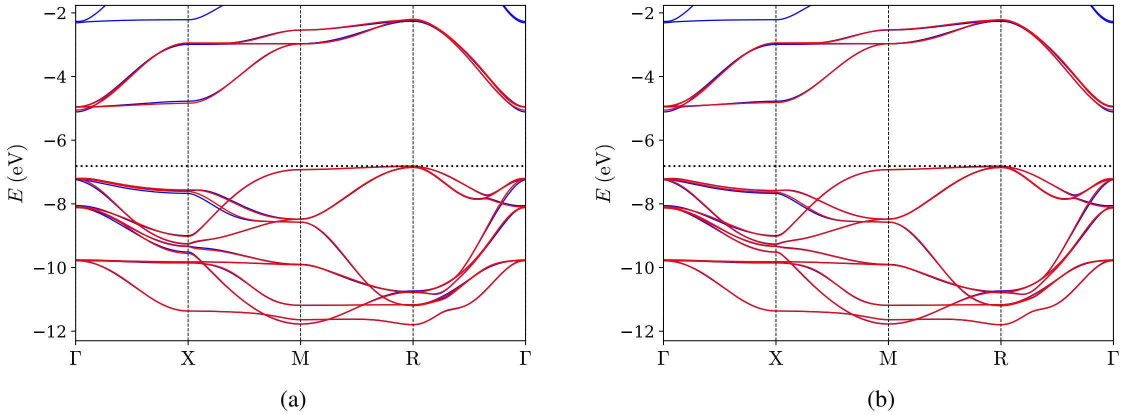


Figure 6.17: (a) Fat band representation of the conduction band for the cubic lattice with the associated atomic displacements of the antiferrodistortive phase of the SrTiO$_3$ system, computed using a $2 \times 2 \times 2$ supercell. The fat bands representation for the bottom of the conduction band is provided, highlighting 24 bands with t_{2g} character (corresponding to three d-orbitals per titanium atom across eight titanium atoms) and 16 bands with e_g character (arising from two d-orbitals per titanium atom across eight titanium atoms), shown in magenta. Panels (b) and (c) show the isolated contributions from the t_{2g} and e_g orbitals, respectively.

Analysis of the error-bar plots reveals that the largest deviations occur near the Γ point for both valence and conduction bands, and around the R point for the conduction band. This behavior can be traced to the narrow energy range over which the t_{2g} and e_g bands overlap in the RAG geometry. As shown in Fig. 6.1, the t_{2g} bands at R overlap energetically with the e_g bands at Γ. Although these manifolds exhibit weak entanglement in the high-symmetry RAG configuration, the octahedra rotations characteristic of the antiferrodistortive phase introduce mixing between the σ-like e_g and π-like t_{2g} states (see Fig. 6.17).

This enhanced entanglement leads to the increased errors observed near the R point in Fig. 6.14. Since the discrepancies appear at relatively high energies—approximately 5 eV above the valence band maximum—their impact on the prediction of optical properties below the vacuum ultraviolet range, or on phenomena such as polaronic behavior, is expected to be minimal. We have explored the possibility of mitigating this error by explicitly including the e_g bands in the model. This approach indeed improves the agreement between second-principles and DFT results in the energy range where the e_g and t_{2g} manifolds overlap. However, due to the strong entanglement of the e_g states with higher-energy bands that are not incorporated into the model, this extension leads to large and unphysical errors at elevated energies. Therefore, we conclude that increasing the number of Wannier functions to include the e_g bands does not offer a viable path toward improving the accuracy of the model.

Finally, the results for the cubic lattice with the displacements characteristic of the ferroelectric phase SrTiO$_3$ are shown in Fig. 6.18. Here, we present the bands computed using two different models, both with $\delta r_{\text{el}} = 4$Å. In Fig. 6.18(a), the model parameters are $\delta f = 0.1\text{eV}/\text{Å}$ and $\delta g = 0.1\text{eV}/\text{Å}^2$, whereas in Fig. 6.18(b), they are $\delta f = 0.01\text{eV}/\text{Å}$ and $\delta g = 0.01\text{eV}/\text{Å}^2$.



(a) (b)

Figure 6.18: Band structure of SrTiO$_3$ for considering off-center displacements (associated to the ferroelectric phase) while keeping the lattice parameters to their cubic values. DFT and second-principles bands are shown in blue and red, respectively. In (a) the energy cutoffs are $\delta f = 0.1\text{eV}/\text{Å}$ and $\delta g = 0.1\text{eV}/\text{Å}^2$ while in (b) they are $\delta f = 0.01\text{eV}/\text{Å}$ and $\delta g = 0.01\text{eV}/\text{Å}^2$ controlling the magnitude of electron-lattice interactions.

With the *methodology II*, where the electron-lattice parameters are computed instead of fitted, we are able to accurately reproduce the splitting in the Γ point for the t_{2g} band states because of the break of the symmetry along the z-direction.

The reproduction of the splitting with the new methodology is a consequence of two facts. Firstly, the new methodology is able to consider the effect of neighbors further than fifth neighbors through the linear constants. Secondly, the model is completely deterministic. Since the variables are computed and not fitted, we do not have to solve a system which is undeterminated. Thus, the electron-lattice variables included in the model for the same step of the finite differences and cutoffs are going to be always the same.

6.3.3 Electron-electron parameters: methodology I

We now proceed to determine the electron-electron interaction parameters, $U_{\boldsymbol{ab},\boldsymbol{a'b'}}$ and $I_{\boldsymbol{ab},\boldsymbol{a'b'}}$, which quantify the variation of the Hamiltonian matrix elements $h_{\boldsymbol{ab}}$ under changes in the total electronic density (sum over both spin channels) and the spin polarization (difference between spin channels), respectively. The standardized training set introduced in *methodology I* Sec. 5.6.2 incorporates both electron and hole doping, as well as magnetized configurations. This allows us to probe how the occupancy of both titanium and oxygen orbitals affects the electronic structure and assess the material's tendency toward spin polarization.

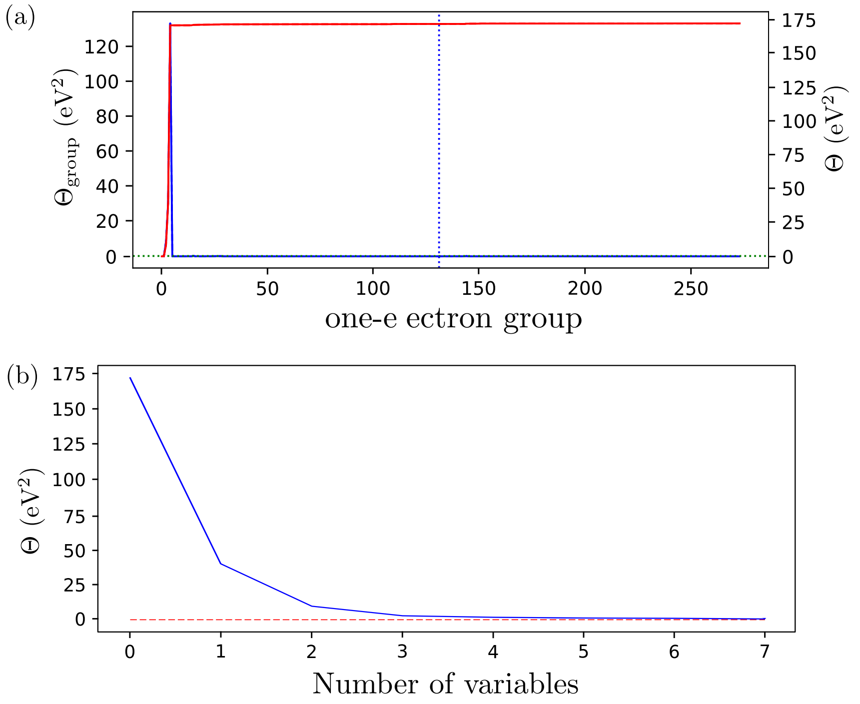


Figure 6.19: (a) Decomposition of the goal function for the electron-electron training set into contributions from each group of symmetry-equivalent $h_{\boldsymbol{ab}}$ matrix elements (blue line), along with the cumulative goal function Θ (red line). The dashed vertical line represents δr_{h}. (b) Evolution of Θ as electron-electron interaction parameters are progressively introduced into the second-principles model.

To apply the fitting protocol described in Sec. 5.6.2.4, we introduce a cutoff parameter, $\delta\Theta$, which defines the set of WF index pairs $(\boldsymbol{a},\boldsymbol{b})$ to be included in the fitting of $U_{\boldsymbol{ab},\boldsymbol{a'b'}}$ and $I_{\boldsymbol{ab},\boldsymbol{a'b'}}$. Figure 6.19(a) displays the contribution of each symmetry-related group of interactions to the electron-electron goal function. The red line represents the total error contribution as a function of group index, ordered by increasing WF separation. We observe a steep rise in the goal function for the initial groups, corresponding to interactions between closely located Wannier functions. This trend quickly saturates, indicating that the dominant electron-electron interactions are intra-atomic, while inter-atomic terms contribute significantly less. Notably, the strongest contributions stem from interactions involving Ti-centered $3d$-like

WFs, consistent with the localized nature of these orbitals. Based on this analysis, we set the cutoff $\delta\Theta = 0.2$ eV^2, as indicated by the green horizontal line in Fig. 6.19(a). This value captures the most significant variations in the Hamiltonian elements, specifically the diagonal terms $h_{\boldsymbol{aa}}$ involving Ti($3d$), O($2p_\sigma$), and O($2p_\pi$) orbitals, as well as the intra-atomic off-diagonal interactions, $h_{\boldsymbol{ab}}$, between O($2p_\sigma$) and O($2p_\pi$).

Following the inclusion of U and I terms in the model, we observe a substantial reduction in the goal function quantifying the difference between DFT and second-principles results, as shown in Fig. 6.19(b). The most significant reduction—from approximately 170 eV^2 to 40 eV^2—is obtained by including the diagonal terms $U_{\boldsymbol{aa},\boldsymbol{aa}}$ and $I_{\boldsymbol{aa},\boldsymbol{aa}}$, where $\boldsymbol{a}$ corresponds to Ti (t_{2g}) WFs. The second most impactful reduction, amounting to approximately 23 eV^2, results from incorporating similar diagonal terms for O($2p_\pi$) orbitals. A third significant reduction is achieved by introducing intra-atomic off-diagonal terms between O($2p_\pi$) and O($2p_\sigma$) WFs. All remaining variables contribute less than 0.5 eV^2 to the goal function and can be safely neglected in the final model.

In the preceding sections, we have demonstrated that a comprehensive electronic model for SrTiO$_3$, a prototypical transition-metal perovskite, can be constructed in a quasi-automated manner. The methodology presented here requires the user to select only a small set of threshold parameters, enabling the generation of accurate second-principles models with relatively modest computational resources.

The training set typically consists of a few hundred single-point calculations (on the order of 100–300) performed on a medium-sized supercell; in the present case, the supercell contains 40 atoms. When combined with a second-principles lattice model constructed following Refs. [36; 51], the result is a fully parameterized model capable of describing both structural and electronic properties.

Importantly, the current electronic model can be coupled with the recently developed real-time time-dependent second-principles DFT (SPDFT) method [173] to compute optical spectra. These predictions show excellent agreement with direct DFT calculations performed using the same exchange-correlation functional employed in the training.

A more challenging application is the description of polaron formation and dynamics in SrTiO$_3$. In this regard, we attempted to induce polaron formation by doping the system with electrons. However, these efforts did not result in charge localization. Benchmark DFT calculations on large supercells indicate that this failure is likely due to the lack of self-interaction correction in the LDA functional, which inhibits polaron formation.

To overcome this limitation, future work will focus on constructing second-principles models based on hybrid functionals. In particular, we plan to exploit the recent implementation of hybrid functionals in SIESTA [174] to develop SrTiO$_3$ models trained using the HSE06 functional [200], which offers a more accurate treatment of electron localization and should enable a realistic description of polaronic effects.

6.4 LiF

In this second example, we consider lithium fluoride (LiF), a prototypical strongly ionic compound whose optical spectrum is characterized by highly localized excitonic states [183; 184; 185]. The presence of a pronounced Stokes shift [184] further indicates significant electron-phonon coupling in the excited state. This unique combination of features makes LiF an ideal candidate for investigation within the second-principles framework, particularly using the recently developed time-dependent propagation formalism [173]. To enable such simulations, however, it is first necessary to construct an accurate second-principles model capable of describing the band structure, electron-lattice coupling, and electron-electron interactions across the valence and conduction manifolds. In this case, the reference atomic geometry (RAG) corresponds to the face-centered cubic (FCC) structure with space group Fm$\bar{3}$m, characteristic of the rock-salt lattice. A conventional unit cell comprising eight atoms and a lattice constant of 4.026 Å is employed for the simulations.

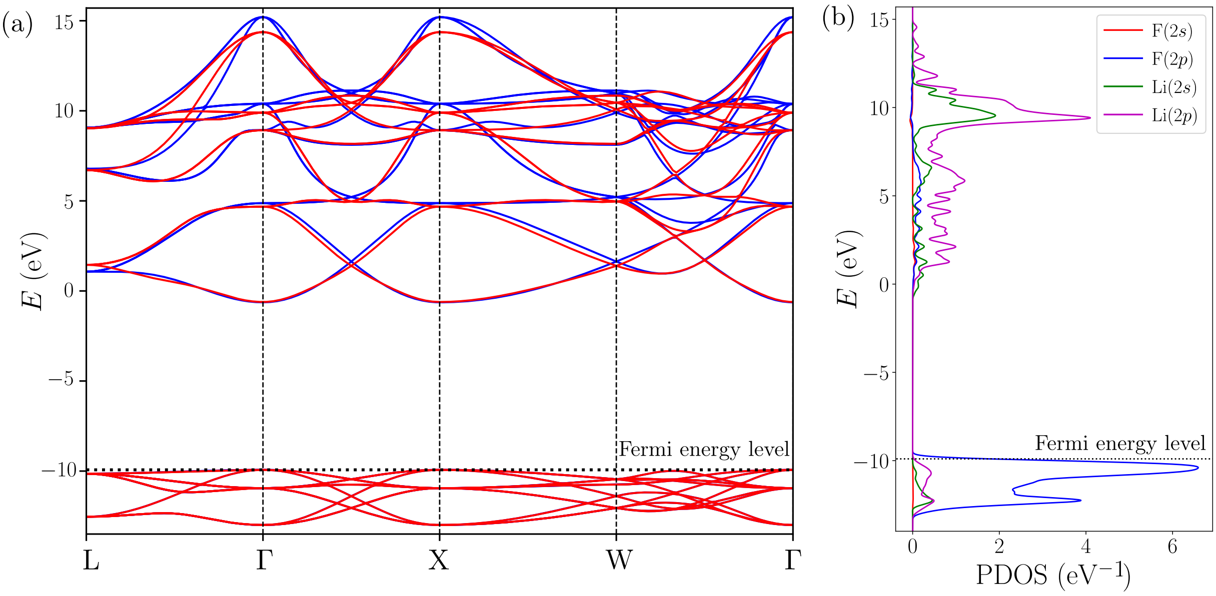


Figure 6.20: (a) Electronic band structure of LiF at the cubic RAG geometry, as obtained from DFT (blue) and second-principles simulations (red). (b) Corresponding projected density of states (PDOS), indicating that the valence bands are predominantly composed of F(2p) orbitals, while the conduction bands exhibit mainly Li(2s) and Li(2p) character. In SIESTA, both the valence and conduction manifolds are disentangled from other states and used to construct the second-principles model.

Figure 6.20 presents the electronic band structure of LiF, computed for the conventional unit cell used to train the second-principles model, using the generalized gradient approximation (GGA) as implemented in SIESTA. The valence bands, primarily of F(2p) character, are relatively narrow, while the conduction bands are significantly more dispersive and exhibit dominant contributions from Li(2s) and Li(2p) orbitals. When comparing the band structure obtained with SIESTA to that produced by a plane-wave-based code, good agreement is observed for both the valence bands and the lower portion of the conduction manifold. At higher energies, however, notable differences emerge: while the plane-wave calculation yields fully entangled conduction bands, the localized basis set used in SIESTA results in a spurious gap between band manifolds. Although this separation is an artifact of the more limited basis set, the lower conduction bands remain reliable due to their close correspondence with the plane-wave results. This separation is advantageous for model construction, as it simplifies the definition of a smooth and consistent Hamiltonian under perturbations such as structural deformations or doping. In contrast, the use of disentanglement procedures can introduce discontinuities in the extracted Hamiltonian parameters, compromising their transferability. Therefore, the naturally disentangled character of the conduction manifold in SIESTA represents a significant benefit for second-principles modeling, enabling a direct and robust evaluation of electron-lattice coupling parameters via finite-difference methods, as described in Sec. 5.6.3.

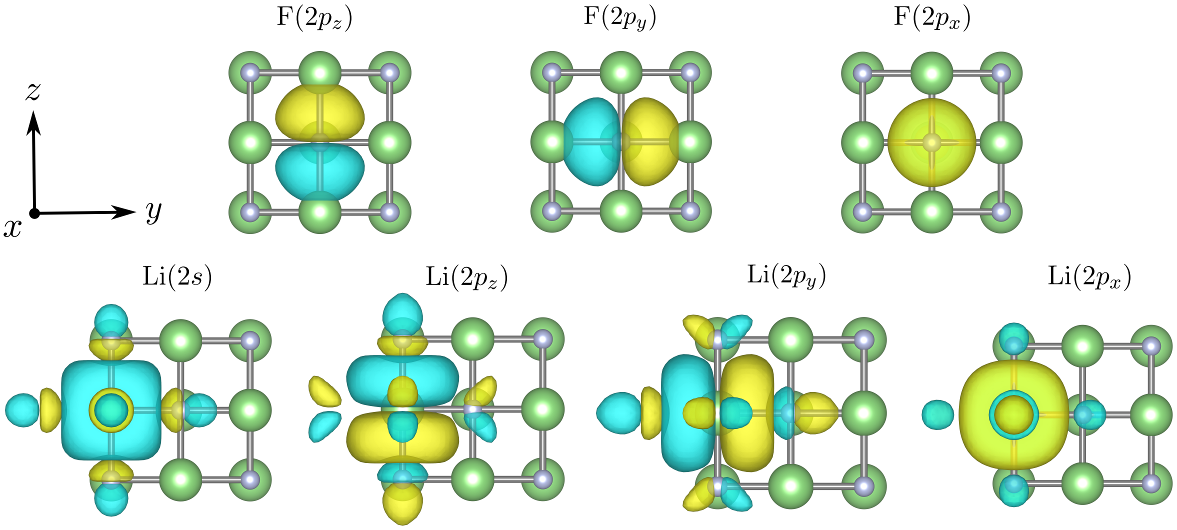


Figure 6.21: Wannier functions of LiF used as the basis for second-principles models that accurately reproduce the top of the valence band with predominant F $(2p)$ character, as well as the lower conduction bands associated with Li$(2s)$ and Li$(2p)$. Green and grey spheres represent Li and F atoms, respectively.

To construct the Wannier functions (WFs) from the DFT data, we project the Bloch wavefunctions of the valence band onto F$(2p)$ atomic orbitals, and those of the conduction band onto Li$(2s)$ and Li$(2p)$ orbitals. Following the procedure used for $\mathrm{SrTiO_3}$, we employ the $H(r)$ function [Eq. (6.1)] to determine a suitable cutoff distance for the Hamiltonian interactions, δr_h, as illustrated in Fig. 6.22. In this case, the spatial extent of the Hamiltonian matrix elements is somewhat larger than that found in $\mathrm{SrTiO_3}$. Nevertheless, adopting $\delta r_\mathrm{h} = 8.0$ Å yields an accurate reproduction of the DFT band structure, with the exception of the uppermost conduction bands located approximately 20 eV above the valence band maximum. These high-energy states lie well beyond the spectral range typically relevant for optical or ultraviolet studies, and also fall within the regime where the accuracy of SIESTA's band structure is limited by the basis set employed.

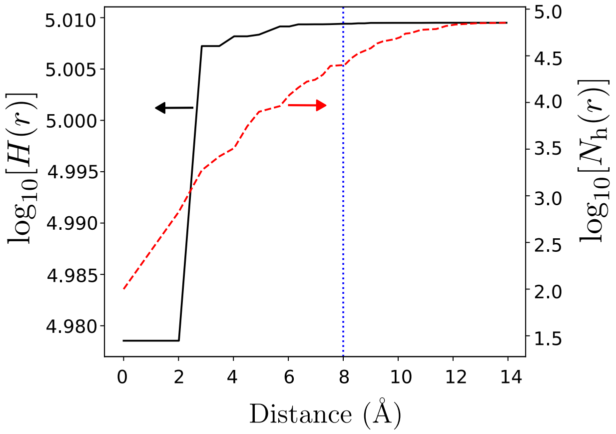


Figure 6.22: Logarithmic plot of the $H(r)$ function (black, left axis) and the number of one-electron Hamiltonian matrix elements as a function of distance (red, right axis). The vertical dashed line indicates the chosen Hamiltonian cutoff distance, $\delta r_\mathrm{h} = 8$Å, used to construct the second-principles model for LiF. This choice corresponds to the band structure shown in red lines in Fig.6.20(a).

6.4.1 Electron-lattice parameters: methodology II

Next, as in the case of SrTiO$_3$, we extend the applicability of the electronic model for LiF beyond the reference atomic geometry (RAG) by incorporating electron-lattice interactions. To this end, we analyze the sensitivity of the model to the electron-lattice cutoff distance, δr_el, which determines the range of atomic pairs contributing to the electron-lattice coupling, as well as the thresholds δf and δg used to discard weak linear and quadratic interactions, respectively. Figure 6.23(a) shows the variation of the goal function as a function of the distortion amplitude d (defined in Sec. 5.6.4) for two different cutoff values, $\delta r_\mathrm{el} = 2.1$ Å and 3.0 Å. The former includes only first-neighbor Li–F interactions, while the latter also captures second-neighbor Li–Li and F–F interactions. Even for relatively large distortions ($d \approx 0.17$ Å), characteristic of thermal motion at approximately 400 K, the total goal function remains low ($\Theta \approx 1$ eV^2), summed over 11964 terms. The model with $\delta r_\mathrm{el} = 3.0$ Å yields an error roughly 60 % smaller than that with $\delta r_\mathrm{el} = 2.1$ Å. However, due to the overall small magnitude of the error, this improvement is of limited practical significance, indicating that the dominant contributions arise from short-range linear interactions. To further reduce the model size, we assess the effect of pruning weak interactions based on the thresholds δf and δg. In Fig. 6.24, we evaluate the model accuracy for several threshold values. A noticeable increase in the goal function is observed when $\delta f = 0.5$ eV/Å and $\delta g = 0.5$ eV/Å^2, compared to the results obtained for $\delta f = 0.1$ eV/Å, $\delta g = 0.1$ eV/Å^2, and the more stringent $\delta f = 0.01$ eV/Å, $\delta g = 0.01$ eV/Å^2. These findings confirm, in line with the SrTiO$_3$ case, that accurate and compact models can be constructed using $\delta f = 0.1$ eV/Å and $\delta g = 0.1$ eV/Å^2 as standard pruning thresholds.

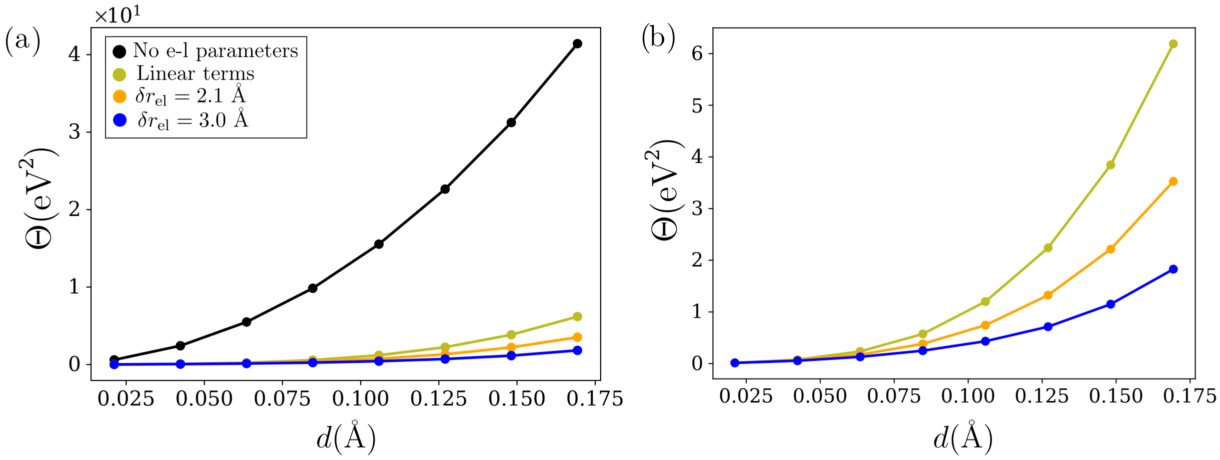


Figure 6.23: Goal function Θ for LiF, measuring the average error per calculation in test set comprising ten configurations with increasing atomic displacements d from the reference atomic geometry (RAG). (a) The thresholds are fixed at $\delta f = 0.1$ eV/Å and $\delta g = 0.1$ eV/Å^2. The black line corresponds to a model excluding electron-lattice interactions; the olive green line includes only linear couplings. Models incorporating both linear and quadratic terms are shown in orange and blue, corresponding to cutoff distances of $\delta r_{\mathrm{el}} = 2.1$ Å and $\delta r_{\mathrm{el}} = 3.0$ Å, respectively. In panel (b), the results corresponding to the model without the electron-lattice coupling correction have been excluded to enable a focused view of the remaining data.

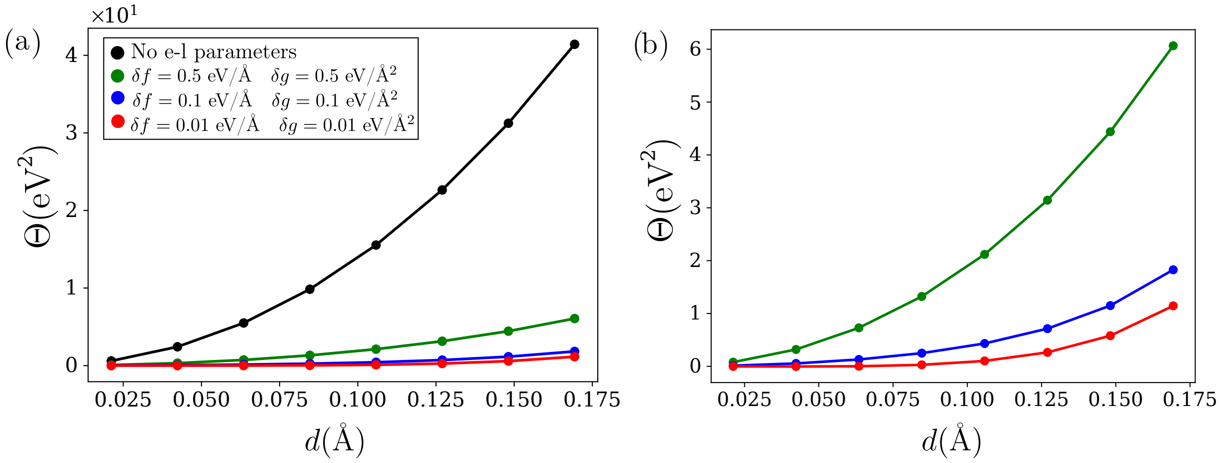


Figure 6.24: Goal function Θ for LiF, measuring the average error per calculation in test set comprising ten configurations with increasing atomic displacements d from the reference atomic geometry (RAG). (a) The electron-lattice cutoff is fixed at $\delta r_{\mathrm{el}} = 3.0$ Å, while different pruning thresholds are explored: $\delta f = 0.5$ eV/Å, $\delta g = 0.5$ eV/Å^2 (green), $\delta f = 0.1$ eV/Å, $\delta g = 0.1$ eV/Å^2 (blue), and $\delta f = 0.01$ eV/Å, $\delta g = 0.01$ eV/Å^2 (red). In panel (b), the results corresponding to the model without the electron-lattice coupling correction have been excluded to enable a focused view of the remaining data.

To assess the distribution of distortion-induced errors across the electronic band structure, we analyze the band-resolved deviations for atomic configurations with a maximum displacement of $d = 0.17$ Å, as shown in Fig. 6.25. These results illustrate how the inclusion of electron-lattice interaction terms progressively improves the model accuracy as the cutoff parameter δr_{el} is increased. Figure 6.25(a) shows the error bars obtained in the absence of electron-lattice corrections, while Figs. 6.25(b) and 6.25(c) correspond to models including electron-lattice interactions with $\delta r_{\mathrm{el}} = 2.1$ Å and $\delta r_{\mathrm{el}} = 3.0$ Å, respectively. The comparison indicates that the dominant contributions to the electron-lattice coupling arise from linear terms and nearest-neighbor Li–F interactions, while the inclusion of second-neighbor Li–Li and F–F couplings yields only modest additional accuracy. Furthermore, we find that the valence bands and the lower part of the conduction manifold exhibit minimal errors across all cases. The largest deviations appear in the upper conduction bands; however, these states lie well above the vacuum ultraviolet range and are thus of limited relevance for most physical applications. Overall, the model demonstrates high reliability for the prediction of spectroscopic and dynamical properties in realistic conditions.

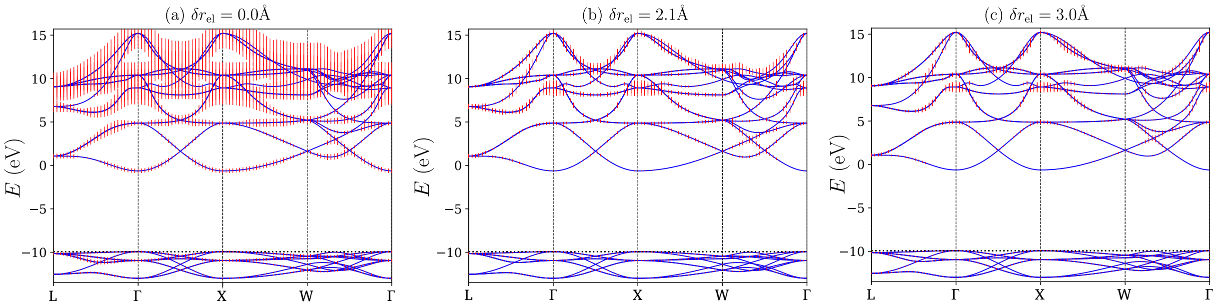


Figure 6.25: Band structures illustrating the statistical error of the second-principles model across a test set of ten distorted geometries. The error bars (in red) represent the standard deviation of the energy levels relative to the reference DFT values. Panel (a) shows results obtained using a second-principles model without electron-lattice interactions. Panels (b) and (c) include linear and quadratic electron-lattice coupling terms with cutoff distances $\delta r_{\mathrm{el}} = 2.1$ Å and 3.0 Å, respectively. All models were generated using thresholds $\delta f = 0.1$ eV/Å and $\delta g = 0.1$ eV/Å^2.

6.4.2 Electron-electron parameters: methodology I

As in the case of SrTiO$_3$ (see Sec. 6.3), the training set for LiF includes electron and hole doping at levels of 0.1, 0.2, and 0.3 electrons, both under spin-restricted (non-magnetic) and spin-polarized conditions. These configurations are used to fit the U and I parameters associated with the valence and conduction bands. Given the known importance of electron-hole interactions in accurately describing the optical properties of LiF, we further include spin-polarized calculations on the neutral system, constraining the total magnetization to 0.5, 1.0, 1.5, and 2.0 electrons. These configurations effectively promote electrons from the valence to the conduction band, enabling a more precise characterization of electron-hole coupling. This comprehensive training strategy allows MODELMAKER to fit interaction terms of the form $U_{\boldsymbol{ab},\boldsymbol{a'b'}}$ and $I_{\boldsymbol{ab},\boldsymbol{a'b'}}$ that couple Wannier functions (WFs) from both the valence and conduction manifolds. When plotting the total goal function Θ for a model that excludes electron-electron interactions, we observe a pattern similar to that found in SrTiO$_3$: the dominant contributions to the error originate from a limited number of interaction groups, primarily associated with intra-atomic interactions [Fig. 6.26(a)]. However, in LiF, we additionally identify several groups corresponding to longer-range interactions that, while smaller, yield non-negligible contributions to Θ. As in the previous case, we adopt a threshold of 0.2 eV^2 to filter relevant electron-electron terms. Beyond the expected intra-atomic contributions, we find that several F($2p$)-F($2p$) Hamiltonian matrix elements exhibit sensitivity to the electronic state. Fitting the U and I constants corresponding to these indices leads to a substantial reduction of the goal function, as shown in Fig. 6.26(b). The most significant improvements stem from three interaction variables: an intra-atomic interaction on Li (contributing 279 eV^2), and two centered on F (together contributing 116 eV^2). Notably, we also identify a significant inter-atomic interaction involving $U_{aa,bb}$ and $I_{aa,bb}$ terms, where WF χ_a has F($2p$) character and WF χ_b corresponds to a first-neighbor Li($2s$) orbital. This interaction alone reduces Θ by approximately 4.5 eV^2, confirming the presence of non-negligible inter-site electron-hole coupling. All other terms contribute less than 0.8 eV^2 to the reduction of the goal function. These results confirm that the method successfully captures strong electron-hole interactions in LiF, in agreement with experimental observations of highly localized excitons in this material [184].

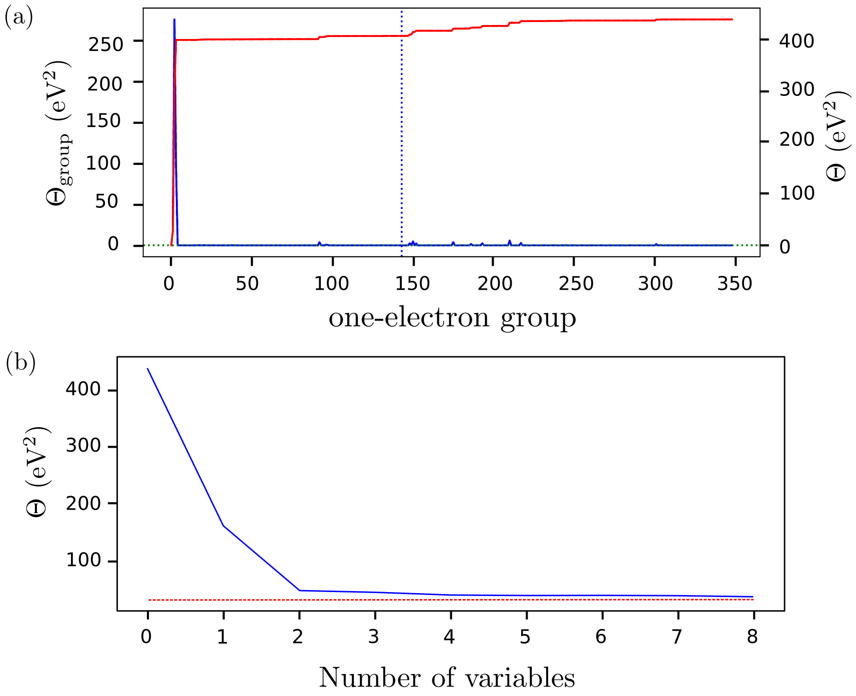


Figure 6.26: (a) Decomposition of the goal function for the electron-electron training set in LiF. The contributions from each group of symmetry-equivalent Hamiltonian matrix elements h_{ab} are shown as a blue curve, Θ_{group}; while the cumulative goal function Θ is plotted in red. The dashed vertical line represents δr_{h}. (b) Evolution of Θ as successive electron-electron interaction terms are incorporated into the second-principles model.

Conclusions and future work

Over the past few decades, computational methods have become a cornerstone in the design and development of advanced materials for technological applications. These include the optimization of battery performance, the design of microelectronic components such as transistors and hard drives, innovations in spintronics, improvements in photovoltaic efficiency, and advances in superconductivity relevant to applications like magnetic resonance imaging and magnetic levitation transport. Within this context, first-principles simulations have revolutionized our ability to predict materials properties from fundamental physical principles. These methods allow the computation of structural, electronic, and thermodynamic properties with remarkable accuracy. However, their practical application remains limited by computational constraints, restricting simulations to small system sizes and short timescales. To overcome these limitations, second-principles methods have emerged as powerful tools that balance quantum-mechanical accuracy with computational efficiency. These methods enable large-scale simulations by systematically constructing effective models based on first-principles data, preserving essential physics while extending reach in both space and time.

This thesis has aimed to contribute to the advancement of second-principles methodologies, particularly in the treatment of electronic degrees of freedom. The work focuses on the development of a robust, quasi-automated framework for constructing second-principles Hamiltonians expressed in a Wannier function basis set, trained on high-fidelity DFT data. Through innovations in training set generation, symmetry handling, and interaction hierarchy, the proposed approach enhances the scalability, accuracy, and interpretability of model construction with respect to the existing methodology [37]—paving the way for predictive simulations of technologically relevant materials.

The first step in this process has been to study the properties and characteristics of Wannier functions, as the methodology relies on expressing the Hamiltonian matrix elements in this basis. For this reason, it has been essential to understand and demonstrate both the fundamental properties of these functions—particularly their real-space localization—and the procedure to obtain the Hamiltonian matrix elements within this basis. Accordingly, we have shown that the interpolated band structures computed with WANNIER90 methodology converge asymptotically to the *ab initio* bands as the density of the $\vec{k}$-point mesh increases, provided it is sufficiently fine for the system under study. Throughout the thesis, we have emphasized the localized character of Wannier functions, illustrating the decay of interaction terms as a function of the distance between them. This localization is a cornerstone in second-principles methodologies: for the resulting models to remain computationally tractable, the interaction parameters must decay with distance. Shorter interaction ranges enable smaller, more efficient training sets, thereby reducing the computational cost of the derived models.

Once the characteristics of the basis functions underlying the method were established, we proceeded to highlight the most important results of the thesis regarding the developed methodology.

In pursuit of the main objective of this thesis, considerable effort has been dedicated to the development of a quasi-automated methodology for constructing second-principles electronic models, enabling efficient and physically grounded parametrization of complex materials. Beyond the methodology used to compute model parameters, the selection of the training set has represented a critical component in ensuring both the efficiency and accuracy of the method.

A major limitation of the *previous methodology* [37] was the absence of a well-defined systematic approach to the construction of the training set, particularly for electron-electron interactions. The process was system-dependent, leading to parameters that were highly reliant on the specific, non-reproducible training set used. Additionally, creating a meaningful and consistent training set required a deep understanding of the system's physics and computational details. In this work, we have been able to address these limitations by establishing a well-defined and systematic framework for training set construction. For the electron-lattice interaction parameters—computed using finite-difference techniques—the training set is rigorously defined, encompassing all symmetry-nonequivalent displacements of one or two atoms along the three Cartesian directions and in both positive and negative senses. This ensures completeness and consistency in the representation of physical perturbations. Regarding electron-electron interactions, although the parameters are computed by a fitting procedure, the training set has been generated through a systematic and reproducible set of calculations, characterized by different charges and spin configurations providing a more robust foundation for parameter extraction.

Furthermore, the implemented methodology accounts for the symmetry properties of the system. This aspect is critical for avoiding numerical artifacts, such as the spurious splitting of energy bands at high-symmetry points in the Brillouin zone where degeneracies are enforced by symmetry. Additionally, the incorporation of symmetry significantly improves the efficiency of the model calculations. As we have seen, its implementation leads to a reduction by several orders of magnitude in both the number of required calculations for the training set and the number of independent parameters that must be computed within the model. In $SrTiO_3$, the number of single-point calculations required to construct the electron-lattice model is 57840 without symmetry. This number decreases to 15 when symmetry is applied with an electron-lattice cutoff of $\delta r_{\text{el}} = 2.0$ Å, or 209 for a cutoff of $\delta r_{\text{el}} = 5.6$ Å. Similarly, in LiF, the calculations reduce from 2353 without symmetry to 21 with symmetry and a cutoff of $\delta r_{\text{el}} = 3.0$ Å.

In this thesis, the validity of the proposed methodology has been rigorously assessed through comparisons with first-principles simulations considering different configurations. The results demonstrate that the model accuracy depends primarily on the character of the electronic states and the magnitude of the applied perturbations. Nonetheless, the valence bands and the lower part of the conduction band-those most relevant for optical and transport phenomena-are consistently well reproduced, even under substantial structural distortions. These results have emerged from the analysis of two representative systems used to validate the proposed methodology: the semicovalent transition-metal perovskite $SrTiO_3$ and the wide-band-gap insulator with a rock-salt structure, LiF. In both cases, the study confirmed the essential role of electron-lattice coupling parameters in accurately capturing the evolution of the electronic band structure across a variety of atomic configurations. Furthermore, the incorporation of electron-electron interaction parameters into the model resulted in a marked reduction of the goal function—which quantifies the quadratic error—highlighting their critical importance in achieving quantitative agreement with first-principles data.

Beyond demonstrating efficiency and accuracy—since both electron-lattice and electron-electron corrections are crucial for reliably reproducing systems with geometric distortions and variations in electronic configurations—the developed methodology also reveals additional potential. With respect to electron-lattice coupling, the systematic refinement of models via variation of cutoffs demonstrates two major advantages: (i) it enables a controlled trade-off between computational cost and accuracy, as less restrictive cutoffs yield more accurate models, and (ii) it provides insights into the relevant physical interactions within the system. For instance, in the case of $SrTiO_3$, our analysis of the interatomic distance cutoff, δr_{el}, revealed that, in addition to the linear electron-lattice coupling parameters, quadratic contributions—particularly those involving first-neighbor Ti–O interactions—are essential. Moreover, while including Ti–O couplings is necessary, it is not sufficient for full accuracy. For example, we have found that, in order to properly capturing the effects of strong electron-lattice coupling—particularly in octahedra rotation modes—it is essential to include quadratic oxygen–oxygen interactions within individual TiO_6 octahedra. These results underline the need to go beyond pairwise interactions limited to nearest neighbors in the five-atom unit cell. In the case of LiF, our methodology identified the dominant contributions as arising from short-range linear interactions, particularly between nearest-neighbor Li and F atoms. The inclusion of second-nearest-neighbor interactions, such as Li–Li and F–F couplings, results in only modest improvements in accuracy, indicating their secondary importance.

The analysis of the developed models has also identified the dominant contributions to electron-electron and electron-hole interactions, which are of critical importance for understanding phenomena such as optical absorption or highly localized excitons in LiF material, considering experimental evidence [184]. In particular, our results demonstrate that intra-atomic interactions are the primary contributors to the electron-electron terms in both systems under study. In the case of the $SrTiO_3$, the methodology captures the highly localized nature of the Ti-centered $3d$-like Wannier functions, as these contribute most significantly to the electron-electron interaction terms. In LiF, intra-atomic interactions also play a major role, although inter-site electron-hole couplings are found to be non-negligible.

An additional advantage of the proposed method is its use of clearly defined cutoff thresholds and a hierarchical model construction, ensuring that larger models systematically improve both predictive accuracy and physical interpretability in comparison to simpler ones. Indeed, this represents a key improvement over the initial implementation in Ref. [37] as well as over machine learning models.

In the current context of model development, it is worth reflecting on the increasing popularity of machine-learning (ML) techniques for materials modeling. ML models have demonstrated impressive accuracy and efficiency, particularly in molecular dynamics and electronic structure predictions. However, these approaches face important limitations relative to physically motivated models. First, ML models typically act as black boxes, lacking transparent traceability between inputs and outputs. Second, their parameters often lack direct physical interpretation, hindering microscopic insight. For example, our linear electron-lattice couplings $\vec{f}_{aa,\lambda}$ directly quantify the force on a Wannier function χ_a due to the motion of atom λ, providing an intuitive connection between model and physics. Third, and most critically, ML models lack standardized metrics for transparency or model comparison. In contrast, our method uses clearly defined cutoff thresholds and hierarchical model construction, ensuring that expanded models systematically improve upon simpler ones—both in predictive power and physical interpretability.

The future work outlined in this thesis has three clear and ambitious directions.

Firstly, theoretical extensions as the model construction algorithm is poised for further refinement, including the incorporation of strain effects and improved treatment of electron-electron interactions.

While this thesis has primarily focused on capturing the effects of geometric distortions due to atomic displacements in electron-lattice interactions, one of the key areas for future development is the incorporation of strain effects into the methodology. Strain is a crucial factor in materials such as ferroelectric perovskite oxides (e.g., $BaTiO_3$ and $PbTiO_3$). Incorporating strain effects into the model generation will enable accurate capture of the electron-lattice interactions in these materials, which are vital for predicting their functional properties. For instance, variations in the Ti–O bond length influence the magnitude and symmetry of the crystal field splitting, while changes in Ti–O–Ti bond angles govern the strength and sign of magnetic superexchange interactions. When combined with a second-principles lattice model, as discussed in Refs. [36; 51], the resulting fully parameterized model will be capable of describing both structural and electronic properties, including those arising from strain effects.

Regarding electron-electron interactions, while electron-lattice couplings can be directly computed from first-principles by analyzing the changes in specific Hamiltonian matrix elements, h_{ab}, under targeted atomic displacements, direct calculation of electron-electron interaction terms, such as $U_{ab,a'b'}$, are not yet available, primarily due to the challenges in controlling charge density perturbations in a physically meaningful manner. Advancements in this area could also lead to improvements in DFT+U methodologies and related approaches.

The second direction of the future work is focused on application of the methodology to new materials. This work has focused on studying semicovalent and ionic crystals, yet an open avenue for future research is validating the methodology in covalent crystals such as silicon, which exhibits an indirect gap relevant for optical spectrum observations, and in metals such as lithium.

Finally, the third direction for future work aims at expanding the methodology to tackle new challenges.

In comparison to earlier SPDFT formulations [37], this methodology represents a substantial advancement in terms of generality, automation, and physical interpretability. Our goal, however, extends beyond merely reproducing DFT results. We aim to develop models capable of reliably predicting properties not included in the training data, such as many-body interactions and finite-temperature effects, which are typically challenging to address with first-principles methods. To achieve this, we propose employing the complementary real-time, time-dependent SPDFT framework developed by our group [173], which allows the computation of dynamic observables such as currents induced by time-dependent electric fields. This framework is particularly well-suited for studying emergent quasiparticles such as polarons and excitons, whose behavior arises from strong electron-lattice and electron-electron (or electron-hole) interactions. A particularly challenging application of this approach is the description of polaron formation in $SrTiO_3$, where conventional functional LDA fails to capture polaronic behavior due to the lack of self-interaction correction. Future work will focus on constructing second-principles models based on hybrid functionals, such as the recently implemented HSE06 functional [200] in SIESTA software, which provides a more accurate treatment of electron localization. This will enable more realistic simulations of polaron formation and dynamics. Although the analysis of these effects falls outside the scope of this thesis, the models developed here have already demonstrated their potential in describing optical phenomena in the studied materials.

In summary, the framework presented here opens a path toward efficient, transparent, and physically insightful second-principles modeling. We anticipate that these tools will prove valuable in the prediction and interpretation of complex phenomena across a wide range of materials in the near future.

Conclusiones y perspectiva futura

Durante las últimas décadas, los métodos computacionales se han consolidado como una herramienta fundamental en el diseño y desarrollo de materiales avanzados con aplicaciones tecnológicas. Entre estas se incluyen la optimización del rendimiento de baterías, el diseño de componentes microelectrónicos como transistores y discos duros, innovaciones en espintrónica, mejoras en la eficiencia fotovoltaica y avances en superconductividad, relevantes para aplicaciones como la resonancia magnética y el transporte por levitación magnética. En este contexto, las simulaciones de primeros principios han revolucionado nuestra capacidad para predecir propiedades de materiales a partir de principios físicos fundamentales. Estos métodos permiten calcular propiedades estructurales, electrónicas y termodinámicas con notable precisión. Sin embargo, su aplicación práctica está limitada por el elevado coste computacional, lo que restringe su uso a sistemas de pequeño tamaño y escalas de tiempo reducidas. Para superar estas limitaciones, han surgido los métodos de segundos principios como herramientas potentes que equilibran la precisión cuántica con una mayor eficiencia computacional. Estos métodos permiten simulaciones a gran escala mediante la construcción sistemática de modelos efectivos basados en datos de primeros principios, conservando la física esencial al tiempo que amplían el alcance en espacio y tiempo.

El objetivo principal de esta tesis ha sido contribuir al desarrollo de metodologías de segundos principios, con un enfoque específico en el tratamiento de los grados de libertad electrónicos. El trabajo se centra en la elaboración de un marco robusto y cuasi-automatizado para la construcción de hamiltonianos de segundos principios en una base de funciones de Wannier, entrenados con datos precisos obtenidos mediante DFT. A través de innovaciones en la generación del conjunto de entrenamiento, el manejo de simetrías y la jerarquía de interacciones, la metodología propuesta mejora la escalabilidad, la precisión y la interpretabilidad de los modelos, abriendo el camino a simulaciones predictivas de materiales tecnológicamente relevantes.

El primer paso de este proceso ha sido el estudio de las propiedades y características de las funciones de Wannier, ya que la metodología se basa en expresar los elementos de matriz del hamiltoniano en dicha base. Ha sido esencial comprender y demostrar tanto las propiedades fundamentales de estas funciones —particularmente su localización en el espacio real— como el procedimiento para obtener los elementos de matriz del hamiltoniano en esta representación. Se ha demostrado que las bandas interpoladas obtenidas mediante la metodología implementada en WANNIER90 convergen asintóticamente a las bandas *ab initio* al incrementar la densidad de la malla de puntos $\vec{k}$, siempre que esta sea suficientemente fina para el sistema en estudio. A lo largo de la tesis se ha enfatizado el carácter localizado de las funciones de Wannier, ilustrando cómo los términos de interacción decaen con la distancia entre ellos. Esta localización es fundamental en los métodos de segundos principios: para que los modelos resultantes sean computacionalmente manejables, los parámetros de interacción deben decaer con la distancia. Rangos de interacción más cortos permiten conjuntos de entrenamiento en superceldas más pequeñas y por lo tanto con un impacto computacional más eficiente.

Una vez establecidas las características de las funciones base del método, se presentan los resultados más relevantes de la metodología desarrollada.

En pos del objetivo principal de esta tesis, se ha dedicado un esfuerzo considerable al desarrollo de una metodología cuasi-automatizada para la construcción de modelos electrónicos de segundos principios, permitiendo una parametrización eficiente y físicamente fundamentada de materiales complejos. Más allá de la metodología utilizada para calcular los parámetros del modelo, la selección del conjunto de entrenamiento ha sido un componente crítico para garantizar tanto la eficiencia como la precisión del método. Una de las principales limitaciones de metodologías previas [37] era la ausencia de un procedimiento sistemático bien definido para construir el conjunto de entrenamiento, especialmente en lo que respecta a las interacciones electrón-electrón. El proceso dependía del sistema y conducía a parámetros altamente sensibles a conjuntos específicos y poco reproducibles. Además, construir un conjunto de entrenamiento significativo y coherente requería un conocimiento profundo de la física del sistema. En este trabajo, se han superado estas limitaciones mediante el establecimiento de un marco sistemático y bien definido para la construcción del conjunto de entrenamiento. Para los parámetros de interacción electrón-red —calculados mediante técnicas de diferencias finitas—, el conjunto de entrenamiento está definido por el método, incluyendo todos los desplazamientos simétricamente no equivalentes de uno o dos átomos en las tres direcciones cartesianas y en ambos sentidos. Esto garantiza una representación completa y coherente de las perturbaciones físicas. En cuanto a las interacciones electrón-electrón, aunque los parámetros se obtienen mediante un procedimiento de ajuste, el conjunto de entrenamiento se ha generado de forma sistemática y reproducible, utilizando configuraciones con diferentes cargas y estados de espín, proporcionando una base más robusta para la extracción de parámetros.

Además, la metodología implementada tiene en cuenta las propiedades de simetría del sistema. Este aspecto es fundamental para evitar problemas numéricos, como la ruptura de la degeneración de las bandas de energía en puntos de alta simetría del espacio recíproco, donde las degeneraciones existentes están impuestas por la simetría del sistema. La incorporación de la simetría mejora notablemente la eficiencia del cálculo de modelos. Como se ha demostrado, su implementación reduce en varios órdenes de magnitud tanto el número de cálculos necesarios en el conjunto de entrenamiento como el número de parámetros independientes. Por ejemplo, en $SrTiO_3$, el número de cálculos necesarios para construir el modelo electrón-red es de 57840 sin simetría, mientras que se reduce a 15 al aplicar simetría con un corte de $\delta r_{\mathrm{el}} = 2.0$ Å, o a 209 con un corte de δr_el $= 5.6$ Å. De forma análoga, en LiF los cálculos se reducen de 2353 a 21 con simetría y un corte de δr_el $= 3.0$ Å.

La validez de la metodología propuesta se ha evaluado rigurosamente mediante comparaciones con simulaciones de primeros principios para diferentes configuraciones. Los resultados muestran que la precisión del modelo depende principalmente del carácter de los estados electrónicos y de la magnitud de las perturbaciones aplicadas. Sin embargo, las bandas de valencia y la parte baja de la banda de conducción —aquellas más relevantes para fenómenos ópticos y de transporte— se reproducen consistentemente bien, incluso ante distorsiones estructurales significativas. Estos resultados se han obtenido del análisis de dos sistemas representativos: el perovskita de transición semicovalente $SrTiO_3$ y el aislante iónico con estructura tipo sal-gema LiF. En ambos casos, se confirma el papel esencial de los parámetros de acoplamiento electrón-red para describir adecuadamente la evolución de la estructura de bandas en diversas configuraciones atómicas. La incorporación de interacciones electrón-electrón también ha contribuido significativamente a reducir la función objetivo —que cuantifica el error cuadrático—, lo cual resalta su importancia para lograr una concordancia cuantitativa con los datos de primeros principios.

Más allá de su eficiencia y precisión, la metodología desarrollada también presenta un gran potencial adicional. En relación al acoplamiento electrón-red, el refinamiento sistemático del modelo mediante la variación del parámetro de corte revela dos ventajas clave: (i) permite ajustar el compromiso entre coste computacional y precisión, y (ii) proporciona información física sobre las interacciones relevantes del sistema. En el caso de $SrTiO_3$, el análisis del parámetro de corte δr_{el} ha mostrado que, además de los términos lineales, es esencial incluir contribuciones cuadráticas, particularmente aquellas que involucran interacciones Ti–O de primeros vecinos. Además, se ha evidenciado que incluir solo interacciones Ti–O no es suficiente; es imprescindible considerar también interacciones cuadráticas oxígeno–oxígeno dentro de cada octaedro TiO_6.

En el caso de LiF, la metodología ha identificado como dominantes las interacciones lineales de corto alcance, especialmente entre átomos de Li y F de primeros vecinos. La inclusión de interacciones de segundos vecinos como Li–Li y F–F aporta mejoras marginales, indicando su menor relevancia.

El análisis de los modelos también ha permitido identificar las principales contribuciones a las interacciones electrón-electrón y electrón-hueco, fundamentales en fenómenos como la absorción óptica o la formación de excitones localizados en LiF, en concordancia con la evidencia experimental [184]. En ambos sistemas, se ha encontrado que las interacciones intraatómicas son las contribuyentes principales. En $SrTiO_3$, la metodología capta la naturaleza altamente localizada de las funciones de Wannier tipo $3d$ centradas en Ti. En LiF, las interacciones intraatómicas también son dominantes, aunque se detectan acoplamientos intersitio electrón-hueco no despreciables.

En el contexto actual, vale la pena reflexionar sobre la creciente popularidad de los enfoques de aprendizaje automático (ML) en la modelización de materiales. Aunque los modelos ML han demostrado gran precisión y eficiencia, especialmente en dinámica molecular y predicciones de estructura electrónica, enfrentan limitaciones significativas frente a modelos físicamente motivados. Primero, actúan como cajas negras, dificultando la trazabilidad entre entrada y salida. Segundo, sus parámetros carecen generalmente de una interpretación física directa. Por ejemplo, los acoplamientos lineales de nuestros modelos

$f_{aa,\lambda}$ cuantifican directamente la fuerza sobre una función de Wannier χ_a provocada por el desplazamiento del átomo λ, estableciendo un vínculo claro entre modelo y física. Tercero, los modelos ML carecen de métricas estandarizadas para evaluar su transparencia o comparabilidad. En contraste, nuestra metodología garantiza la mejora sistemática de los modelos mediante umbrales definidos y jerarquías estructuradas.

En relación a lo anterior, una ventaja adicional del método propuesto es el uso de umbrales de corte bien definidos y una construcción jerárquica del modelo, asegurando que modelos más complejos mejoren sistemáticamente tanto la precisión predictiva como la interpretación física respecto a versiones más simples. Esto representa una mejora sustancial con respecto a implementaciones previas [37] y frente a modelos basados en técnicas de aprendizaje automático.

El trabajo futuro atendiendo a los avances realizados en esta tesis contempla tres direcciones ambiciosas. La primera se orienta al perfeccionamiento teórico del algoritmo de construcción del modelo, incluyendo la incorporación de efectos de deformación (strain) y una mejor descripción de las interacciones electrón-electrón. Aunque esta tesis se ha centrado en desplazamientos atómicos, los efectos de strain son cruciales en materiales como perovskitas ferroeléctricas. Su incorporación permitirá capturar variaciones en la simetría del campo cristalino y en las interacciones magnéticas de superintercambio. Junto con modelos de red de segundos principios [36; 51], se obtendrán modelos completos capaces de describir tanto propiedades estructurales como electrónicas.

En cuanto a las interacciones electrón-electrón, aún no se dispone de cálculos directos para los términos $U_{\boldsymbol{ab},\boldsymbol{a}'\boldsymbol{b}'}$ debido a la dificultad de controlar perturbaciones en la densidad de carga de forma físicamente significativa. Avances en este aspecto también podrían beneficiar metodologías como DFT+U.

La segunda dirección futura se enfoca en la aplicación de la metodología a nuevos materiales. Aunque esta tesis ha abordado cristales semicovalentes e iónicos, queda abierta la validación en materiales covalentes como el silicio o metálicos como el litio.

Finalmente, la tercera dirección de trabajo futuro consiste en extender la metodología a nuevos desafíos, como predecir propiedades fuera del conjunto de entrenamiento, incluyendo efectos de muchos cuerpos y de temperatura finita. Para ello, se propone utilizar el marco complementario desarrollado recientemente por nuestro grupo [173], basado en SPDFT dependiente del tiempo en tiempo real. Este enfoque permite calcular observables dinámicos, como corrientes inducidas por campos eléctricos, y es adecuado para estudiar cuasipartículas emergentes como polarones y excitones. Un reto particularmente desafiante es la formación de polarones en $SrTiO_3$, donde los funcionales convencionales (LDA) no los describen adecuadamente. En el futuro se desarrollarán modelos de segundos principios basados en funcionales híbridos, como HSE06 [200], recientemente implementado en SIESTA, lo que permitirá simulaciones más realistas de formación y dinámica de polarones.

En resumen, el marco metodológico presentado en esta tesis abre una vía prometedora hacia una modelización eficiente, transparente y físicamente fundamentada basada en segundos principios. Se anticipa que estas herramientas serán valiosas en el futuro próximo para la predicción e interpretación de fenómenos complejos en una amplia gama de materiales.

Bibliography

[1] JPT. Training a single ai model can emit as much carbon as five cars in their lifetimes, 2023. Accessed: 2025-02-05.

[2] National Renewable Energy Laboratory (NREL). Nrel joins effort to advance data center cooling efficiency, 2023. Accessed: 2025-02-05.

[3] The sustainable development goals of the United Nations can be found in https://sdgs.un.org/goals.

[4] G. Ceder, K. Persson, D. Levitan, M. Fessenden, L. Greenemeier, L. Billings, K. H. Courage, C. Q. Choi, D. Yuhas, and D. F. Maron. World Changing Ideas. SCI AM **309**, 34 (2013).

[5] G. Moody and M. S. Islam. Materials for ultra-efficient, high-speed optoelectronics. MRS Bull. **47**, 475 (2022).

[6] S. C. Tian, M. Ahamed, and D. Bimberg. Progress in short wavelength energy-efficient high-speed vertical-cavity surface-emitting lasers for data communication. MDPI, 2023. Photonics **10**, 410 (2023).

[7] G. E. Moore. Cramming more components onto integrated circuits. IEEE Solid-State Circuits Society Newsletter **11**, 33 (2006).

[8] T. Gao, Y. Guo, B. Zhang, P. Cicotti, Y. Lu, P. Balaji, and M. Taufer. Memory-Efficient and Skew-Tolerant MapReduce Over MPI for Supercomputing Systems. IEEE Transactions on Parallel and Distributed Systems **31**, 2734 (2020).

[9] Leonardo, one of the most powerful supercomputers in the world, reached a performance of over 250 PFlops.

[10] P. F. Zubieta Rico, L. Schneider, G. R. Pérez-Lemus, R. Alessandri, S. Dasetty, T. D. Nguyen, C. A. Menéndez, Y. Wu, Y. Jin, Y. Xu, S. Varner, J. A. Parker, A. L. Ferguson, J. K. Whitmer, and J. J. de Pablo. PySAGES: flexible, advanced sampling methods accelerated with GPUs. Npj Comput. Mater. **10**, 35 (2024).

[11] K. Liegener, O. Morsch, and G. Pupillo. Solving quantum chemistry problems on quantum computers. Phys. Today **77**, 34 (2024).

[12] C. Lichtensteiger, P. Zubko, M. Stengel, P. Aguado-Puente, J. M. Triscone, P. Ghosez, and J. Junquera. Ferroelectricity in Ultrathin-Film capacitors. Oxide Ultrathin Films (2012).

[13] J. E. Lennard-Jones. Cohesion. Proc. Phys. Soc. **43**, 461 (1931).

[14] J. Tersoff. New empirical approach for the structure and energy of covalent systems. Phys. Rev. B **37**, 6991 (1988) .

[15] T. P. Senftle, S. Hong, M. M. Islam, S. B. Kylasa, Y. Zheng, Y. K. Shin, C. Junkermeier, R. Engel-Herbert, M. J. Janik, H. M. Aktulga, T. Verstraelen, A. Grama, and A. C. T. van Duin. The ReaxFF reactive force-field: development, applications and future directions. Npj Comput. Mater. **2**, 15011 (2016).

[16] K. M. Rabe and J. D. Joannopoulos. Theory of the structural phase transition of GeTe. Phys. Rev. B **36**, 6631 (1987).

[17] W. Zhong and D. Vanderbilt. Competing Structural Instabilities in Cubic Perovskites. Phys. Rev. Lett. **74**, 2587 (1995).

[18] W.A. Harrison. Tight-binding methods. Surf. Sci. **299**, 298 (1994).

[19] K. I. Kugel and D. I. Khomskii. Crystal-structure and magnetic properties of substances with orbital degeneracy. Zh. Eksp. Teor. Fiz **64**, 1429 (1973).

[20] J. Spałek. Effect of pair hopping and magnitude of intra-atomic interaction on exchange-mediated superconductivity. Phys. Rev. B **37**, 533 (1988).

[21] J. Hubbard. Electron correlations in narrow energy bands. Proc. R. soc. Lond. Ser. A **276**, 238 (1963).

[22] S. Lifson and A. Warshel. Consistent force field calculations. II. crystal structures, vibrational frequencies, and thermodynamic quantities of cyclic polyhydrides and polyhalides. J. Chem. Phys. **49**, 5116 (1968).

[23] I. D. Brown. Recent Developments in the Methods and Applications of the Bond Valence Model. Chem. Rev. **109**, 6858 (2009).

[24] W. Zhong, D. Vanderbilt, and K. M. Rabe. Phase Transitions in $BaTiO_3$ from First Principles. Phys. Rev. Lett. **73**, 1861 (1994).

[25] W. Zhong, D. Vanderbilt, and K. M. Rabe. First-principles theory of ferroelectric phase transitions for perovskites: The case of $BaTiO_3$. Phys. Rev. B **52**, 6301 (1995).

[26] A. J. Millis, B. I. Shraiman, and R. Mueller. Dynamic Jahn-Teller Effect and Colossal Magnetoresistance in $La_{1-x}Sr_xMnO_3$. Phys. Rev. Lett. **77**, 175 (1996).

[27] M. Stengel. First-Principles Modeling of Electrostatically Doped Perovskite Systems. Phys. Rev. Lett. **106**, 136803 (2011).

[28] I. A. Kornev, S. Lisenkov, R. Haumont, B. Dkhil, and L. Bellaiche. Finite-Temperature Properties of Multiferroic $BiFeO_3$. Phys. Rev. Lett. **99**, 227602 (2007).

[29] J. Schmidt, M.R.G. Marques, S. Botti, and M.A.L. Marques. Recent advances and applications of machine learning in solid-state materials science. Npj Comput. Mater. **5**, 83 (2019).

[30] Z. J. Baum, X. Yu, P. Ayala, Y. Zhao, S. P. Watkins, and Q. Zhou. Artificial intelligence in chemistry: Current trends and future directions. J. Chem. Inf. Model. **61**, 3197 (2021).

[31] H. Li, Z. Wang, N. Zou, M. Ye, R. Xu, X. Gong, W. Duan, and Y. Xu. Deep-learning density functional theory Hamiltonian for efficient *ab initio* electronic-structure calculation. Nat. Comput. Sci. **2**, 367 (2022).

[32] Deeph (deep hamiltonian) method. `https://github.com/mzjb/DeepH-pack`. Accessed: 1 April 2025.

[33] Y. Zhuo, A. Mansouri T., and J. Brgoch. Predicting the band gaps of inorganic solids by machine learning. J. Phys. Chem. Lett. **9**, 1668 (2018).

[34] J. D. Evans and F.-X. Coudert. Predicting the mechanical properties of zeolite frameworks by machine learning. Chem. Mater. **29**, 7833 (2017).

[35] Ph. Ghosez and J. Junquera. Modeling of Ferroelectric Oxide Perovskites: From First to Second Principles. Annu. Rev. Condens. Matter Phys. **13**, 325 (2022).

[36] J. C. Wojdeł, P. Hermet, M. P. Ljungberg, Ph. Ghosez, and J. Íñiguez. First-principles model potentials for lattice-dynamical studies: general methodology and example of application to ferroic perovskite oxides. J. Phys.: Cond. Matter **25**, 305401 (2013).

[37] P. Garcia-Fernandez, J. C. Wojdeł, J. Íñiguez, and J. Junquera. Second-principles method for materials simulations including electron and lattice degrees of freedom. Phys. Rev. B **93**, 195137 (2016).

[38] 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. J. Condens. Matter Phys. **14**, 2745 (2002).

[39] A. A. Mostofi, J. R. Yates, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari. Wannier90: A Tool for Obtaining Maximally-Localised Wannier Functions. Comput. Phys. Commun. **178**, 685 (2008).

[40] G. Pizzi, V. Vitale, R. Arita, S. Blügel, F. Freimuth, G. Géranton, M. Gibertini, D. Gresch, Ch. Johnson, T. Koretsune, J. Ibañez-Azpiroz, H. Lee, J.-M. Lihm, D. Marchand, A. Marrazzo, Y. Mokrousov, J. I. Mustafa, Y. Nohara, Y. Nomura, L. Paulatto, S. Poncé, Th. Ponweiser, J. Qiao, F. Thöle, S. S. Tsirkin, M. Wierzbowska, N. Marzari, D. Vanderbilt, I. Souza, A. A Mostofi, and J. R. Yates. WANNIER90 as a community code: new features and applications. J. Phys.: Condens. Matter **32**, 165902 (2020).

[41] M. E. Lines. Statistical Theory for Displacement Ferroelectrics. Phys. Rev. **177**, 797 (1969).

[42] D. Amoroso, A. Cano, and Ph. Ghosez. First-principles study of $(\mathrm{Ba,Ca})\mathrm{TiO_3}$ and $\mathrm{Ba(Ti,Zr)O_3}$ solid solutions. Phys. Rev. B **97**, 174108 (2018).

[43] W. Zhong and D. Vanderbilt. Effect of quantum fluctuations on structural phase transitions in $\mathrm{SrTiO_3}$ and $\mathrm{BaTiO_3}$. Phys. Rev. B **53**, 5047 (1996).

[44] J. Íñiguez and D. Vanderbilt. First-Principles Study of the Temperature-Pressure Phase Diagram of $\mathrm{BaTiO_3}$. Phys. Rev. Lett. **89**, 115503 (2002).

[45] U. V. Waghmare and K. M. Rabe. *Ab initio* statistical mechanics of the ferroelectric phase transition in $\mathrm{PbTiO_3}$. Phys. Rev. B **55**, 6161 (1997).

[46] U. V. Waghmare, K. M. Rabe, H. Krakauer, R. Yu, and C.-Z. Wang. Effective hamiltonian for the ferroelectric phase transitions in $\mathrm{KNbO_3}$. AIP Conf. Proc. **436**, 32 (1998).

[47] D. Vanderbilt and W. Zhong. First-principles theory of structural phase transitions for perovskites: Competing instabilities. Ferroelectrics **206**, 181 (1998).

[48] Y. Yang, B. Xu, C. Xu, W. Ren, and L. Bellaiche. Understanding and revisiting the most complex perovskite system via atomistic simulations. Phys. Rev. B **97**, 174106 (2018).

[49] X. Gonze, B. Amadon, G. Antonius, F. Arnardi, L. Baguet, J.-M. Beuken, J. Bieder, F. Bottin, J. Bouchet, E. Bousquet, N. Brouwer, F. Bruneval, G. Brunin, T. Cavignac, J.-B. Charraud, W. Chen, M. Côté, S. Cottenier, J. Denier, G. Geneste, Ph. Ghosez, M. Giantomassi, Y. Gillet, O. Gingras, D. R. Hamann, G. Hautier, X. He, N. Helbig, N. Holzwarth, Y. Jia, F. Jollet, W. Lafargue-Dit-Hauret, K. Lejaeghere, M. A.L. Marques, A. Martin, C. Martins, H. P.C. Miranda, F. Naccarato, K. Persson, G. Petretto, V. Planes, Y. Pouillon, S. Prokhorenko, F. Ricci, G.-M. Rignanese, A. H. Romero, M. Marcus Schmitt, M. Torrent, M. J. van Setten, B. Van Troeye, M. J. Verstraete, G. Zérah, and J. W. Zwanziger. The ABINIT project: Impact, environment and recent developments. Comput. Phys. Commun. **248**, 107042 (2020).

[50] M.M. Schmitt. *First-and second-principles studies of perovskites*. PhD thesis, Université de Liège (Belgium), 2020.

[51] C. Escorihuela-Sayalero, J. C. Wojdeł, and J. Íñiguez. Efficient systematic scheme to construct second-principles lattice dynamical models. Phys. Rev. B **95**, 094115 (2017).

[52] X. Gonze and J.-P. Vigneron. Density-functional approach to nonlinear-response coefficients of solids. Phys. Rev. B **39**, 13120 (1989).

[53] X. Gonze, Ph. Ghosez, and R.W. Godby. Density-polarization functional theory of the response of a periodic insulating solid to an electric field. Phys. Rev. Lett. **74**, 4035 (1995).

[54] X. Gonze. First-principles responses of solids to atomic displacements and homogeneous electric fields: Implementation of a conjugate-gradient algorithm. Phys. Rev. B **55**, 10337 (1997) .

[55] J.A. Pople and D.L. Beveridge. *Approximate Molecular Orbital Theory*. McGraw-Hill series in advanced chemistry. McGraw-Hill, 1970.

[56] J. Kohanoff. *Electronic structure calculations for solids and molecules: theory and computational methods*. Cambridge University Press, 2006.

[57] M. Stengel. First-principles modeling of electrostatically doped perovskite systems. Phys. Rev. Lett. **106**, 136803 (2011) .

[58] N. W. Ashcroft and N. D. Mermin. *Solid State Physics*. Saunders College Publishing, Philadelphia, 1976.

[59] D. A. Papaconstantopoulos. *Handbook of electronic structure of elemental solids*. Plenum (New York), 1986.

[60] W. A. Harrison. *Elementary electronic structure*. World Publishing (Singapore), 1999.

[61] D. Porezag, Th. Frauenheim, Th. Köhler, G. Seifert, and R. Kaschner. Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon. Phys. Rev. B **51**, 12947 (1995).

[62] J. C. Slater and G. F. Koster. Simplified LCAO method for the periodic potential problem. Phys. Rev. **94**, 1498 (1954).

[63] D. Vanderbilt. *Berry Phases in Electronic Structure Theory: Electric Polarization, Orbital Magnetization and Topological Insulators*. Cambridge University Press, 2018.

[64] G. Wannier. The structure of electronic excitation levels in insulating crystals. Phys. Rev. **52**, 191 (1937).

[65] N. Marzari, A.A. Mostofi, J.R. Yates, I. Souza, and D. Vanderbilt. Maximally localized wannier functions: Theory and applications. Rev. Mod. Phys. **84**, 1419 (2012).

[66] N. Marzari, D. Vanderbilt, and M. C. Payne. Ensemble density-functional theory for *Ab Initio* molecular dynamics of metals and finite-temperature insulators. Phys. Rev. Lett. **79**, 1337 (1997).

[67] Jacques Des Cloizeaux. Energy bands and projection operators in a crystal: Analytic and asymptotic properties. Phys. Rev. **135**, A685 (1964).

[68] N. Marzari and D. Vanderbilt. Maximally localized generalized wannier functions for composite energy bands. Phys. Rev. B **56**, 12847 (1997).

[69] I. Souza, N. Marzari, and D. Vanderbilt. Maximally localized wannier functions for entangled energy bands. Phys. Rev. B **65**, 035109 (2001).

[70] Xiaolong Yang, Zhouyi He, and Xiao Zheng. Unit cell consistency of maximally localized Wannier functions. Electron. Struct. **2**, 014001 (2020).

[71] H.J. Monkhorst and J.D. Pack. Special points for Brillouin-zone integrations. Phys. Rev. B **13**, 5188 (1976).

[72] Alan V Oppenheim, Alan S Willsky, and Syed Hamid Nawab. *Signals & systems*. Pearson Educación, 1997.

[73] I. Bersuker. *The Jahn-Teller Effect*. Cambridge University Press, 2006.

[74] F. Giustino. Electron-phonon interactions from first principles. Rev. Mod. Phys. **89**, 015003 (2017).

[75] J. M. Ziman and P. W. Levy. Electrons and phonons, 1961.

[76] M. Bartomeu. Vibrational averages along thermal lines. Phys. Rev. B **93**, 14302 (2016).

[77] F. Bloch. Über die quantenmechanik der elektronen in kristallgittern. Z. Phys. **52**, 55 (1929).

[78] J. Bardeen. Electrical conductivity of metals. J. Appl. Phys. **11**, 88 (1940).

[79] J. Frenkel. Wave mechanics: elementary theory. Oxford University Press, 1933.

[80] S. Pimputkar, J. S. Speck, S. P. DenBaars, and S. Nakamura. Prospects for LED lighting. Nat. Photonics **3**, 180 (2009).

[81] B. A. Gregg. Excitonic solar cells. J. Phys. Chem. B **107**, 4688 (2003).

[82] M. A. Green, A. Ho-Baillie, and H. J. Snaith. The emergence of perovskite solar cells. Nat. Photonics **8**, 506 (2014).

[83] G. J. Snyder and E. S. Toberer. Complex thermoelectric materials. Nat. Mater. **7**, 105 (2008).

[84] A. Fujishima and K. Honda. Electrochemical photolysis of water at a semiconductor electrode. Nature **238**, 37 (1972).

[85] A. De Sio and C. Lienau. Vibronic coupling in organic semiconductors for photovoltaics. Phys. Chem. Chem. Phys. **19**, 18813 (2017).

[86] L. N. Cooper. Bound electron pairs in a degenerate Fermi gas. Phys. Rev. **104**, 1189 (1956).

[87] A. S. Alexandrov and A. B. Krebs. Polarons in high-temperature superconductors. Sov. Phys. Usp. **35**, 345 (1992).

[88] C. Franchini, M. Reticcioli, M. Setvin, and U. Diebold. Polarons in materials. Nat. Rev. Mater. **6**, 560 (2021).

[89] J. T. Devreese, S. N. Klimin, J. L. M. van Mechelen, and D. van der Marel. Many-body large polaron optical conductivity in $SrTi_{1-x}Nb_xO_3$. Phys. Rev. B **81**, 1252 (2010).

[90] Y. Obara, H. Ito, T. Ito, N. Kurahashi, S. Thürmer, H. Tanaka, T. Katayama, T. Togashi, S. Owada, Y. Yamamoto, et al. Femtosecond time-resolved X-ray absorption spectroscopy of anatase TiO_2 nanoparticles using XFEL. Struct. Dyn. **4**, 044033 (2017).

[91] E. Possenriede, H. Kröse, T. Varnhorst, R. Scharfschwerdt, and O. F. Schirmer. Shallow acceptor and electron conduction states in $BaTiO_3$. Ferroelectrics **151**, 199 (1994).

[92] A. Joshi, S. Saikia, S. Badola, A. Nag, and S. Saha. Defect-mediated electron–phonon coupling in halide double perovskite. Appl. Phys. Lett. **126**, 021902 (2025).

[93] G. Santamaría, T. Fernández-Ruiz, J. M. García-Lastra, P. García-Fernández, I. Sánchez-Movellán, M. Moreno, and J. A. Aramburu. Understanding pressure effects on structural, optical, and magnetic properties of $CsMnF_4$ and other $3d^n$ compounds. Inorg. Chem. **63**, 13231 (2024).

[94] D. Carrasco-Busturia, I. Sánchez-Movellán, A.S. Tygesen, A. Bhowmik, J. M. García-Lastra, J.A. Aramburu, and M. Moreno. Red shift in optical excitations on layered copper perovskites under pressure: role of the orthorhombic instability. Chem. Eur. J. **29**, e202202933 (2023).

[95] K. S. Song and R. T. Williams. Self-trapped excitons. Chapter: Alkali Halides, 123 (1993).

[96] R. H. Borcherts, H. Kanzaki, and H. Abe. EPR Spectrum of a Jahn-Teller System, NaCl: Cu^{+2}. Phys. Rev. B **2**, 23 (1970).

[97] R. E. Coffman. Jahn-Teller Effect in the EPR Spectrum of Cu^{++}: MgO at 1.2K. J. Chem. Phys. **48**, 609 (1968).

[98] J. P. Gordon, H. J. Zeiger, and C. H. Townes. The Maser—New Type of Microwave Amplifier, Frequency Standard, and Spectrometer. Phys. Rev. **95**, 282 (1954).

[99] C. D. Yang and S. Y. Han. Tunneling quantum dynamics in ammonia. Int. J. Mol. Sci. **22**, 8282 (2021).

[100] F. Bouakline. Umbrella inversion of ammonia redux. Phys. Chem. Chem. Phys. **23**, 20509 (2021).

[101] D. A. Morton-Blake. Quantum Mechanical Tunnelling: A chemical example of quantum tunnelling. Senior Sophister, Trinity College Dublin (2011).

[102] I. B. Bersuker, N. B. Balabanov, D. Pekker, and J. E. Boggs. Pseudo Jahn–Teller origin of instability of molecular high-symmetry configurations: Novel numerical method and results. JCP **117**, 10478 (2002).

[103] J. H. Van Vleck. The Jahn-Teller effect and crystalline Stark splitting for clusters of the form XY_6. J. Chem. Phys. **7**, 72 (1939).

[104] P. García-Fernández, I. B. Bersuker, J. A. Aramburu, M. T. Barriuso, and M. Moreno. Origin of warping in the $E\otimes e$ Jahn-Teller problem: Quadratic vibronic coupling versus anharmonicity and application to NaCl: Rh^{2+} and triangular molecules. Phys. Rev. B **71**, 184117 (2005).

[105] J. A. Aramburu, P. García-Fernández, J. M. García-Lastra, and M. Moreno. Jahn–Teller and non-Jahn–Teller systems involving ${CuF_6}^{4-}$ units: role of the internal electric field in Ba_2ZnF_6: Cu^{2+} and other insulating systems. J. Phys. Chem. C **121**, 5215 (2017).

[106] V. V. Sliznev, O. A. Pimenov, and G. V. Girichev. Jahn-Teller effect versus spin-orbit coupling: The structure of the free molybdenum pentafluoride molecule. J. Mol. Struct. **1199**, 126884 (2020).

[107] M. Kataoka. Theory of the antiferroorbital ordering due to the cooperative jahn–teller effect. J. Phys. Soc. Jpn. **70**, 2353 (2001).

[108] J. Varignon, M. Bibes, and A. Zunger. Origins versus fingerprints of the Jahn-Teller effect in d-electron ABX_3 perovskites. Phys. Rev. Res. **1**, 033131 (2019).

[109] T. Fernández-Ruiz, I. Sánchez-Movellán, J. M. García-Lastra, M. Moreno, J. A. Aramburu, and P. García-Fernández. Many-body model for the cooperative Jahn-Teller effect in crystals and its associated orbital ordering. Phys. Rev. B **109**, 205150 (2024).

[110] M. Derzsi, P. Piekarz, and W. Grochala. Structures of Late Transition Metal Monoxides from Jahn-Teller Instabilities in the Rock Salt Lattice. Phys. Rev. Lett. **113**, 025505 (2014).

[111] T. Fernández. *Acoplamiento electrón-vibración en sólidos: desarrollo de técnicas de Segundos Principios dependientes del tiempo para simular transporte y una teoría multielectrónica del Jahn-Teller cooperativo*. PhD thesis, Universidad de Cantabia (España), 2024.

[112] Eduard L Nagaev. *Colossal magnetoresistance and phase separation in magnetic semiconductors*. World Scientific, 2002.

[113] K. I. Kugel and D. I. Khomskii. Crystal-structure and magnetic properties of substances with orbital degeneracy. Zh. Eksp. Teor. Fiz. **64**, 1429 (1973) .

[114] K.I. Kugel and D.I. Khomskii. The Jahn-Teller effect and magnetism: transition metal compounds. Sov. Phys. Uspekhi **25**, 231 (1982).

[115] Y. Tokura and N. Nagaosa. Orbital physics in transition-metal oxides. Sci. **288**, 462 (2000).

[116] J. Van den Brink. Orbital excitations in $LaMnO_3$. Phys. Rev. Lett. **87**, 217202 (2001).

[117] J. Kanamori. Crystal distortion in magnetic compounds. J. Appl. Phys. **31**, S14 (1960) .

[118] I. B. Bersuker and V. Z. Polinger. *Vibronic interactions in molecules and crystals*, volume 49. Springer Science & Business Media, 2012.

[119] V. Polinger. Orbital ordering versus the traditional approach in the cooperative Jahn–Teller effect: a comparative study. In *The Jahn-Teller effect: fundamentals and implications for Physics and Chemistry*, pages 685–725. 2009. .

[120] G. A. Gehring and K. A. Gehring. Co-operative Jahn-Teller effects. Rep. Prog. Phys. **38**, 1 (1975).

[121] H. Kuhn. A quantum-mechanical theory of light absorption of organic dyes and similar compounds. J. Chem. Phys. **17**, 1198 (1949).

[122] C. Carter. Electronic Spectra of Transition Metal Complexes. Lecture Notes, University of Massachusetts Boston.

[123] E. Petryayeva, W. R. Algar, and I. L. Medintz. Quantum dots in bioanalysis: a review of applications across various platforms for fluorescence spectroscopy and imaging. Appl. Spectrosc. **67**, 215 (2013).

[124] M Tanaka and T. Kushida. Cooperative vibronic transitions in eu^{3+}-doped oxide glasses. 60(21):14732, 1999. Phys. Rev. B **60**, 14732 (1999).

[125] G. Viera, S. Huet, and L. Boufendi. Crystal size and temperature measurements in nanostructured silicon using Raman spectroscopy. J. Appl. Phys. **90**, 4175 (2001).

[126] L. M. Malard, M. A. Pimenta, G. Dresselhaus, and M. S. Dresselhaus. Raman spectroscopy in graphene. Phys. Rep. **473**, 51 (2009).

[127] W. Domcke, H. Köppel, and L. S. Cederbaum. Spectroscopic effects of conical intersections of molecular potential energy surfaces. Mol. Phys. **43**, 851 (1981).

[128] M. Fox. *Optical properties of solids*, volume 3. Oxford university press, 2010.

[129] A. Sharma, D. Saklani, K.R.J. Thomas, Shahnawaz, S.S. Swayamprabha, and J.-H. Jou. Synthesis and characterization of multi-substituted carbazole derivatives exhibiting aggregation-induced emission for OLED applications. Org. Electron. **86**, 105864 (2020).

[130] A. Slodek, D. Zych, A. Maron, R. Gawecki, A. Mrozek-Wilczkiewicz, K. Malarz, and et al. Phenothiazine Derivatives – Synthesis, Characterization, and Theoretical Studies with an Emphasis on the Solvatochromic Properties. J. Mol. Liq. **285**, 515–525 (2019).

[131] A. Slodek, D. Zych, A. Maron, J.G. Malecki, S. Golba, G. Szafraniec-Gorol, and et al. Does the length matter? Synthesis, photophysical, and theoretical study of novel quinolines based on carbazoles with different length of alkyl chain. Dyes Pigments **160**, 604 (2019).

[132] M. D. Khitrov, D.N. Platonov, A.Y. Belyy, K.P. Trainov, J.A. Velmiskina, M.G. Medvedev, R.F. Salikov, and Y.V. Tomilov. A unique small molecule class of fluorophores with large stokes shift based on the electron deficient 9-methoxypyrroloisoquinolinetrione core. Dyes and Pigments **203**, 110344 (2022).

[133] David Emin. *Polarons*. Cambridge University Press, 2013.

[134] A. S. Alexandrov and N. F. Mott. *Polarons and bipolarons*. World Scientific, 1996.

[135] A. J. E. Rettie, W. D. Chemelewski, D. Emin, and C. B. Mullins. Unravelling small-polaron transport in metal oxide photoelectrodes. J. Phys. Chem. Lett. **7**, 471 (2016).

[136] R. von Helmolt, J. Wecker, B. Holzapfel, L. Schultz, and K. Samwer. Giant negative magnetoresistance in perovskitelike $La_{2/3}Ba_{1/3}MnO_x$ ferromagnetic films. Phys. Rev. Lett. **71**, 2331 (1993).

[137] H. A. Jahn and E. Teller. Stability of polyatomic molecules in degenerate electronic states-I-Orbital degeneracy. Proc. R. Soc. Lond. Ser. A **161**, 220 (1937).

[138] M. Puppin, S. Polishchuk, N. Colonna, A. Crepaldi, D. N. Dirin, O. Nazarenko, R. De Gennaro, G. Gatti, S. Roth, T. Barillot, L. Poletto, R. P. Xian, L. Rettig, M. Wolf, R. Ernstorfer, M. V. Kovalenko, N. Marzari, M. Grioni, and M. Chergui. Evidence of large polarons in photoemission band mapping of the perovskite semiconductor $CsPbBr_3$. Phys. Rev. Lett. **124**, 206402 (2020) .

[139] S. I. Pekar. Local quantum states of electrons in an ideal ion crystal. Zh. Eksp. Teor. Fiz. **16**, 341 (1946).

[140] L. D. Landau. Über die Bewegung der Elektronen in Kristallgitter. Phys. Z. Sowjetunion **3**, 664 (1933).

[141] H. Fröhlich, H. Pelzer, and S. Zienau. Properties of slow electrons in polar materials. Philos. Mag. **41**, 221 (1950).

[142] H. Fröhlich. Electrons in lattice fields. Adv. Phys. **3**, 325–361 (1954).

[143] T. Holstein. Studies of polaron motion: Part I. The molecular-crystal model. Ann. Phys. **8**, 325 (1959).

[144] T. Holstein. Studies of polaron motion: Part II. The “small” polaron. Ann. Phys. **8**, 343 (1959).

[145] R. W. Hellwarth and I. Biaggio. Mobility of an electron in a multimode polar lattice. Phys. Rev. B **60**, 299–307 (1999).

[146] J. M. Frost. Calculating polaron mobility in halide perovskites. Phys. Rev. B **96**, 195202 (2017).

[147] R. Rosenfelder and A. Schreiber. On the best quadratic approximation in Feynman's path integral treatment of the polaron. Phys. Lett. A **284**, 63 (2001).

[148] Yukio Ōsaka. Polaron state at a finite temperature. Progr. Theor. Phys. **22**, 437 (1959).

[149] W. Meissner and R. Ochsenfeld. Ein neuer effekt bei eintritt der supraleitfähigkeit. Sci. Nat. **21**, 787 (1933).

[150] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. Theory of superconductivity. Phys. Rev. **108**, 1175 (1957).

[151] J. F. Annett. *Superconductivity, Superfluids and Condensates*. Oxford Master Series in Physics. Oxford University Press, Oxford, 2004.

[152] E. K. Kudinov. Model of a polaron superconductor. Phys. Solid State **44**, 692 (2002).

[153] B. K. Chakraverty. Bipolarons and superconductivity. J. Phys. France **42**, 1351 (1981).

[154] A.W. Leissa. The historical bases of the rayleigh and ritz methods. JSV **287**, 961 (2005).

[155] R. P. Feynman. Forces in molecules. Phys. Rev. **56**, 340 (1939) .

[156] Attila Szabo and Neil S Ostlund. *Modern quantum chemistry: introduction to advanced electronic structure theory*. Courier Corporation, 1996.

[157] H. Hellman. Einführung in die quantenchemie. Franz Deuticke, **285**, 90 (1937).

[158] Michael Spivak. *Calculus*. Cambridge University Press, 2006.

[159] P. Garcia-Fernandez, J.A. Aramburu, M. Moreno, M. Zlatar, and M. Gruden-Pavlović. A practical computational approach to study molecular instability using the Pseudo-Jahn–Teller effect. J. Chem. Theory Comput. **10**, 1824 (2014).

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

[161] P. Pulay. *Ab initio* calculation of force constants and equilibrium geometries in polyatomic molecules. Mol. Phys. **17**, 197 (1969) .

[162] R. M. Sternheimer. Electronic polarizabilities of ions from the Hartree-Fock wave functions. Phys. Rev. **96**, 951 (1954).

[163] X. Gonze, B. Amadon, P.-M. Anglade, J.-M. Beuken, F. Bottin, P. Boulanger, F. Bruneval, D. Caliste, R. Caracas, M. Côté, et al. Recent developments in the ABINIT software package. Comput. Phys. Commun. **205**, 106 (2016).

[164] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, et al. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. J. Phys.: Condens. Matter **21**, 395502 (2009).

[165] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, et al. Gaussian 16, Revision C.01. Gaussian, Inc., Wallingford CT (2016).

[166] F. Neese. The ORCA program system. WIREs Comput. Mol. Sci. **2**, 73 (2012).

[167] D. G. A. Smith, L. A. Burns, A. C. Simmonett, R. M. Parrish, M. C. Schieber, R. Galvelis, P. Kraus, H. Kruse, R. Di Remigio, A. Alenaizan, et al. Psi4 1.4: Open-Source Software for High-Throughput Quantum Chemistry. J. Chem. Phys. **152**, 184108 (2020).

[168] Z. P. Yin, A. Kutepov, and G. Kotliar. Correlation-enhanced electron-phonon coupling: Applications of *gw* and screened hybrid functional to bismuthates, chloronitrides, and other high-T_c superconductors. Phys. Rev. X **3**, 21011 (2013).

[169] W. A. Saidi, S. Poncé, and M. Bartomeu. Temperature dependence of the energy levels of methylammonium lead iodide perovskite from first-principles. JPC letters **7**, 5247 (2016).

[170] R. M Pick, M. H. Cohen, and R. M. Martin. Microscopic theory of force constants in the adiabatic approximation. Phys. Rev. B **1**, 910 (1970).

[171] A.L. Fetter and J.D. Walecka. *Theoretical mechanics of particles and continua*. Courier Corporation, 2003.

[172] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, Th. Frauenheim, S. Suhai, and G. Seifert. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. Phys. Rev. B **58**, 7260 (1998).

[173] T. Fernández-Ruiz, N. Carral-Sainz, J. Íñiguez, J. Junquera, and Pablo Garcia-Fernandez. To be sent to Phys. Rev. B (2025).

[174] A. García, N. Papior, A. Akhtar, E. Artacho, V. Blum, E. Bosoni, P. Brandimarte, M. Brandbyge, J. I. Cerdá, F. Corsetti, R. Cuadrado, V. Dikan, J. Ferrer, J. Gale, P. García-Fernández, V. M. García-Suárez, S. García, G. Huhs, S. Illera, R. Korytár, P. Koval, I. Lebedeva, L. Lin, P. López-Tarifa, S. G. Mayo, S. Mohr, P. Ordejón, A. Postnikov, Y. Pouillon, M. Pruneda, R. Robles, D. Sánchez-Portal, J. M. Soler, R. Ullah, V. W.-Z. Yu, and J. Junquera. SIESTA: Recent developments and applications. J. Chem. Phys. **152**, 204108 (2020).

[175] M. I. Aroyo, J. M. Perez-Mato, D. Orobengoa, E. Tasci, G. De La Flor, and A. Kirov. Crystallography online: Bilbao crystallographic server. Bulg. Chem. Commun. **43**, 183 (2011).

[176] M. I. Aroyo, A. Kirov, C. Capillas, J. M. Perez-Mato, and H. Wondratschek. Bilbao Crystallographic Server. II. Representations of crystallographic point groups and space groups. Acta Crystallogr. A **62**, 115 (2006).

[177] M. E. Lines and A. M. Glass. *Principles and applications of ferroelectrics and related materials*. Oxford university press, 2001.

[178] E.K.H. Salje, M.C. Gallardo, J. Jiménez, F.J. Romero, and J. Del Cerro. The cubic-tetragonal phase transition in strontium titanate: excess specific heat measurements and evidence for a near-tricritical, mean field type transition mechanism. J. Phys.: Condens. Matter **10**, 5535 (1998).

[179] L. Gigli, M. Veit, M. Kotiuga, G. Pizzi, N. Marzari, and M. Ceriotti. Thermodynamics and dielectric response of $BaTiO_3$ by data-driven modeling. Npj Comput. Mater. **8**, 209 (2022).

[180] Ludger Wirtz, Andrea Marini, Myrta Grüning, and Angel Rubio. Excitonic effects in optical absorption and electron-energy loss spectra of hexagonal boron nitride. arXiv preprint cond-mat/0508421 (2005).

[181] B. Kaduk, T. Kowalczyk, and T. Van Voorhis. Constrained Density Functional Theory. Chem. Rev. **112**, 321 (2012).

[182] Y. L. Li, S. Choudhury, J. H. Haeni, M. D. Biegalski, A. Vasudevarao, A. Sharan, H. Z. Ma, J. Levy, Venkatraman Gopalan, S. Trolier-McKinstry, D. G. Schlom, Q. X. Jia, and L. Q. Chen. Phase transitions and domain structures in strained pseudocubic (100) $SrTiO_3$ thin films. Phys. Rev. B **73**, 184112 (2006).

[183] D. Block and A. Wasiela. Self-trapped exciton in alkali fluorides: ODMR study. Solid State Commun. **28**, 455 (1978).

[184] K. S. Song and H. T. Williams. *Self-trapped excitons*. Springer, 1996.

[185] R.T. Williams and K. S. Song. The self-trapped exciton. J. Phys. Chem. Solids **51**, 679 (1990).

[186] D. M. Ceperley and B. J. Alder. Ground state of the electron gas by a stochastic method. Phys. Rev. Lett. **45**, 566 (1980).

[187] J. P. Perdew and Alex Zunger. Self-interaction correction to density-functional approximations for many-electron systems. Phys. Rev. B **23**, 5048 (1981).

[188] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. Phys. Rev. Lett. **77**, 3865 (1996).

[189] J. Junquera, O. Paz, D. Sánchez-Portal, and E. Artacho. Numerical atomic orbitals for linear-scaling calculations. Phys. Rev. B **64**, 235111 (2001).

[190] D. R. Hamann, M. Schlüter, and C. Chiang. Norm-conserving pseudopotentials. Phys. Rev. Lett. **43**, 1494 (1979).

[191] N. Troullier and J. L. Martins. Efficient pseudopotentials for plane-wave calculations. Phys. Rev. B **43**, 1993 (1991).

[192] L. Kleinman and D. M. Bylander. Efficacious form for model pseudopotentials. Phys. Rev. Lett. **48**. 1425 (1982).

[193] F. Sanchez, C. Ocal, and J. Fontcuberta. Tailored surfaces of perovskite oxide substrates for conducted growth of thin films. Chem. Soc. Rev. **43**. 2272 (2014).

[194] M. Kawasaki, A. Ohtomo, T. Arakane, K. Takahashi, M. Yoshimoto, and H. Koinuma. Atomic control of $\mathrm{SrTiO_3}$ surface for perfect epitaxy of perovskite oxides. App. Surf. Sci. **107**, 102 (1996).

[195] S. Das, Y. L. Tang, Z. Hong, M. A. P. Gonçalves, M. R. McCarter, C. Klewe, K. X. Nguyen, F. Gómez-Ortiz, P. Shafer, E. Arenholz, et al. Observation of room-temperature polar skyrmions. Nature **568**, 368 (2019).

[196] A. Ohtomo and H. Y. Hwang. A high-mobility electron gas at the $\mathrm{LaAlO_3/SrTiO_3}$ heterointerface. Nature **427**, 423 (2004).

[197] K. van Benthem, C. Elsässer, and R. H. French. Bulk electronic structure of $\mathrm{SrTiO_3}$: Experiment and theory. J. Appl. Phys. **90**, 6156 (2001).

[198] P. Aguado-Puente, P. Garcia-Fernandez, and J. Junquera. Interplay of couplings between antiferrodistortive, ferroelectric, and strain degrees of freedom in monodomain $\mathrm{PbTiO_3/SrTiO_3}$ superlattices. Phys. Rev. Lett. **107**, 217601 (2011).

[199] Y.J. Chang, A. Bostwick, Y.S. Kim, K. Horn, and E. Rotenberg. Structure and correlation effects in semiconducting $\mathrm{SrTiO_3}$. Phys. Rev. B **81**, 235109 (2010).

[200] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria. Influence of the exchange screening parameter on the performance of screened hybrid functionals. J. Chem. Phys. **125**, 224106 (2006).