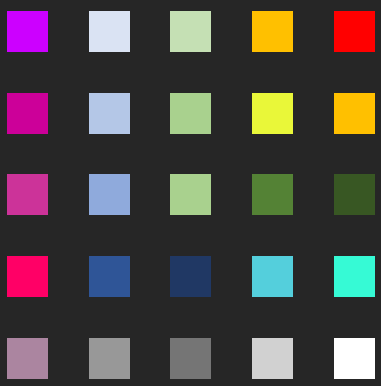
A Journey Through Property Space: DFT and QTAIM Applications in Solid-State Physics

Carl-Friedrich
Schön



A Journey Through Property Space: DFT and QTAIM Applications in Solid-State Physics

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften
der RWTH Aachen University zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von
M.Sc. Carl-Friedrich Schön
aus Siegen

Berichter:
Univ. Prof. Dr. rer. nat. Matthias Wuttig
Univ. Prof. Dr. rer. nat. Stefan Blügel

Tag der mündlichen Prüfung: 18.04.2024

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek verfügbar

Zusammenfassung

Durch die computergestützte Klassifizierung von Materialien anhand ihrer Eigenschaften wird gezeigt, dass der Eigenschaftsraum eine intrinsische Struktur aufweist und dass die Art der chemischen Bindung die Eigenschaften eines Materials definiert. Weiterhin kann durch diesen Ansatz bewiesen werden, dass neben den klassischen Bindungstypen - kovalent, ionisch, metallisch - die sogenannte metavalente Bindung ebenfalls existiert. Zur Kartographierung des Eigenschaftsraum werden die Koordinaten "Electrons Shared" (ES) und "Electrons Transferred" (ET) verwendet. Die aus diesen Parametern resultierende Karte separiert Materialien nach dem Typen ihrer chemischen Bindung und beschreibt Eigenschaftstrends. ES und ET eignen sich daher als Bindungsdeskriptoren und Eigenschaftsprädiktoren. Zudem wurde eine interaktive Version der ES/ET Karte entwickelt, um den Eigenschaftsraum zu visualisieren und so die Einstiegsbarriere zur Verwendung der ES/ET Karte zu verringern. Weiterhin werden die ES/ET Karte in Verbindung mit Dichtefunktionaltheorie (DFT) und "Quantum Theory of Atoms in Molecules" (QTAIM) Rechnungen sowie dem Konzept der metavalenten Bindung eingesetzt, um die charakteristischen Eigenschaften einer Vielzahl von Materialklassen zu untersuchen. Zu diesen Materialien gehören unter anderem Perowskite, Chalkogenide sowie Materialien mit schichtartiger Struktur. Zur Optimierung der DFT und QTAIM Rechnungen wurden Computerprogramme geschrieben, mit welchen ES/ET, "Domain Averaged Fermi Holes", Bandstrukturen, Zustandsdichten sowie dielektrische Funktionen automatisiert berechnet werden können. ES und ET Werte von ca. 4000 Materialien werden dann für maschinelles Lernen eingesetzt, um die Anwendbarkeit dieser Parameter für die Vorhersage von Materialeigenschaften zu demonstrieren.

Abstract

By classifying a database of compounds using an algorithmic approach, an intrinsic structure in property space is implied and chemical bonding is reconfirmed as the fundamental mechanism that determines the properties of solids. In addition to the archetypical types of chemical bonding, covalent, ionic, and metallic bonding, the existence of the recently proposed metavalent bonding is confirmed by the classification results as well. In order to navigate property space, the concepts of "Electrons Shared" (ES) and "Electrons Transferred" (ET) are presented. A map can be drawn that separates materials by their chemical bonding type and exhibits distinct property trends using these parameters. ES and ET can thus be considered bonding descriptors and property predictors that can be used to understand property behavior of materials. To facilitate a low-threshold entry into the ES/ET map concept, an interactive version of the map has been developed. The ES/ET map in conjunction with calculations from density functional theory (DFT) and quantum theory of atoms in molecules (QTAIM) are furthermore utilized to investigate a multitude of material classes and their link to metavalent bonding to explain their characteristic properties. These material classes include inter alia halide and oxide perovskites, chalcogenides as well as layered systems. A supplementary code framework is presented which streamlines the DFT and QTAIM computations to enable high-throughput calculations of ES/ET, domain averaged Fermi holes, band structures, densities of states, and dielectric functions. Roughly 4000 compounds have been calculated in an automatic fashion using this framework and subsequently been employed for a machine learning approach to show the applicability of ES/ET to predict material properties.

Preface

Chemical bonding is a fundamental concept in chemistry that describes the way atoms interact with each other to form molecules. It is the basis for understanding the properties of matter and the behavior of chemical reactions. In material science, chemical bonding is particularly significant because it determines the structure and properties of materials. In material science, the study of chemical bonding is important for the development and optimization of materials for various applications. For example, the chemical bonds in a material can affect its strength, conductivity, and other properties that make it suitable for a particular use. Materials with strong chemical bonds are often used in construction and other applications where strength and durability are important, while materials with weaker chemical bonds are often used in applications where flexibility and malleability are desired. The study of chemical bonding is also important for the development of new materials with novel properties. For example, scientists are currently exploring the potential of using chemical bonds to create superconductors, materials that can conduct electricity with zero resistance at low temperatures. This could revolutionize the way we generate and transmit electricity, as well as have many other applications in fields such as transportation and medicine. Another reason why chemical bonding is a topic of ongoing research is that there are still many mysteries surrounding it. For example, scientists are still trying to understand how chemical bonds form and break under different conditions, and how they can be controlled and manipulated. Additionally, researchers are constantly discovering new types of chemical bonds and exploring their potential applications. In conclusion, chemical bonding is a fundamental and vitally important concept in material science and chemistry. Its significance lies in its role in determining the structure and properties of materials and in the development of new materials with novel properties. It is a topic of ongoing research because there are still many mysteries surrounding it and because of the potential for new discoveries and applications.

-ChatGPT[1], 2022

[1]Used Prompt: *Write a 400 word perspective on chemical bonding with a focus on material science. Include why it is significant and should be a topic of research even in modern physics.*

Author's Note

This thesis is split into three parts. Part I contains the scientific story, describing research conducted to date and what has been achieved. While the interpretation of the methods used and results obtained is motivated and explained, the theoretical background of the conducted DFT (density functional theory) and QTAIM (quantum theory of atoms in molecules) calculations are discussed in Part II. Part III contains a detailed description of how the DFT and QTAIM calculations are performed in practice, providing hands-on examples with which any results of Part I can be recreated. All parts are generally speaking interchangeable and do not need to be read in consecutive order.

This three-way split has been employed to enable readers to tailor this thesis to their individual needs. For a reader proficient in DFT and QTAIM, it should be sufficient to focus on Part I alone, while for scholars from a another field of study the theoretical backgrounds presented in Part II might be of interest as well. Part III is mainly aimed at individuals that want to learn and apply the calculations presented in this thesis themselves. References are provided to link scientific results with their theoretical foundations and practical implementation across the three separate parts.

Contents

Part I

Chapter 1

Introduction - The Discovery of Property Space

Each tool or device is constructed for a certain application and is limited by the properties of the material that it is made from. For every such device or tool, there is an optimal set of material properties that maximizes performance. Take a power line, for example: one would desire a material that is superconducting up to typical earthly temperatures, lightweight, durable, recyclable, and non-toxic as well as cheap to produce. For applications in photovoltaics, a material needs to have large optical absorption, a high charge carrier mobility, small effective masses and a band gap of around 1.1 eV[1]. Optimally, it would even exhibit a layered structure, where the top layer absorbs light with low energy, while being transparent for higher energy photons, subsequently absorbed in the layers below to achieve up to 100% energy conversion. For most applications, defining the required properties for optimal performance is rarely problematic. The challenge is finding a material exhibiting this set of properties.
Calculating the properties of a given material, however, is in principle trivial, as all information is contained in the solution of the Schrödinger equation:

$$H\Psi = E\Psi, \tag{1.1}$$

which only requires the atomic structure of a material as input to calculate the wave function Ψ. Neglecting the feasibility of solving it for systems more complex than a hydrogen atom, the Schrödinger equation provides a concrete mathematical framework to work with. No such approach even exists for the inverse problem, that is stating a desired set of properties and obtaining a material with said properties as the result. In practice, novel, better materials are either obtained by the "Cook'n'Look" approach, i.e. testing variations of known compounds to optimize properties, or by browsing databases for potentially suitable compounds for the desired application. These approaches are heuristic at best, relying on educated guesses and experience and do not promise to yield materials yet completely unknown. Therefore, the question arises as to why there is no methodical process to design materials with a specific property set. Does it not exist, or has it just not yet been discovered? This question is reminiscent of the famous *P versus NP* problem in mathematics and computer science, which poses the question of whether a solution that can "easily" be verified can also be "easily" obtained - the most prominent example being the prime factorization of whole numbers. Finding the prime factors becomes exponentially harder the larger the number to be factorized. However, multiplying prime factors to check their correctness remains (comparatively) trivial, regardless of the size of the number. Transposing this to the problem at hand, the Schrödinger

equation provides a framework to "easily" check the results, i.e. the properties of a compound, but the inverse algorithm has yet to be discovered.
The contents of this chapter are, when not stated otherwise, based on the publications *Classification of properties and their relation to chemical bonding: Essential steps toward the inverse design of functional materials*[2] and *Revisiting the nature of chemical bonding in chalcogenides to explain and design their properties*[3], which were respectively authored and co-authored by the author of this thesis.

1.1 Materials by Design - The Vision

The task of finding an inverse method to tailor property sets can be rephrased to instead ask how we can navigate the property space, what it looks like, what rules it abides by and how it is linked to actual materials. For example, it is unlikely that a material with a large band gap will exhibit a high electrical conductivity. This kind of relation suggests the presence of property correlations. The desired goal would be some sort of navigation algorithm that takes the targeted property set as "destination", and then provides a path through property space, arriving at a material with the desired property portfolio. To evaluate whether such navigation is even possible in principle, a necessary condition is that the property space be intrinsically structured. To investigate whether there is structure in property space, and if so,

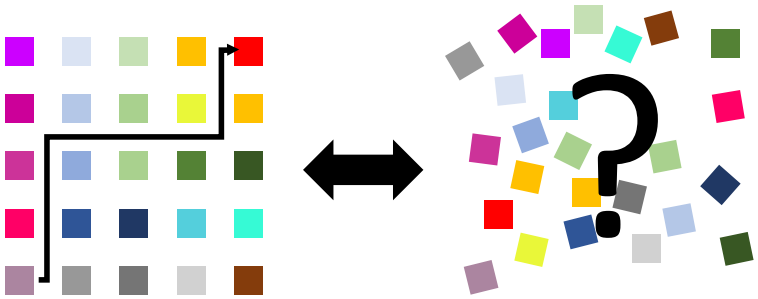


Figure 1.1: Sketch of a structured arrangement (left), that can be navigated, and an unstructured arrangement (right), where navigation is hardly possible.

what it looks like, a database of for the most part relatively simple elemental materials and binary compounds has been compiled. For each of these compounds, the following properties have been derived, either from literature or by means of calculation:

1. The electrical conductivity σ
2. The Born effective charge Z^*
3. The effective coordination number ECoN
4. The band gap E_G
5. The melting point T_M
6. The density ρ
7. The atomic density ρ_A

The Born effective charge Z^* is a measure of the bond polarizability of a compound, while the atomic density ρ_A refers to the number of atoms per unit volume. These seven properties were chosen, as they are easily accessible (comparatively), and are defined for all material classes. While adding more properties promises to offer deeper insight, it would also reduce the number of compounds for which the complete set of properties can be obtained. In sum, about 130 unique compounds could be compiled, where values for each property exist[1]. As the conductivity σ varies by more than 20 orders of magnitude, the logarithm of the conductivity $\log(\sigma)$ is taken into account. The Born effective charge Z^* is furthermore normalized by the formal oxidation state of the respective compound. For a regular ionic compound (like NaCl), the Born effective charge Z^* generally approaches the formal oxidation state. Hence, by defining the excess Born effective charge $Z^*_+ = Z^*/[\text{OxidationState}]$, one can gauge the deviation from the norm ($Z^*_+ \approx 1$). The effective coordination number (ECoN) describes a distance-weighted average over the number of all neighbors of an atomic site (see chapter 7.4). It seeks to provide a numeric measure to assess the structure of a given compound.
To augment the database and increase the total number of datapoints, additional datapoints were created by means of interpolation, motivated by Vegard's law, given that there is a phase diagram indicating that compounds are miscible. This procedure elevates that total number of compounds to about 330 (see tables 11.17 to 11.24 in the appendix).
Using the excess Born effective charge Z^*_+ and the logarithm of the conductivity $\log(\sigma)$ (which will be the standard hereafter, if not stated otherwise), the material database spans a 7D property space consisting of about 330 datapoints. Uncovering correlation and concealed information within structures of data is at the heart of computer science. A data-driven approach from computer science hence promises to yield insights into the structure of the 7D property space. To this end, an instance of the expectation maximization algorithm (EMA)[4] is employed to fit a Gaussian mixture model (GMM) to the datapoints and classify them into distinct clusters[2]. The number of (allowed) clusters to be obtained by the EMA is controlled by varying the number of modes, with the aim being to reveal the emerging structure of the 7D property space step by step. It should be noted that this approach is exclusively evidence-driven, i.e. does not take into account any physical or chemical biases regarding the similarity or dissimilarity of individual properties or compounds. Figure 1.2 shows the EMA results for different numbers of clusters $N = 2, 3, 4$. Four clusters ($N = 4$) can be considered the optimal number of clusters for the set of datapoints used. This can be motivated by the average log likelihood (ALL) metric, shown in figure 1.3. The value of the ALL increases monotonously with increasing numbers of allowed clusters N, which is in line with the general rule that increasing the number of free parameters of a fit improves the (numerical) goodness. However, considering the individual steps going from $N = n \rightarrow n + 1$, it can be observed that the clustering results improve considerably going from $N = 2 \rightarrow 3$ and $N = 3 \rightarrow 4$, while the improvement decreases significantly for the steps $N = 4 \rightarrow 5$ and $N = 5 \rightarrow 6$. This kink at $N = 4$ indicates that 4 is the optimal number of clusters for the analyzed

[1] In this study, only ideal crystal structures are considered, hence the influence of defects, etc., is omitted.
[2] The choice of algorithm and implementation was conducted by Steffen van Bergerem, Christian Mattes, Martin Grohe and Leif Kobbelt[2].

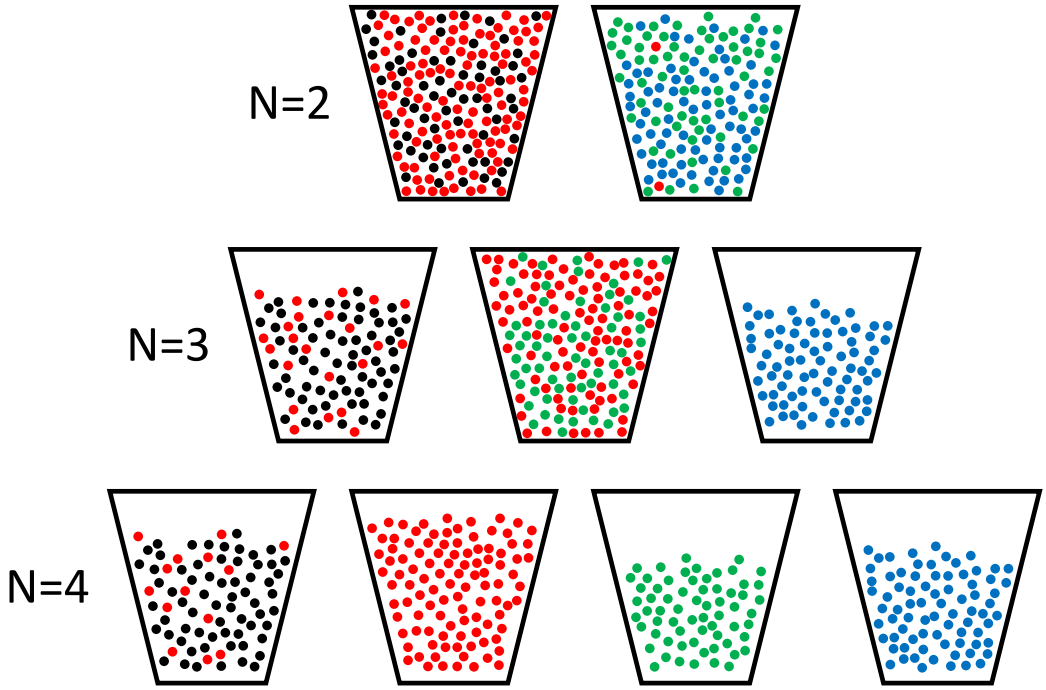


Figure 1.2: Results of the expectation maximization algorithm for different numbers N of clusters. Each compound is represented by a circle. The colors represent the type of chemical bonding within a compound: Covalent bonding (red), ionic bonding (black), metallic bonding (blue), metavalent bonding (green).

dataset. Table 1.1 shows the full list of compounds used, as well as their classification for $N = 4$ clusters. The results of the EMA are only helpful if the common feature (in terms of properties) for each compound within a cluster (that leads to its formation) can be understood. This is indeed possible, as the EMA seems to form clusters corresponding to the chemical bonding type present with the respective compound, as denoted by the color of the circles in figure 1.2 and font color in table 1.1. Red corresponds to covalent bonding, blue to metallic bonding, black to ionic bonding and green to metavalent bonding, a novel bonding type recently proposed[3][5]. While additional bonding types such as hydrogen and van der Waals bonding exist, no such compounds are contained in the database (e.g. ice and solid helium) and are hence not assigned a color. The $N = 4$ row of figure 1.2 indicates that the clusters correspond almost perfectly to the chemical bonds present within the compounds. While there are no misclassifications for the metavalent (green) and metallic (blue) compounds, 14 covalent (red) compounds are assigned to the cluster corresponding to the ionic (black) compounds. Looking at $N = 2$ clusters, covalent and ionic compounds as well as metavalent and metallic compounds are lotted together. This changes for $N = 3$ clusters, where metals and (almost) all ionic compounds are split off into their own clusters, while the metavalent compounds then form a cluster with the covalent compounds. This switching of clusters going from $N = 2 \rightarrow 3$ indicates that the metavalent compounds are about equally different from covalent and metallic

[3]The assignment of the "true" bonding types (represented by the colors) is either obtained from the literature directly, or inferred from typical characteristic property ranges (see table 1.2).

Table 1.1: Results of the EMA clustering in comparison to the expert classification. The rows indicate a cluster found by the EMA, named after the most common bonding type within them. The font color of a compound denotes the expert classification. The stable phases of all compounds are used, while metastable phases are indicated in brackets. Alloys/blends of compounds are denoted with colons. Full tables along with properties are included in the appendix (tables 11.17 to 11.24). Adapted from [2].

Cluster 1 (Ionic)	$AgSbS_2$, AlN, AlP, BaO, BaS, BaSe, BaTe, BeO, BeS, BeSe, BeTe, BN, C-Diamond, CaO, CaS, CaSe, CaTe, CsBr, CsCl, CsF, CsF ($Pm\bar{3}m$), CsI, GaN, GeSe, HgS ($F4\bar{3}m$), KBr, KCl, KF, KI, MgO, MgS, MgSe, MgTe, NaCl, NaBr, NaF, NiO, PbO, RbBr, RbBr ($Pm\bar{3}m$), RbCl, RbI, Sb_2S_3, Sb_2Se_3, SnO, SnS, SrO, SrS, SrSe, SrTe, ZnO, ZnS, ZnSe, RbBr:RbI, NaBr:NaCl, NaCl:KCl, CaS:CaSe, CsCl:CsBr, CsBr:CsI
Cluster 2 (Covalent)	AlBi, AlSb, CdS, CdSe, CdTe, GaP, GaAs, GaSb, GaSe, Ge, HgSe, HgTe, InAs, InP, InSb, Si, SnSe ($Pnma$), ZnTe, CdS:CdSe, CdTe:HgTe, HgSe:HgTe, InP:GaP, InAs:GaAs, InP:InAs, GaP:GaAs, InP:GaP:InAs:GaAs, InP:GaP:InSb:GaSb, InAs:GaAs:InSb:GaSb
Cluster 3 (Metavalent)	$AgSbSe_2$, $AgBiSe_2$, $AgBiTe_2$, $AgSbTe_2$, As_2Te_3 ($R\bar{3}m$), Bi_2Se_3, Bi_2Te_3, GeTe, PbS, PbSe, PbTe, Sb_2Te_3, SnTe, PbS: PbSe, PbTe:PbSe, PbTe:SnTe, GeTe:SnTe, Sb_2Te_3:Bi_2Te_3, Bi_2Se_3:Bi_2Te_3, PbTe:$AgSbTe_2$, GeTe:$AgSbTe_2$, $AgSbTe_2$:$AgBiTe_2$, GeTe:Sb_2Te_3, SnTe:Sb_2Te_3, PbTe:Sb_2Te_3, PbTe:Bi_2Te_3
Cluster 4 (Metallic)	Ag, $AgSnSe_2$, $AgSnTe_2$, Al, Au, Ca, Ca ($Im\bar{3}m$), Cu, CuAu, Hf, K, Li, Mg, Na, Nb, Ni, NiPt, Pb, Pd, Sc, Sr, Ta, Y, Zr, AlAu, AlCu, AlPd, AlPt, Co, Cu, CuZr, Fe, Ga, GaPd, GaPt, In, In_3SbTe_2, Ir, La, Mn, Mo, NbC, NbN, Ni, NiAl, Pt, Re, Rh, TiO, Zn, Ni:Pd, Ni:Cu, Sr:Ca, Hf:Zr, Au:Co, Co:Ni, Ta:Nb

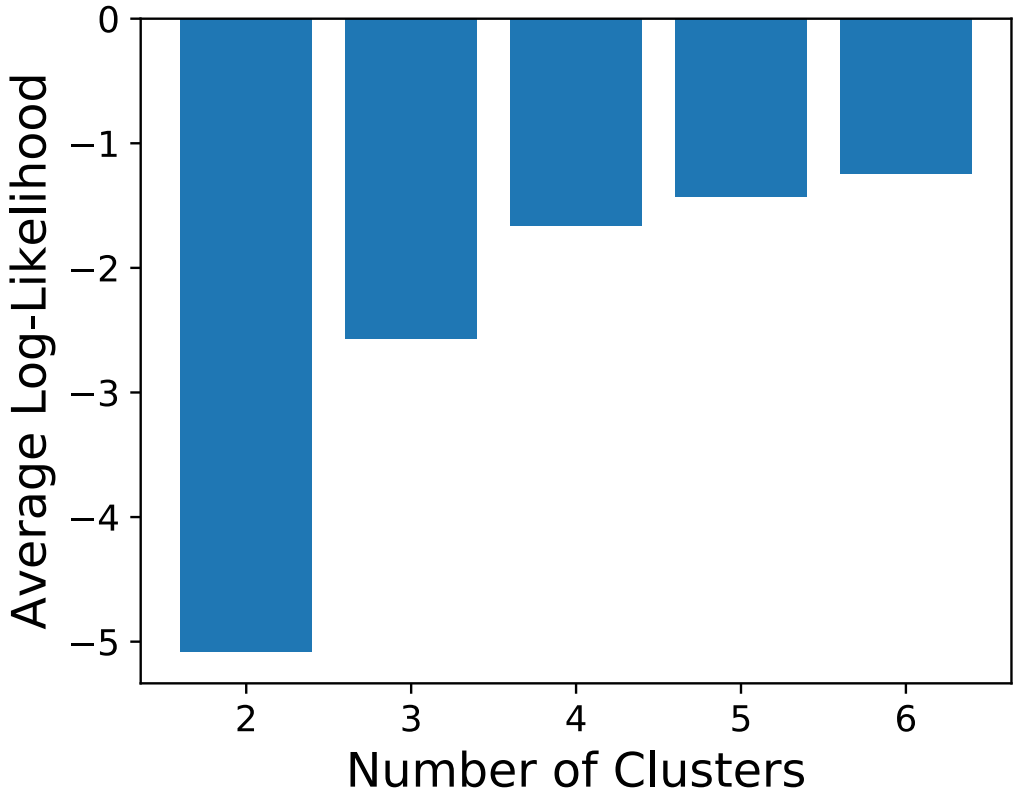


Figure 1.3: The average log likelihood (ALL) metric used to evaluate the goodness of the EMA clustering for different numbers of cluster $N = 2, 3, 4, 5, 6$ (the higher the better). The ALL increases monotonously with increasing N, as adding more parameters is bound to improve the quality of a fit. However, it can be observed that the ALL increases considerably going from $N = 2 \rightarrow 3$ and $N = 3 \rightarrow 4$, while the improvement decreases significantly for the steps $N = 4 \rightarrow 5$ and $N = 5 \rightarrow 6$. This "elbow" analysis indicates that $N = 4$ is the optimal number of clusters.

bonding, and putting them in the covalent or metallic cluster is almost equally incorrect. It should further be noted that before the proposition of metavalent bonding, the respective compounds were considered to be covalently bonded. The clustering with $N = 3$ clusters therefore reproduces this perspective, while the results for $N = 4$ clearly suggest the existence of metavalent bonding, underlined by the ALL metric shown in figure 1.3.
Based on the data presented, the concept of chemical bonding appears to play a crucial role within the property space. The next section therefore explores this concept in more detail, with the focus being placed on metavalent bonding, as this is a bonding type discovered much later than covalent, ionic and metallic bonding.

1.2 Chemical Bonding - The Foundation

One of the most influential works regarding chemical bonding is Linus Pauling's book *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry*, published in 1960[6]. Pauling states that it is "convenient to consider three general extreme types of chemical bonds: *electrostatic bonds*, *covalent bonds* and *metallic bonds*"[6]. In modern terminology electrostatic bonds within solids is generally referred to as ionic bonding. Ionic, covalent, and metallic bonding are prototypical corner cases which (usually) do not manifest in nature in pure form. In most cases, a mixture of bonds occurs, with at least slight contributions from the non-dominant bonding types.

Ionic Bonding

Ionic bonding originates from electrostatic forces attracting oppositely charged ions. The commonly used prime example for this bonding type is NaCl, i.e. sodium chloride. The metallic element (Na) is prone to removing the electron in its outermost shell, thereby emptying it, while the non-metallic element (Cl) adds it to its outer-most shell, filling it up completely. This results in a positively charged cation (Na^+) and a negatively charged anion (Cl^-), which then exert electrostatic forces on their respective atomic environments, stabilizing the crystal structure[6].

Covalent Bonding

While ionic bonding relies on the transfer of electrons, covalent bonding is based on the concept of sharing electron pairs between two nuclei in order to lower the total energy and achieve a stable crystal structure. The covalent bond generally occurs between nuclei with similar affinity to attract electrons (as quantified by the electronegativity), which is best fulfilled for identical constituents, e.g. two carbon atoms forming a bond: C−C. Sharing a full pair of electrons matches the description of a single covalent bond in the Lewis picture (bond order BO = 1). Somewhat similar to the ionic bond, each atom aims to fill the outer shell (octet rule). However, in the case of covalent bonding, this is through sharing electrons instead of transferring them[7]. Covalent bonds are furthermore linked to the overlap of (atomic) orbitals of the constituents, e.g. by forming σ or π bonds (see figure 1.4).

Metallic Bonding

Metallic bonding is characterized by delocalized valence electrons and can be considered as the sharing of free electrons within an arrangement of cations (positively charged ions). Considering the energetic benefits of this configuration, it can be argued that the delocalization

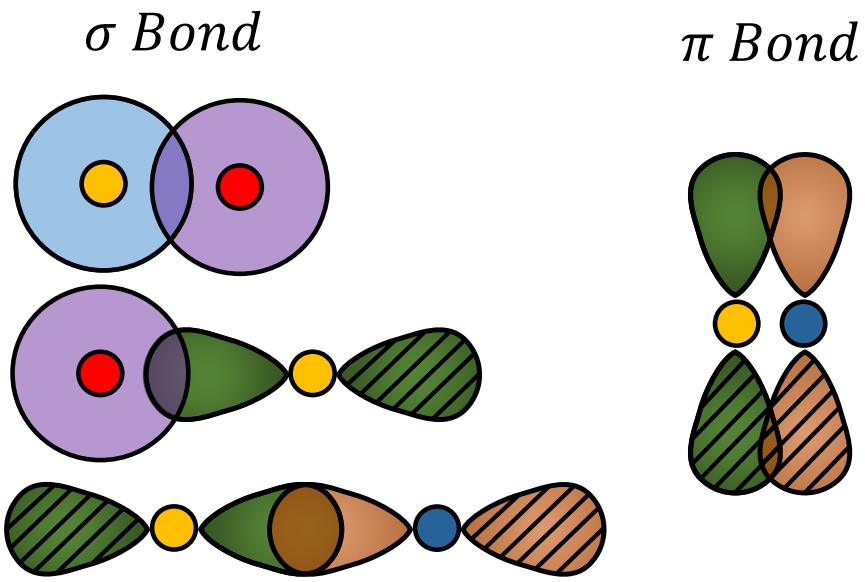


Figure 1.4: Sketch of s and p orbital overlaps in covalent bonding.

of electrons results in a reduction of kinetic energy. This can be motivated by Heisenberg's uncertainty principle:

$$\sigma_x \cdot \sigma_p \geq \frac{\hbar}{2}, \tag{1.2}$$

with σ_x and σ_p being the standard deviation of position x and momentum p, respectively. The larger the delocalization and consequently σ_x is, the smaller σ_p can be and hence the average momentum p (and thus velocity) of an electron, lowering the total kinetic energy $E_{kin.} \propto p^2$.

However, the virial theorem as valid for electrostatic repulsion, states[8]:

$$\langle E \rangle = \frac{\langle V \rangle}{2} = -\langle T \rangle, \tag{1.3}$$

where $\langle E \rangle$, $\langle V \rangle$ and $\langle T \rangle$ are the average total, potential and kinetic energies respectively. This implies that the potential energy $\langle V \rangle$ is responsible for bond formation, as it decreases during bond formation, while the kinetic energy $\langle T \rangle$ increases twice as much. It can be argued that the virial theorem in this form is only valid for the initial (unbonded) and final (bonded) state, and that during bond formation both kinetic and potential energy play a crucial role[4]. Which energy contribution is ultimately the decisive factor for bond formation is still an ongoing debate among quantum chemists[9].

[4] A prominent example being the H_2 molecule, for which the atomic distance is about $d_{H-H} \approx 0.74\,\text{Å} < 2a_0$, where $a_0 \approx 0.53\,\text{Å}$ is the Bohr radius, the radius of an isolated hydrogen atom. While the total energy $\langle E \rangle$ is minimized at d_{H-H}, neither the potential $\langle V \rangle$ nor the kinetic energy term $\langle T \rangle$ are.

Metavalent Bonding

As already mentioned before, metavalent bonding (MVB) is a new bonding type that has recently been proposed[5]. Its name originates from the Greek word *μετά* (meta), meaning *transcending* as a prefix. It additionally is a play on the combination of *metallic* and *covalent*, indicating a bonding type that is related to covalent and metallic bonding, but transcends these, hence constituting something new. Metavalently bonded compounds are also often referred to as *"incipient metals"*, as they are characterized by a half-filled conduction band, which would generally result in metallic behavior. However, due to symmetry breaking of the structure (by slight distortions and/or charge transfer) a small band gap opens up, rendering them semi-conductors[10]. A prominent example is cubic (octahedral) GeTe, as depicted in figure 1.5, where the average number of valence electrons per atomic site is three; hence only half an electron pair (i.e. one electron) is shared between the six bonding partners (as opposed to the full electron pair, i.e. two electrons, in covalent compounds). Materials employing MVB can thus be thought of as being in the competition zone between electron localization and delocalization.

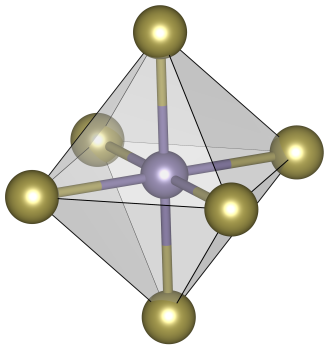


Figure 1.5: Depiction of cubic GeTe (octahedral configuration). Each atomic site forms bonds with six neighbors, sharing one electron per bond.

Still, visualizing the different types of bonding would be insightful, as one could argue that metallic, covalent, and ionic bonding (and metavalent bonding for that matter) are merely textbook limiting cases without any significance in terms of real-world application. The domain averaged Fermi hole (DAFH) approach can be utilized to visualize to the orbitals taking part in bonding (see chapters 6.2 and 7.2.3), as shown for exemplary systems in figure 1.6. A clear difference in the shape for the different bonding types can be observed. Aluminum employs metallic bonding, which is dominated by the outermost 3s orbital. It is quite delocalized, extending (with significant contribution) up to the second coordination shell. The covalent bond of C (diamond) in contrast is localized between the atomic sites partaking in bonding. Each atom is providing about one electron to each bond, forming an electron pair as expected for an archetypical covalent bond. For ionic bonding as exemplified by NaCl, the orbitals show very little overlap and resemble the atomic orbitals of isolated elements. The outermost orbitals of Na and Cl are almost fully occupied (2.0 e^-/1.9 e^-), as the charge is transferred from the Na 3s states to the Cl 3p states. Metavalent bonding shows directed orbitals. However, the orbitals are more delocalized than for covalent bonding, but not as delocalized as in metallic bonding. Ge contributes about 0.5 e^- and Te 1.3 e^-. This total amount of about 1.8 e^- is shared amongst two bonding partners and the central site, each bond thus consisting of only 0.9 e^- or half an electron pair each, forming a 3 center - 2 electron (3c - 2e) arrangement. This configuration differs distinctly from what is found in metallic of covalent bonding (and ionic bonding, respectively).

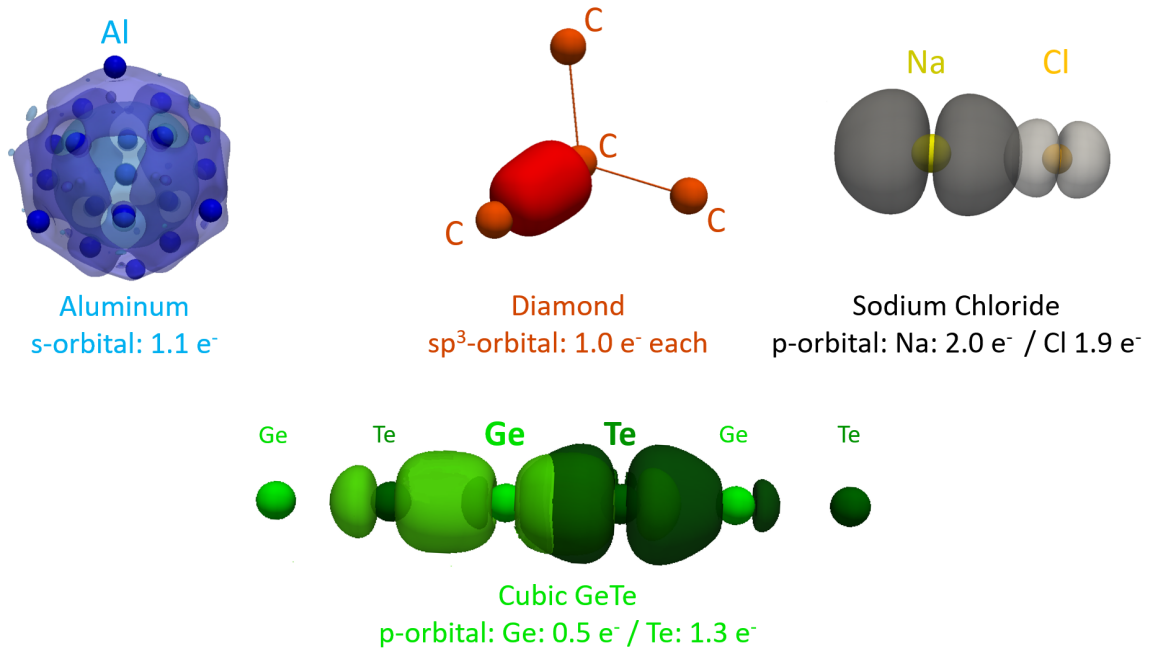


Figure 1.6: Domain averaged Fermi hole (DAFH) orbital visualizations (see chapters 6.2 and 7.2.3) for Al, C (diamond), NaCl and cubic GeTe. Aluminum employs metallic bonding, which is dominated by the outermost 3s orbital. It is quite delocalized, extending (with significant contribution) up to the second coordination shell. The covalent bond of C (diamond) in contrast is localized between the atomic sites partaking in bonding. For ionic bonding as exemplified by NaCl, the orbitals show very little overlap and resemble the atomic orbitals of isolated elements. Metavalent bonding shows directed orbitals, however more delocalized than for covalent bonding, but not as delocalized as in metallic bonding. The occupation of the respective orbitals is stated below each figure. Isosurfaces are set to: Al: 0.005, C (diamond): 0.2, NaCl: 0.03, GeTe: 0.03. The DAFH occupations and overlaps for Al, C and NaCl are shown in tables 11.7, 11.8 and 11.9 in the supplement, while the values for cubic GeTe are shown in table 2.5. Taken from [3].

Figure 1.7 shows comparable information in regard to figure 1.6, using identical values, however, for the orbital isosurfaces: $|0.025|$, in order to illustrate the different sizes and spatial extent of the orbitals in a more comparable manner.
Instead of diving deeper into the underlying mechanisms behind chemical bonding, it might provide more insight in the given context to change the perspective once again and look at chemical bonding from the point of view of properties. Table 1.2 shows the characteristic ranges for a subset of (characteristic) properties.

This set of properties has been used to help attribute the "correct" bonding types to the database of materials (expert classification), in accordance with what is generally reported in literature for the respective compounds, and was used to determine the colors of the compounds in figure 1.2 and table 1.1:

1. **Electrical Conductivity σ at Room Temperature**: This property is a measure of the electronic structure of a given compound, and is especially useful in separating metals, as they exhibit high room temperature conductivities of $\sigma > 1 \times 10^5\,\mathrm{S\,cm^{-1}}$.

2. **Number of Nearest Neighbors**: In order to gauge the structure of a material, the number of nearest neighbors is assessed. This property is most useful in detecting covalent bonding, as the number of nearest neighbors usually follows the $8 - N$ rule, where N is the number of valence electrons[6, 11].
3. **Dielectric Constant ϵ_∞**: This property describes the optical response of a material in the limit of high frequencies. It tends to be high for compounds employing MVB.
4. **Born Effective Charge Z^***: As already stated in the previous paragraphs, the Born effective charge Z^* is a measure of the bond polarizability and high values are characteristic of metavalently bonded materials.
5. **Grüneisen Parameter γ_{TO}**: The Grüneisen parameter for transversal optical (TO) modes is employed to evaluate the lattice anharmonicity of a given solid.

The notion of MVB being a unique, separate bonding type is further corroborated by comparing characteristic ranges of the property sets for the bonding types as provided in table 1.2. The property set of MVB is not a linear combination of any other bonding type (neither are the sets of the other types of chemical bonding), which would be expected if it was merely an intermediate state between other bonding types.

Table 1.2: Typical property ranges for ionic, covalent, metavalent and metallic bonding. No set of properties can be expressed as a linear combination of the other property sets, providing strong evidence that each bonding type (especially MVB) is indeed unique and not an intermediate state or mixture of the other bonding types. Adapted from [2].

Property	*Ionic*	*Covalent*	*Metavalent*	*Metallic*
Electrical Conductivity σ	Very Low	Low-Moderate	Moderate	High
Number of Nearest Neighbors	4, 6 or 8	8-N Rule Satisfied	8-N Rule Unsatisfied	8 or 12
Optical Dielectric Constant ϵ_∞	Low (2-3)	Moderate (5-15)	High (>15)	-
Born Effective Charge Z^*	Low (1-2)	Moderate (2-3)	High (4-6)	Vanishes (0)
Grüneisen Parameter γ_{TO}	Moderate (2-3)	Low (0-2)	High (>3)	Low (0-2)

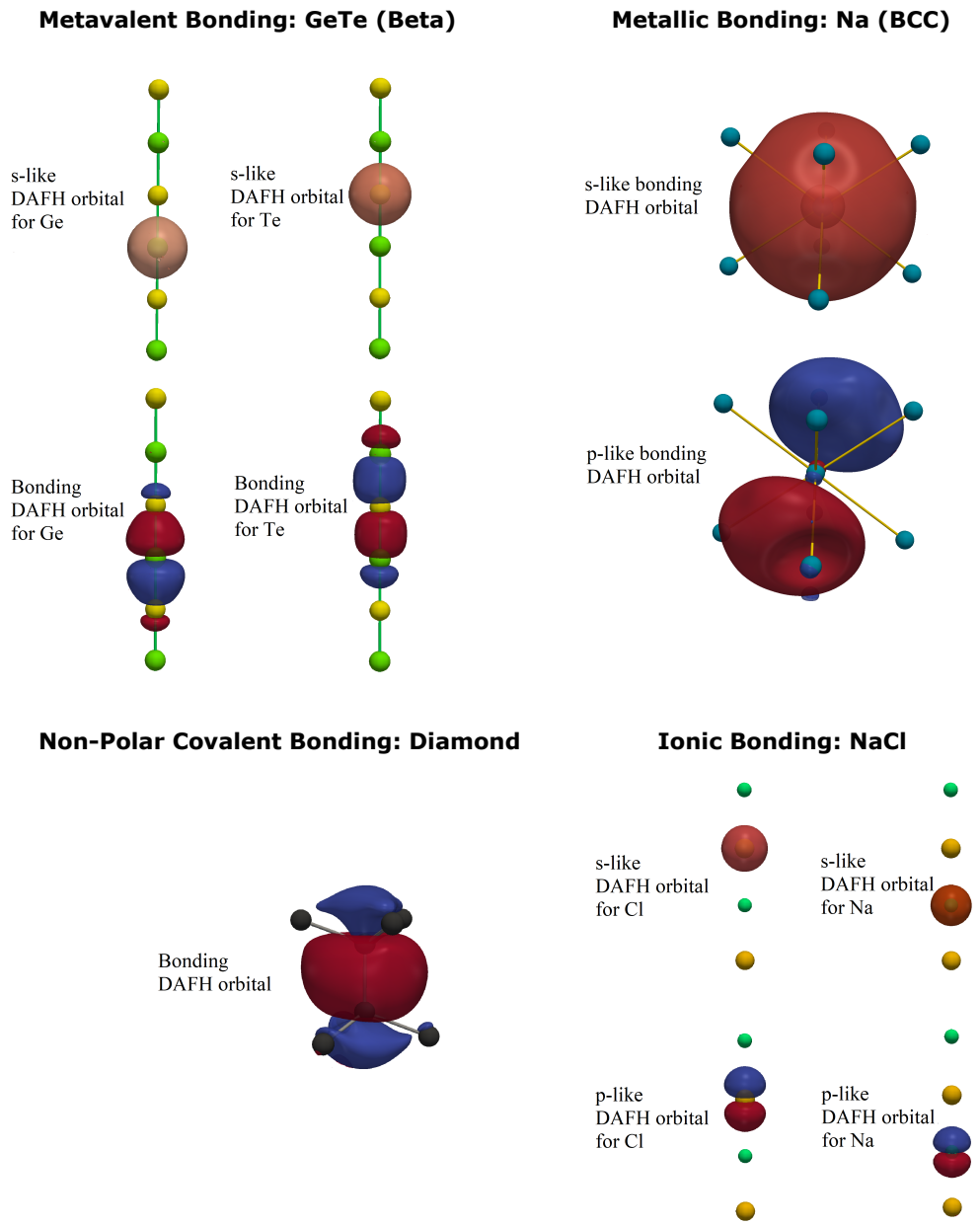


Figure 1.7: Domain averaged Fermi holes (DAFH) orbital visualizations (see chapters 6.2 and 7.2.3) for cubic GeTe, Na, C (Diamond) and NaCl. Only valence shells are shown: C - (3s/3p), Al - (3s, 3p), Na - (3s, 3p), Cl - (3s, 3p), Ge- (4s, 4p), Te - (5s, 5p). As the diamond orbitals hybridize (sp^3), only one orbital is visualized. In contrast to figure 1.6, isosurfaces are set to $|0.025|$ for all systems to illustrate the different sizes and spatial extent of the orbitals in a more comparable manner. Taken from [3], created by Pavlo Golub.

While some properties like the Born effective charge Z^* coincide with the properties used for the EMA, it also includes other properties like the Grüneisen parameter γ_{TO}. This disparity origins from the fact that not all properties of table 1.2 are readily available for many compounds, or are not even defined for all bonding types (e.g. the dielectric constant ϵ_∞ is undefined for metals). While the EMA requires a complete set of properties for each compound, the manual "expert" classification conducted by humans can work with incomplete property sets. Using table 1.2, it is relatively straight forward to manually assign a bonding type to a given compound. The fundamental difference between the two approaches should be stressed at this point: The numeric approach of the EMA is by design an a priori and hence unbiased classification of compounds. The "expert" classification is also mainly based on properties, but relies on chemical background knowledge as well. While the EMA would be free to consider Aluminum to be ionic, no property anomalies would convince the "expert" to label it anything other than a metal, as he *knows* it to be one.
That fact that chemical bonding is often considered a somewhat "soft" descriptor makes the agreement between the EMA and the "expert" classification even more notable.

1.3 Properties - The Landscape

The previous chapter illustrated that in property space there must be structure that is closely linked to chemical bonding as an elemental building block. In this chapter, the correlation of various properties with each other will be investigated in more detail in order to better understand the topography of property space.
For that matter, looking at the property ranges of the compounds used for classification should offer some insight. Table 1.3 shows the minimum and maximum value of each property used for the EMA, partitioned into the different bonding types as assigned by the expert classification. Additionally, figure 1.8 shows the distribution of properties within each bonding type (as assigned by the expert classification).

Table 1.3: Minimum and maximum value of each of the properties used for the classification algorithm. The compounds are separated by their respective bonding types, as assigned by the expert classification. Most of the property ranges overlap among the different bonding types, implying multiple properties are required to assign a compound. Adapted from [2].

Property	Ionic		Covalent		Metavalent		Metallic	
	Min	*Max*	*Min*	*Max*	*Min*	*Max*	*Min*	*Max*
Conductivity Log(σ)	-22.2	-4.2	-14.0	3.0	1.4	4.0	3.3	5.8
Elev. Born Eff. Charge Z^*_+	0.6	1.4	0.0	1.7	1.7	3.3	0.0	0.0
ECoN	4.0	12.0	3.1	6.0	4.8	6.0	6.0	12
Band Gap E_G (eV)	1.6	10.6	>0	5.5	0.1	0.7	0.0	0.0
Melting Point T_M (K)	853	3643	794	4100	828	1387	301	3873
Density ρ (g/cm^3)	1.9	5.8	2.3	8.6	5.9	7.9	0.6	21.9
Atomic Density ρ_A (10^{-2}/Å^3)	1.7	16.8	2.7	17.5	2.8	4.1	1.4	10.1

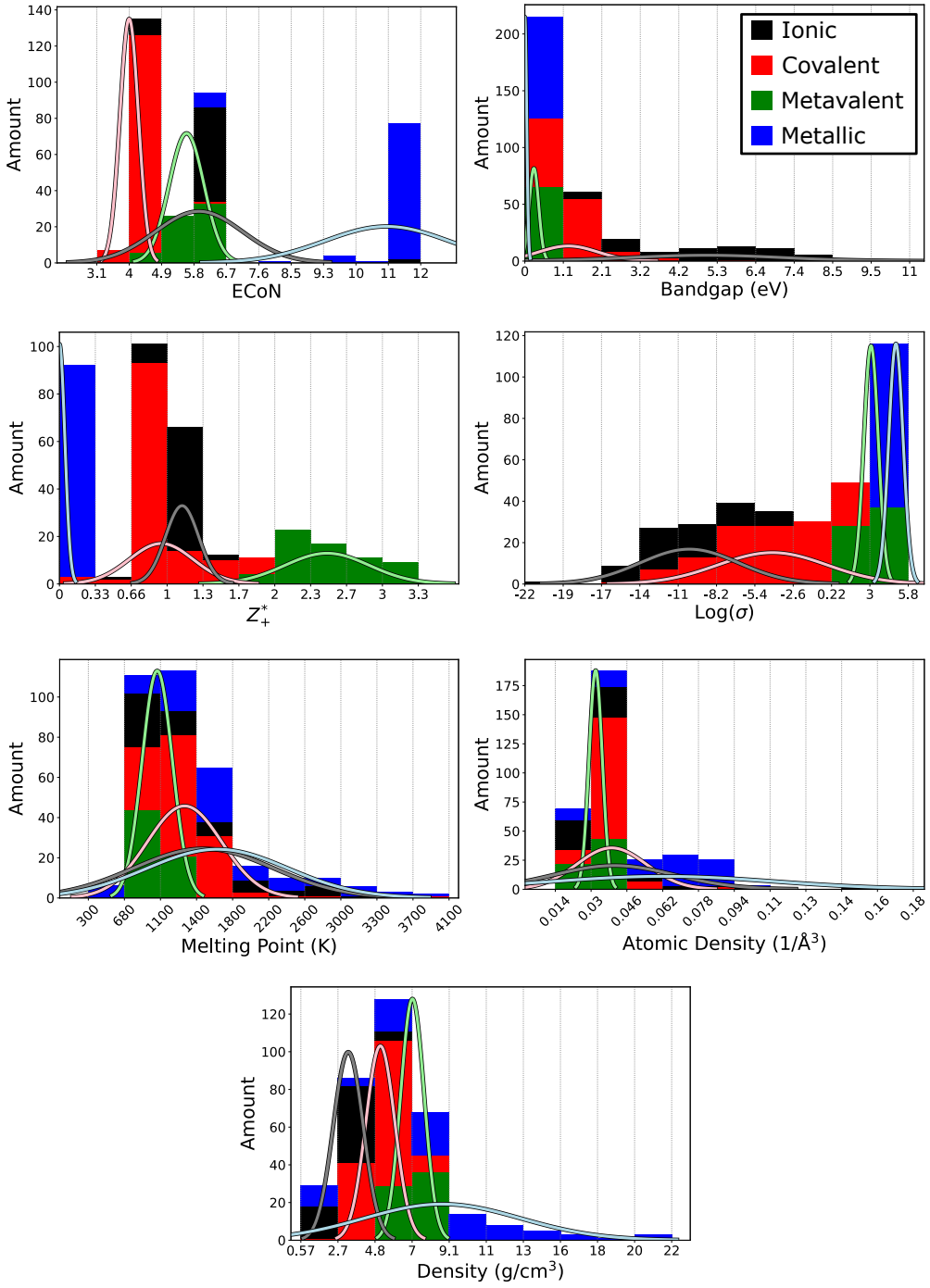


Figure 1.8: The property distribution within each bonding type as assigned by the expert classification. The respective property ranges are subdivided into 10 equally large bins each. The minimum and maximum values of each bin are denoted on the x-axis. The height of the bar in each bin is then obtained by counting the number of compounds that exhibit property values in the corresponding bin range. The bar is colored according to the ratio of bonding types of all compounds within the respective bin. A Gaussian distribution is furthermore fitted to the data and plotted alongside the bins, showing that no property alone is sufficient to infer the bonding type due to (significant) overlap. However, some properties exhibit more narrow ranges for specific bonding types, e.g. the conductivity $\log(\sigma)$ for metals and MVB compounds. Adapted from [2].

Table 1.3 and figure 1.8 in conjunction show that no property can be used by itself to determine the underlying bonding type, as the distributions/ranges do overlap. This holds true for all properties, apart from the elevated Born effective charge Z^*_+ and the band gap E_G, which are always and exclusively zero for metallic compounds. Hence, metals can be identified relatively easily, while for the other bonding types the combination of individual properties must be considered to discriminate between bonding types. It also becomes evident that some properties seem to contain more discriminatory power than others[5]:

1. The **melting point T_M**, the **atomic density ρ_A** and the **mass density ρ** do not appear to be of significant use in the separation into bonding types, as the ionic, covalent and metallic distributions cover wide ranges with significant overlap. The atomic density ρ_A, for example, ranges from $0.0017 - 0.0168\,\text{Å}^{-3}$ for ionic compounds, from $0.0027 - 0.0175\,\text{Å}^{-3}$ for covalent compounds and from $0.0014 - 0.0101\,\text{Å}^{-3}$ for metals. A comparable (relative) spread is, however, also observed for the melting point T_M and the mass density ρ. Only the metavalent compounds are restricted to a by comparison narrower window (see table 1.3).

2. The **logarithmic electrical conductivity $\log(\sigma)$** is a quantity that seems to separate metavalent ($\log(\sigma) = 1.4 - 4.0$) and metallic compounds ($\log(\sigma) = 3.3 - 5.8$) reasonably well, as the distributions form relatively narrow windows for these two bonding types. For ionic and covalent compounds, however, the predictive power is much weaker, with wider $\log(\sigma)$ ranges, respectively.

3. The **elevated Born effective charge Z^*_+** is especially useful to detect metals, as the delocalized electrons screen the dipoles created by lattice vibrations, yielding $Z^*_+ = 0$. This is different in the metavalent compounds, which exhibit the highest values of Z^*_+, ranging from $1.7 - 3.3$, with almost no overlap with the covalent and ionic compounds. Among the ionic and covalent compounds however, the discriminatory power is again much weaker, indicating a potential origin of the difficulties the EMA showed in separating covalent and ionic compounds/properties.

4. The **band gap E_G** seems to underline this notion even further, as ionic and covalent bonding both can produce relatively large band gaps E_G of up to 5.5 eV. Values above this threshold are exclusively reserved for ionic compounds, while values below 1.6 eV only occur for covalent and metavalent solids. Overlap between covalent and metavalent bonding does exist. However, MVB compounds exhibit a narrower range of roughly between $0.1 - 0.7$ eV. Regarding metals, the band gap E_G is (similar to the elevated Born effective charge Z^*_+) an excellent predictor, as $E_\mathrm{G} = 0$ holds true for all and only for metals.

5. The **effective coordination number ECoN** appears to be of moderate discriminatory power, as overlap is present for all bonding types. Nevertheless, the range for covalent compounds is relatively narrow, while it does exhibit overlap with ionic and metavalent

[5] At least by human assessment. Machine learning approaches might come to a different conclusion.

solids. While the ECoN range for metallic compounds is relatively wide ($6.0 - 12.0$), a tendency towards large ECoN values of $11.0 - 12.0$ causes the overlap with the other bonding types to be relatively small.

A classification based on an individual property is only possible for metals, while it is not feasible for covalent, ionic or MVB compounds. Hence the characteristics of the complete property set for a given compound are the decisive aspect enabling classification as performed by the EMA and human experts, as already mentioned.
To further examine the structure of the property space, it is essential to consider correlations among properties, within and across bonding types. Correlation plots for all possible combinations of properties are shown in figure 1.9. While not all property pairs are significantly correlated, the ones that do are:

1. **Electrical Conductivity $\log(\sigma)$ - Band Gap E_G**: For metals, the band gap is zero $E_G = 0\,\mathrm{eV}$ regardless of the value for electrical conductivity $\log(\sigma)$, as expected. For metavalent compounds, the electrical conductivity $\log(\sigma)$ decreases with increasing band gap E_G. The same relation holds true for covalent and ionic compounds, for which the conductivities decrease even further. Also, two distinct lines with different slopes are present for ionic and covalent compounds, on which almost all (ionic and covalent) compounds are located, with the slope of the ionic line being much steeper. Ionic compounds hence exhibit higher band gaps E_G as covalent compounds for the same conductivity $\log(\sigma)$.

2. **Electrical Conductivity $\log(\sigma)$ - Elevated Born Effective Charge Z_+^***: Metals are located within a range of $\log(\sigma) = 3.3 - 5.8$, for which the elevated Born effective charge Z_+^* is zero by definition. Adjacent in conductivity $\log(\sigma)$ to the metallic compounds are the metavalent compounds, which exhibit the highest values of Z_+^* of all bonding types in general. Interestingly, the highest values of Z_+^* within the metavalent bonds are close to the border to the metallic compounds (in terms of electrical conductivity $\log(\sigma)$). This discontinuity implies that fundamental mechanisms change at the transition between metavalent and metallic compounds, further highlighting the distinctiveness of metavalent and metallic bonding. For covalent and ionic compounds however, no clear trend can be observed, as the values of the elevated Born effective charge are mostly constant, with values around $Z_+^* = 1$ for both bonding types. (The values for covalent compounds are generally slightly below $Z_+^* = 1$, while ionic compounds feature slightly higher values.)

3. **Electrical Conductivity $\log(\sigma)$ - Melting Temperature T_M**: A linear trend can be observed for covalent compounds, where the melting temperature T_M increases with decreasing conductivity $\log(\sigma)$. The ionic compounds, up to this point behaving similar to covalent compounds, do not exhibit the same trend, as no distinct correlation seems to be present. Similarly, for metallic and metavalent compounds, the melting temperature T_M seems to be rather uncorrelated to the electrical conductivity $\log(\sigma)$.

4. **Electrical Conductivity $\log(\sigma)$ - Effective Coordination Number ECoN**: For all bonding types, no clear correlation between $\log(\sigma)$ and ECoN can be observed. However, metals and metavalent compounds are characterized by high values of $\log(\sigma)$ and ECoN respectively, with both properties being slightly lower for compounds utilizing MVB.

5. **Band Gap E_G - Elevated Born Effective Charge Z^*_+**: Again, no clear trend for ionic and covalent compounds is visible. For metals, both properties are zero by definition. For materials employing metavalent bonding, the elevated Born effective charge Z^*_+ tends to decrease for increasing band gap E_G.

For most other property combinations, no distinct correlation can be inferred directly. Especially properties that are not particularly characteristic of any bonding type, i.e. melting temperature T_M, atomic density ρ_A and mass density ρ, show only very weak correlations, if at all. As listed above, the electrical conductivity at room-temperature $\log(\sigma)$, the band gap E_G and the elevated Born effective charge Z^*_+ generally offer greater insight.

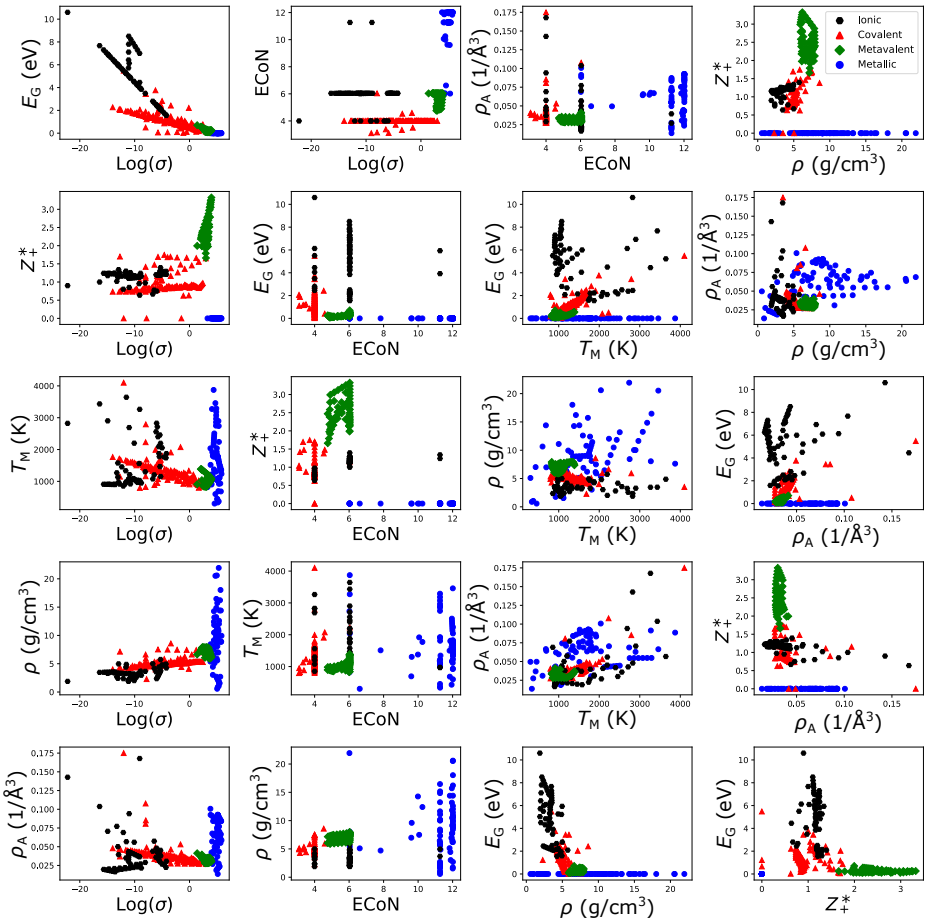


Figure 1.9: Correlation plots across all properties used for classification. The colors indicate the bonding types (as defined by the expert classification): Black - Ionic, Red - Covalent, Blue - Metallic, Green Metavalent. Taken from [2].

1.4 Abstraction of Property Space - The Cartography

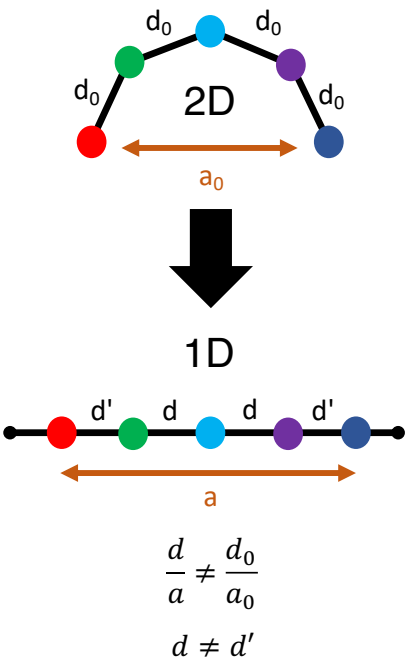


Figure 1.10: Depiction of the working principle of the *t*-distributed stochastic neighbor embedding (t-SNE) in a 2D $\rightarrow$ 1D model example. Datapoints are originally distributed in 2D (top part), while the distance between each neighboring point is d_0, and the distance between the leftmost red datapoint and the rightmost dark blue datapoint amounts to a_0. The t-SNE reduces the dimensionality of this distribution to 1D (bottom part) by trying to keep the relative distances between all datapoints equal, i.e. amongst others $\frac{d_0}{a_0} \stackrel{!}{=} \frac{d}{a} \stackrel{!}{=} \frac{d'}{a}$. As this is generally impossible to achieve, the t-SNE minimizes the deviation from this optimum by distributing the datapoints in the lower dimensional space accordingly (joint probabilities are used to assess the similarity of datapoints, followed by the minimization of the Kullback-Leibler divergence between them.)[12, 13].

In the previous section the correlation between properties within and across different bonding types was discussed. However, the scope of these considerations was narrowed by limiting them to 2-dimensional representations. Different, more distinct and/or more intricate correlations might be present in higher dimensional representations, occluded in two dimensions. In order to make visualization beyond the third dimension feasible, a nonlinear low-dimensional embedding technique, in this case the so-called *t*-distributed stochastic neighbor embedding (t-SNE), is employed. The concept underlying this approach is shown in figure 1.10 for a 2D $\rightarrow$ 1D model example. The general idea of t-SNE is to reduce the dimensionality of a set of datapoints, while keeping the relative distances among all datapoints as equal as possible. Hence, datapoints that are located close to each other in the 7D property space used so far will also be located close to each other when reduced to a 2D representation using this method. This 2D representation of the property space is shown in figure 1.11. In the t-SNE representation, different sets of properties (and thus different bonding types) are reasonably well separated. Especially the metallic cluster is distinctly partitioned off, which is to be expected, as metals feature the most characteristic properties, in particular the zero band gap $E_G = 0\,\text{eV}$ and elevated Born effective charge $Z^*_+ = 0$ in addition to high electrical conductivities, distributed over a relatively narrow range (see table 1.3). Interestingly, there is a split-off cluster of metallic compounds, which is located closer to the central covalent group than to the main metallic cluster. This split-off cluster contains Ga, $AgSnTe_2$, $AgSnSe_2$ and In_3SbTe_2. These materials are different from the remaining metals, as $AgSnTe_2$, $AgSnSe_2$ and In_3SbTe_2 are considered so-called *strange metals* or *bad metals*, encompassing inter alia lower electrical conductivities compared to regular metals, motivating their separate location on the t-SNE representation. Regarding Ga it can be argued that it is an elemental metal, which

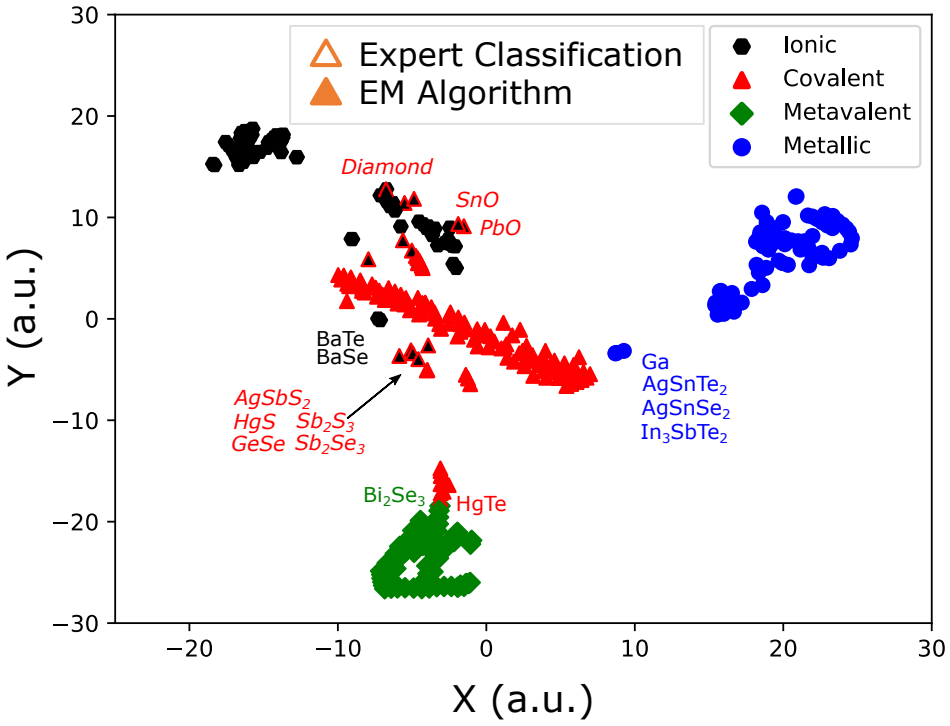


Figure 1.11: The *t*-distributed stochastic neighbor embedding (t-SNE) representation of the 7D property space, reducing it to 2D. The relative distances between datapoints in 7D property space are preserved as much as possible. The expert classification is indicated by the color of the outline of the markers, while the EMA classification is denoted by the body color of the markers. Different bonding types are separated, with especially metallic and MVB compounds forming distinct clusters. The separation of covalent and ionic compounds is less pronounced. Italic font indicates misclassified compounds, while roman script is used to label compounds close to the border of a different bonding type. It should be noted that the t-SNE distribution of the datapoints is independent of the clustering results by the EMA (or the expert classification). The t-SNE does not classify the datapoints, nor do the results of the EMA influence the position of datapoints in this figure. Taken from [2].

separates it from the main cluster. However, other elemental metals like Al and Ni are not split off, contradicting this notion. The MVB cluster is also located quite remotely from the other clusters. This is possibly due to the uniquely large elevated Born effective charge Z^*_+ and similar electric conductivities $\log(\sigma)$ and effective coordination numbers within this group of compounds. The top end of the mainly metavalent cluster consists of covalently bonded materials. This proximity could be caused by similarly small band gaps and moderate electrical conductivities of covalent materials like HgTe[6]. They differ however in coordination and hence ECoN, as the metavalent compounds are octahedral-like, while the covalent materials are

[6]That is in comparison with the metavalent compounds in general, and Bi_2Se_3 in particular, as it borders the covalent outpost.

tetrahedrally coordinated. The covalent and ionic clusters feature the most overlap, which comes to no surprise considering figure 1.8. While overlap does exist, the top left ionic cluster and the central covalent cluster are separated quite well. The regions between these and around the covalent central cluster are less defined, which is where the materials are located for which EMA and expert classification deviate. These shortcomings once again underscore the challenge of rigorously distinguishing ionic from covalent compounds.
Figure 1.11 shows that the EMA outperforms the t-SNE method in terms of assigning the datapoints to clusters, at least when employing the expert classification as benchmark. Considering only the t-SNE representation (ignoring color), one would probably assign the covalent points at the top of the bottom cluster to that cluster, and Ga, $AgSnTe_2$, $AgSnSe_2$ and In_3SbTe_2 to the central covalent cluster, which would also include all datapoints scattered around it. Nevertheless, clear borders do exist separating out metavalent and metallic compounds within the t-SNE, meaning the t-SNE promises to provide insights into bonding - property relations when used in conjunction with the EMA. Figure 1.12 depicts the t-SNE representation in combination with the EMA results for 2-5 allowed clusters. While the depiction for 2 and 3 clusters provides little new information, it is interesting to note that for 5 allowed clusters the additional cluster (orange) mainly consists of ionic compounds located above the main covalent cluster, indicating that these materials share a similarity in properties. Also, a couple compounds close to the covalent central cluster misclassified as ionic within the 4-cluster picture are now correctly classified as covalent, implying the EMA as well perceives them to be at the threshold between covalent and ionic.
So far, it has been shown that property space has inherent structure, and that the various types of chemical bonding appear to be fundamental building blocks of property space. While these are noteworthy findings in themselves, the goal set at the beginning of this chapter was to be able to navigate property space to tailor materials and their properties by design. Figure 1.11 and 1.12 already provide representations similar to a map. One might hence assume that the t-SNE approach could be used as a roadmap of property space. Unfortunately, this notion is deceiving, as the artificial 2D[7] coordinates of the t-SNE representations hold no physical meaning. They do not contain any information on how to modify a given material in order to move its position on the t-SNE "map" in a specific direction. After all, the t-SNE coordinates are derived from the properties to be tailored. Hence consulting the t-SNE representation to tailor properties will result in circular reasoning. This can be exemplified by imagining a t-SNE representation that transforms a 7D $\rightarrow$ 7D property representation, i.e. not changing property space at all. Attempting to use property space itself to, for example, modifying a material to increase its electrical conductivity σ then only yields the result to "move on the map" in the direction of higher conductivities, which is a circular argument. Hence, the t-SNE representation unfortunately does not provide any (direct) insights into how tailoring of properties can be achieved and remains a tool of visualization.
The previous paragraph concluded that another set of coordinates is required to draw a usable map of property space. With chemical bonding being an integral part of property space, as

[7] Or any other dimensionality

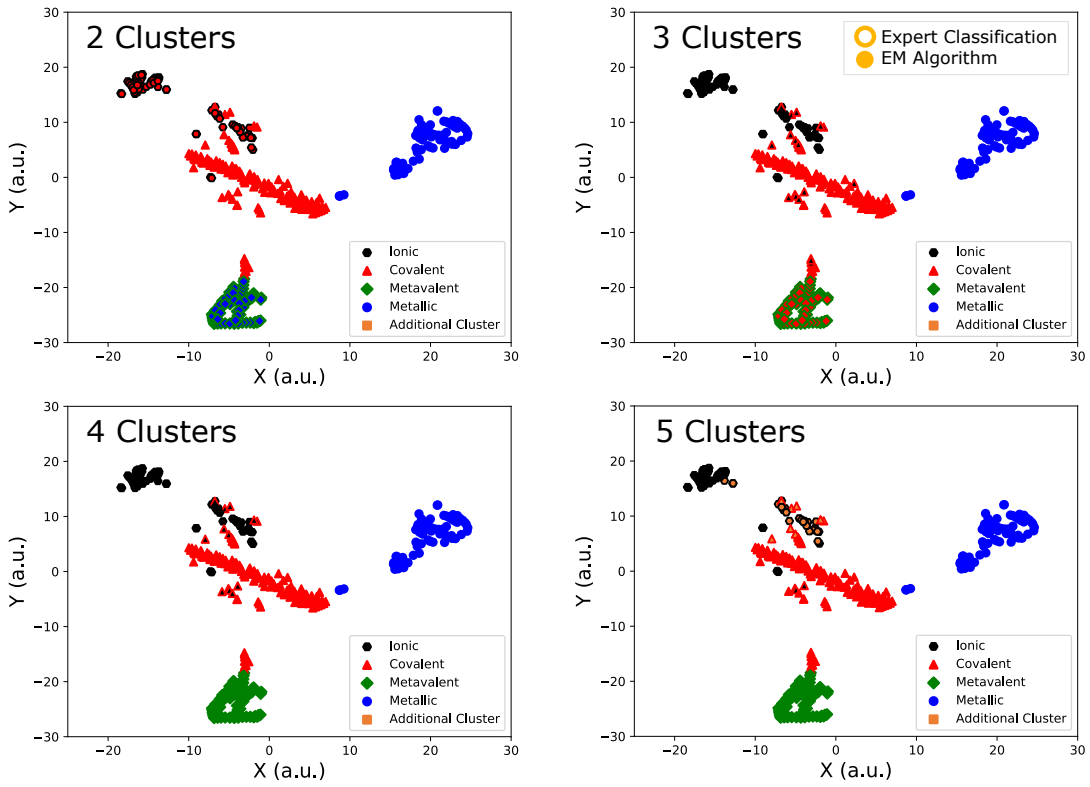


Figure 1.12: *t*-distributed stochastic neighbor embedding (t-SNE) representation of the 7D property space, reducing it to 2D. Varying numbers of allowed clusters are depicted, as found by the EMA (compare figure 1.2). The expert classification is indicated by the color of the outline of the markers, while the EMA classification is denoted by the body color of the markers. Two clusters: Ionic and covalent, as well as metavalent and metallic compounds are merged. Three clusters: Metals and ionic compounds form their own cluster, metavalent materials join the covalent cluster. Also, some covalent compounds at the top end of the bottom group are misclassified as ionic. Four clusters: Identical to figure 1.11. Five clusters: Some additional ionic and covalent materials are separated out. Taken from [2].

has been shown above, using parameters related to this concept appears promising. Indeed, the properties *Electrons Transferred* (ET) and *Electrons Shared* (ES) have been shown by Wuttig et al.[10, 14] to describe property trends and chemical bonding itself. ET describes the relative number of electrons transferred between two atomic sites, i.e. the total number of electrons transferred (TET) divided by the formal oxidation state, while ES denotes the number of electrons shared between adjacent atoms. A value of ET = 1 for NaCl hence implies that one electron is transferred between sodium and chlorine, while the same value of ET = 1 for MgO requires two electrons to be transferred, as the formal oxidation state of MgO is 2, while it is 1 for NaCl. An ES value of ES = 2 denotes that two electrons are shared, which corresponds to one electron pair. The ES and ET values for a given compound can be calculated by means of the quantum theory of atoms in molecules (QTAIM), which has been developed by Richard Bader over time since the 1960s[15–18]. A more detailed description

of QTAIM and the calculation of ES and ET in practice can be found in chapters 6 and 7, respectively. Figure 1.13 employs ES/ET as coordinates to span a map of compounds using the materials utilized for classification with the EMA. This type of map again separates materials

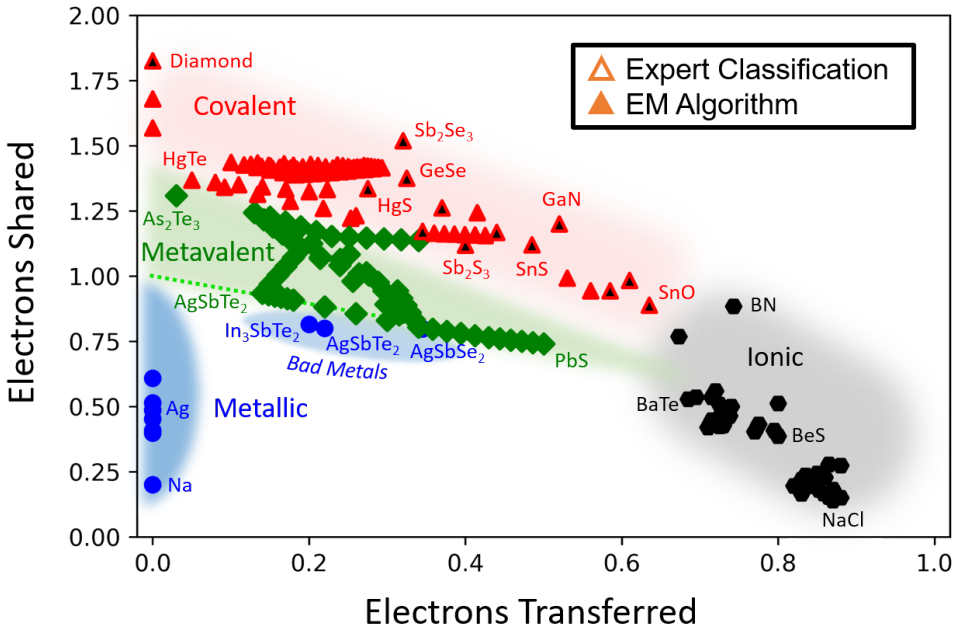


Figure 1.13: A map using the (relative) number of electrons transferred (ET) and electrons shared (ES) as coordinates. The different bonding types are denoted by color and symbol and form separate clusters using this set of coordinates (the border color denotes the expert classification; the body color indicates the results of the EMA). The broken green line indicates the suspected border between metallic and metavalent compounds. Compounds utilizing d-electrons (i.e. materials with occupied d shells, e.g. Au or Cr) have been excluded, as the ES value of these compounds behave differently from s- and p- bound systems. Taken from [2].

employing different types of chemical bonding. Ionic materials are located in the bottom right-hand corner, which corresponds to high values for ET and low values for ES. This is to be expected, as ionic bonding is driven by transferring charge to create electrostatic forces. The opposite occurs in covalent bonding. Low values for ET and a high number of ES position themselves at the top left-hand section of the map. Again, this meets with expectations, as the ideal covalent compound is characterized by the sharing of one electron pair between adjacent sites (2 center - 2 electron bonding). Metavalent compounds are positioned in the center-left region of the map, sharing only about 1 electron between sites, rendering ES ≈ 1 and featuring relatively low ET values of ET < 0.5. Delocalized electrons constitute metallic bonding, resulting in low values of both ES and ET. While this separation of bonding types underlines the fundamental character of chemical bonding once again, the usefulness of the ES/ET map must be judged by its ability to not only describe bonding, but by its capacity to predict properties. Figure 1.14 shows several 3D versions of the ES/ET map for a selection

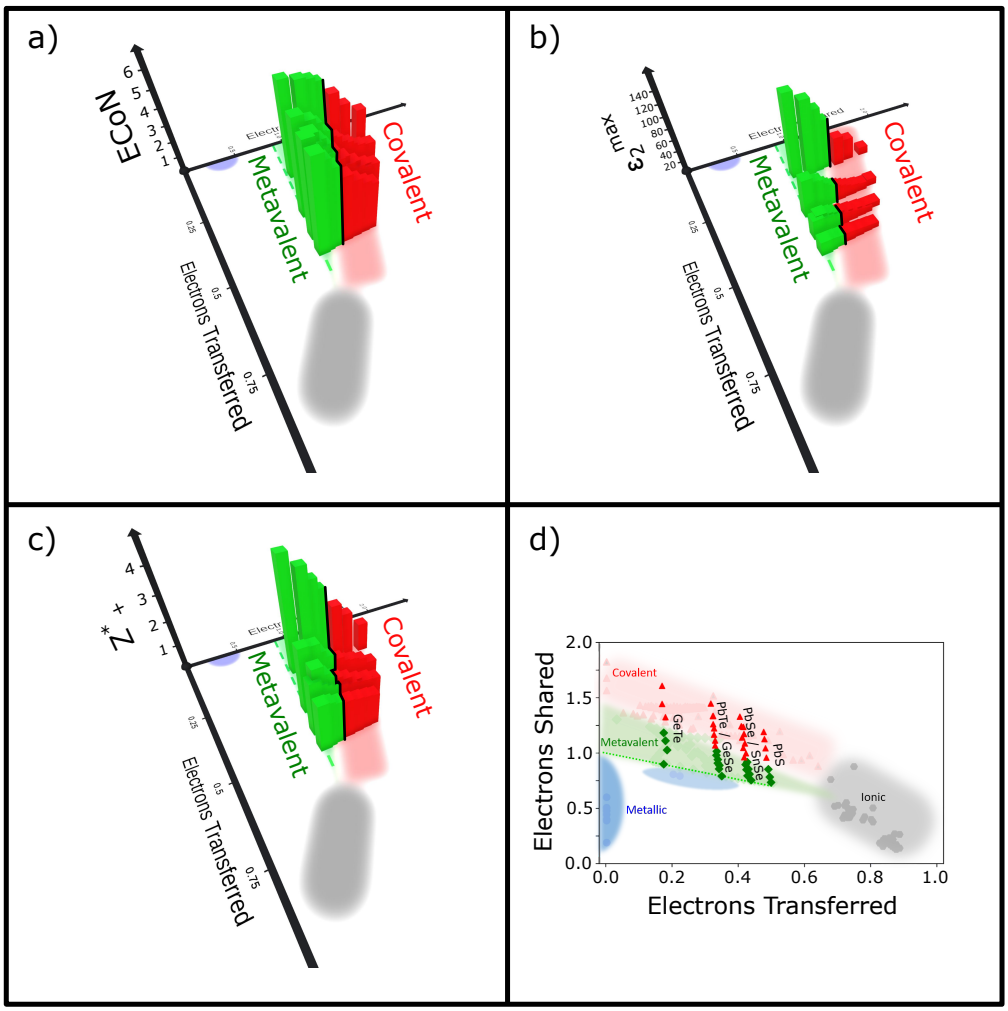


Figure 1.14: 3D version of the ES/ET map, showing the ECoN (a), maximum of the imaginary dielectric function $\epsilon_2^{\max}$ (b) and the elevated Born Effective Charge Z_+^* (c) on the z-axis for various monochalcogenides (PbTe, PbSe, PbS, SnTe, SnSe, SnS, GeTe) with different degrees of distortion. Panel d) shows the 2D version of the ES/ET map with the monochalcogenides highlighted. Taken from [2].

of monochalcogenides (PbTe, PbSe, PbS, SnTe, SnSe, SnS, GeTe) with different degrees of distortion. Monochalcogenides have relevant applications as thermoelectric and phase-change materials as well as topological insulators[19–22], so exploring their property portfolio using ES and ET promises to offer insight. The perfectly cubic, octahedrally coordinated compounds (no Peierls distortion) are located on the broken green line ($\mathrm{ES} = 1 - 0.52 \times \mathrm{ET}$), as shown on Panel d) (or figure 1.13). Increasing the (Peierls-) distortion of any given monochalcogenide increases the ES value, while keeping the ET value almost constant. Panel a) of figure 1.14 features the ECoN (see chapter 7.4) on the z-axis. The octahedrally coordinated compounds at the broken green line display an ECoN of 6, which then decreases with increasing ES.

Panel b) shows the monochalcogenides' relation for the maximum photon absorption $\epsilon_2^{\max}$ with ES and ET. A distinct trend can be observed, as $\epsilon_2^{\max}$ decreases with increasing ES and ET, respectively. The highest values of $\epsilon_2^{\max}$ can be found on the dashed green line at low ET values. This behavior is caused by the change of the joint density of states (JDOS) and the matrix elements of the optical transitions, as postulated in Fermi's golden rule (see also chapter 5.7). Increasing ES (i.e. distorting the system) decreases the overlap of the wave functions of the initial and final state (orbitals are less aligned), hence reducing the matrix elements for the optical transitions and consequently $\epsilon_2^{\max}$[23]. Increasing ET increases the band gap which also reduces the overlap of atomic orbitals and therefore reduces $\epsilon_2^{\max}$[24] (see chapter 2.1.2, in particular figure 2.7). Panel c) depicts the elevated Born effective charge Z_+^* as a function of ES and ET. Similarly to Panel a) and b), a clear trend can be observed, as, on average, Z_+^* decreases with increasing ES and ET. The correlations for ECoN and Z_+^* with other properties in figure 1.9 are weaker compared to the correlation of ES and ET to ECoN and Z_+^*. This suggests that ES and ET are suitable property predictors for the monochalcogenides, and that ES and ET are "natural" variables to describe property trends. In contrast to the t-SNE coordinates (see figure 1.11), ES and ET can be used to deliberately navigate property space, as a recipe to modify them does exist. A simple method to alter ET is to replace an atomic site within the compound with another element from the same group. As seen in figure 1.14, replacing the tellurium in cubic PbTe (located on the broken green line) with selenium or sulfur (PbTe $\rightarrow$ PbSe $\rightarrow$ PbS) increases the ET value from 0.34 $\rightarrow$ 0.43 $\rightarrow$ 0.50 respectively, while keeping the ES value almost constant (0.80 $\rightarrow$ 0.76 $\rightarrow$ 0.74). Conversely, ES can be increased by gradually distorting the system in question (see for example chapter 2.4).
An interactive version of the 3D ES and ET map has been developed within the scope of this thesis and can be accessed at[8]:

materials-map.rwth-aachen.de

Selecting a compound reveals the numeric values of its properties. This interactive map also features older, alternative types of bonding maps, i.e. the van Arkel/Ketelaar triangle and the Littlewood map (see figures 8.1 and 8.2 in the appendix)[10] to enable comparison between the different generations of bonding maps in terms of bonding description and property prediction. Compared to the ES/ET map, the separation of different bond types is much weaker within the framework of the van Arkel/Ketelaar triangle and the Littlewood map. Nor can they distinguish between allotropes of the same structure, which is another downside compared to the ES/ET map. Sample images of the interactive map can be found in chapter 10 in the appendix.
It can be concluded that the ES/ET map is a powerful framework for navigating property space. Its application will be illustrated in the following chapters.

[8] as of May 2023

Chapter 2

Results - The Exploration of Property Space

In the previous chapter, it was shown that the property space of effective coordination number ECoN, electrical conductivity σ, band gap E_G, elevated Born effective charge Z^*_+, density ρ, atomic density ρ_A and melting point T_M is intrinsically structured and that the parameters ES and ET can be used to effectively navigate this property space by describing bonding and predicting property changes. The algorithmic approaches shown also underscored the existence of metavalent bonding as an independent and unique bonding type. Within this chapter, various applications for the ES/ET bonding map and metavalent bonding will be explored, while highlighting the usefulness of these concepts as tools to explain and describe miscellaneous phenomena in material science.
For this purpose, each section of this chapter will feature a publication or thesis that has been co-authored or supervised by the author of this dissertation.

2.1 The Nature of Chemical Bonding in Chalcogenides

The review paper "Revisiting the Nature of Chemical Bonding in Chalcogenides to Explain and Design their Properties" was first-authored by Matthias Wuttig and published in the *Advanced Materials Hall of Fame* - series[3]. Within this publication, the previously introduced bonding descriptors ES and ET are utilized for a quantitative description of bonding in chalcogenides as well as to predict property changes. Furthermore, the concept of metavalent bonding is discussed more thoroughly and compared to alternative concepts like hypervalent bonding and lone pairs.

2.1.1 Chemical Warfare

> In the midst of chaos, there is also opportunity.
>
> *Sun Tzu, The Art of War*

The advent of quantum mechanics brought unrest into the world of chemistry. While it was soon widely recognized as an indispensable tool to describe chemical phenomena, especially bonding, the interpretation of the results sparked fierce disputes[25–27]. Although significant advancements were made in understanding of the covalent bond, which Linus Pauling captured

in his book "The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry"[6], he inevitably found himself in a verbal tug-of-war regarding the valence bond (VB) versus molecular orbital (MO) interpretation of the covalent bond[28]. The VB and MO concepts are both one-electron orbital-based descriptions, rewriting the wave function Ψ as a single determinant, while the one-electron orbitals are determined by constructing an average effective field created by the other electrons and nuclei in the system. As the MO and VB approaches are based on the same concept, they even yield identical results in the limit of complete convergence. Still, being this closely related did not suffice to settle the disputes among chemists, as the chosen basis set of orbitals can impact the respective outcomes of the calculations when full convergence is not feasible. Additionally, in the case of electron correlation becoming relevant, the single-determinant approaches are no longer accurate, and MO and VB descriptions, when limited to the single-determinant approximation, are no longer sufficient.
Rivaling the MO and VB approaches is the probabilistic interpretation of the wave function Ψ. Within this framework, the probability densities of electrons, stochastically describing the location of electrons and/or pairs of electrons, is analyzed. This approach has the advantage of not requiring a choice of a set of orbitals, being independent of computational models and methods and yielding measurable properties, enabling confirmation by experiment. Based on these concepts, there are the so-called *Quantum Chemical Topological* (QCT) approaches, with the *Quantum Theory of Atoms in Molecules* (QTAIM) being among the most prominent examples (see chapter 6)[17].
While the advantages of the QCT approaches appear compelling, the orbital-based methods remain relevant, partly due to their being an established practice, but also due to their providing of an alternative perspective which can yield complementary insights. Still, as the conclusions drawn from these methods can also be incompatible, the controversies regarding interpretation are far from resolved. The Nobel laureate Roald Hoffmann stated in 1988: "Many solid-state chemists have isolated themselves from their organic or even inorganic colleagues by choosing not to see bonds in their materials"[29]. With solid-state physicists focusing primarily on the atomic arrangement and the resulting electronic band structure (see chapters 5.1 and 5.6), the necessity of chemical bonding has been challenged on a fundamental level[30]. Still, the concept of chemical bonding is being used to explain the atypical behavior of functional materials, e.g. IV- VI group thermoelectrics, which convert thermal to electrical energy, or phase-change materials, which feature a distinct optical and electrical contrast upon switching from the amorphous to the crystalline phase. To explain the differences between these unique group IV - VI type semiconductors and the more conventional group III - V and II - VI compounds, differences in chemical bonding have been cited. For mono chalcogenides[1], e.g. GeTe, a surprising variety of bonding concepts have been invoked to explain their unusual set of properties:

[1]Mono chalcogenides have the chemical formula of the shape ME, where M is a transition metal like Germanium (Ge), and E a chalcogen, which roughly translates to "ore-forming", and refers to the chemical elements of group 16: Oxygen (O), sulfur (S), selenium (Se), tellurium (Te), polonium (Po) and livermorium (Lv).

1. **Hypervalent (3-center - 4-electron)**: The so-called octet rules states that elements tend to bond in such a way that the outermost valence shell is occupied by 8 electrons. Hypervalence in molecules describes the situation where this octet rule is violated. An example of a hypervalent molecule is SF_6, which features 12 electrons in the valence shell of sulfur. GeTe too has been ascribed this characteristic of forming electron-rich, hypervalent multi-center bonds, in particular 3-center - 4-electron bonds[31, 32]. Yet, with three orthogonal bond axes, GeTe would require 12 electrons to form 3 orthogonal hypervalent bonds.

2. **10 Electron Solids**: Within this framework, all formal valence electrons are considered to take part in bonding. This hence includes the electrons of the Ge $4s^2$ and $4p^2$ states, as well as the electrons of the $5s^2$ and the $5p^4$ state of Te.

3. **Lone Pairs**: A lone pair describes a pair of valence electrons that is not shared with (or transferred to) another atomic site. While oxygen features 6 valence electrons, 2 of them form a shared pair in the H_2O molecule, while the remaining 4 form 2 lone pairs, responsible for the bond-angle of 104.5°[33–35]. Regarding GeTe, it is considered to be an ionic material where the Ge $4s^2$ state forms a lone pair, impacting the material properties.

4. **Metavalent Bonding (2-center - 1-electron)**: In contrast to the "10 Electron Solids" picture, metavalent bonding (MVB, see chapter 1.2) considers only the electrons of the Ge $4p^2$ states and the electrons of the Te $5p^4$ states, in total 6 electrons, to be actively partaking in bond formation. As both the Ge and Te sites in GeTe feature six nearest neighbors, each bond consists of only one shared electron, or half an electron pair (2-center - 1-electron). The bond can hence be described as electron-deficient.

Table 2.1 summarizes the different bonding approaches in terms of required electron count. Four different variants of chemical bonding, which greatly disagree in terms of the total number of electrons involved, are being proposed to describe a presumably simple, binary compound. It is remarkable how little consent can be reached for a seemingly elemental question within the field of chemistry.
To assess which mechanism is present in GeTe, several methods can be employed. For the sake of completeness, the cubic ($Fm\bar{3}m$), rhombohedral ($R3m$), and orthorhombic ($Pnma$) phases of GeTe will be discussed.

Table 2.1: Bonding types proposed to describe GeTe. For hypervalent bonding, no information is available on which orbitals would provide the 12 electrons required. Modified from [3].

Bonding Scheme	Contributing States	Total Number of Electrons
Hypervalent	-	12
10 Electron Solids	Ge $4s^2$ $4p^2$ - Te $5s^2$ $5p^4$	10
Lone Pair	Ge $4s^2$ $4p^2$ - Te $5p^4$	8
MVB	Ge $4p^2$ - Te $5p^4$	6

Band Structure, Density of States (DOS) and Integrated Density of States (IDOS)

The density of states (DOS) is closely related to the band structure, as it denotes the number of states per unit volume within a given energy range (see chapter 5.6). Consequently, flat bands in the band structure result in a large DOS at the corresponding energy and vice versa. The integrated DOS (IDOS) is a measure of how many electrons are implemented in a specific orbital up to an energy value in question. To calculate the DOS, the DFT implementation QUANTUM ESPRESSO has been employed (see chapters 7.1 and 7.1.3). The IDOS for the three GeTe phases are depicted in figure 2.1, while the band structures are depicted in figure 2.2. Non-overlapping (atomic) orbitals do not hybridize. Hence their

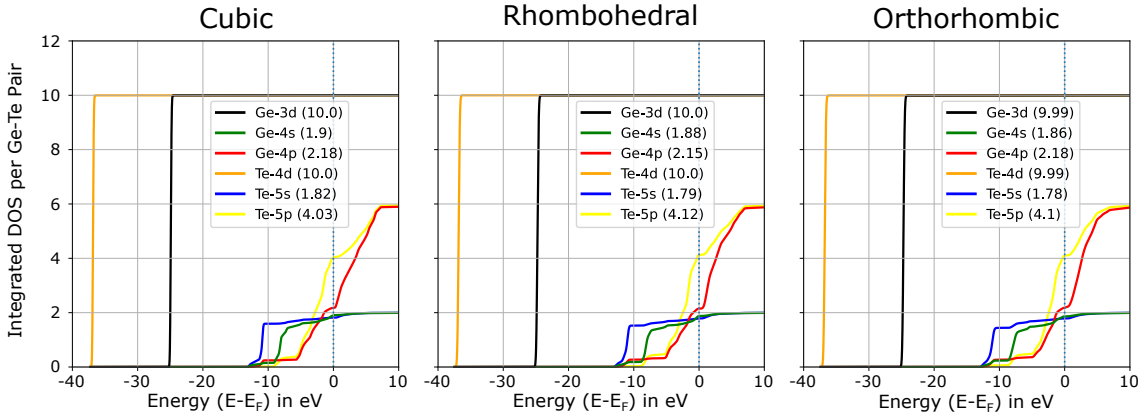


Figure 2.1: Integrated density of states (IDOS) for three phases of GeTe: Cubic ($Fm\bar{3}m$), rhombohedral ($R3m$) and orthorhombic ($Pnma$). The total electron occupations of the respective states up to the Fermi energy E_F are denoted in brackets. The IDOS and occupations are almost identical for all three phases. Adapted from [3].

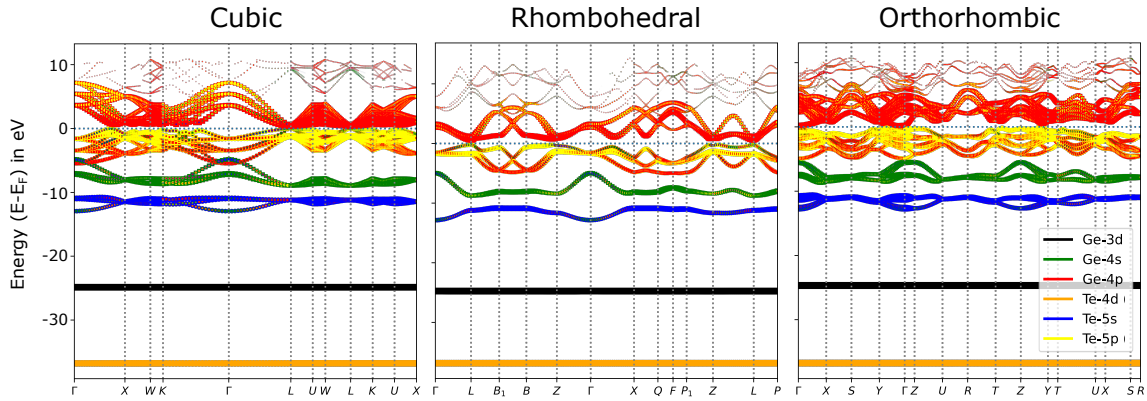


Figure 2.2: Orbital-resolved band structure for three phases of GeTe: Cubic ($Fm\bar{3}m$), rhombohedral ($R3m$) and orthorhombic ($Pnma$). The size of the markers indicates the relative contribution of the respective state.

energy levels resemble the discrete energy steps of an isolated atom. This leads to flat bands in the band structure, sharp peaks in the DOS and a step-like behavior in the IDOS, as can be seen for the Ge 3d and Te 4d states for all phases. The almost discrete jump of the IDOS for the GeTe phases hence implies that the d states do not overlap, interact and/or form electron pairs with any adjacent sites and thus do not contribute to chemical bonding. This notion is further reinforced by considering the IDOS value of the d states at the Fermi energy IDOS(E_F) = 10, meaning that the d states are completely filled, underlining their inert character[2]. Orbital overlap leads to the dispersion of bands and hence broader peaks in the DOS and smoother increases in the IDOS. This behavior starts to appear for the Ge 4s and Te 5s states, albeit not that pronounced. The s states are also almost filled at the Fermi energy E_F, implying that while some overlap with the Ge 4p and Te 5p states does exist[3], their contribution to bonding is presumably minor, but cannot be ruled out rigorously from these considerations. The Ge 4p and Te 5p states show significant overlap and are also only about half-filled at the Fermi energy E_F. It can hence be assumed that p states are the main contributors to chemical bonding within GeTe. The occupation of all (valence) states as inferred from the IDOS is summarized in table 2.2. The occupations do not vary substantially among the three phases, indicating that atomic arrangement does not impact the IDOS significantly. Figure 2.3 shows the IDOS of silicon, where neither 3s nor 3p states are filled completely, implying that the s states make a more pronounced contribution to the forming of chemical bonds (as expected from an sp^3 hybridized compound). GeTe forms three (almost) perpendicular bonds. Twelve electrons in total would be necessary to form three hypervalent bonds, which would be required for such a system. Table 2.2 and figure 2.1 contradict the notion of hypervalent bonding (3 center - 4 electron bonding), as not enough electrons are available. On the contrary, the data offers much more evidence for electron-deficient metavalent bonding, as (almost) exclusively the p states feature occupied and unoccupied states around the Fermi energy E_F. Still, the contribution of the s states cannot be dismissed with absolute certainty, as a small fraction of them remains

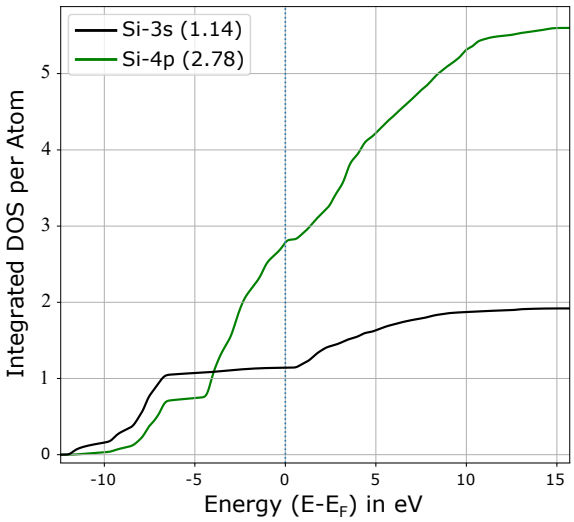


Figure 2.3: Integrated density of states (IDOS) for silicon ($Fd\bar{3}m$). The total electron occupations of the respective states up the Fermi energy E_F are denoted in brackets.

[2] With only one steep increase in the corresponding IDOS, they are already filled at the respective energies where that increase occurs, i.e. ≈ -38 eV for the Te 4d states and ≈ -25 eV for the Ge 3d states, far below the Fermi energy E_F.

[3] That s and p states overlap can be inferred from the band structure, where bands with s- and p- contributions exist, as well as from the IDOS, where s and p state values increase at identical energy ranges.

Table 2.2: Occupation of the different states of the three phases of GeTe, as obtained from the integrated density of states (IDOS). The s states are almost completely filled, reaching about 1.8-1.9 out of 2 electrons, while the p states are about half-filled, with the Ge 4p states at around 2.2 out of 6 and the Te 5p states at around 4.1 out of 6 electrons, respectively. The occupations do not vary substantially among the three phases, indicating that atomic arrangement does not impact the IDOS significantly. The total number of electrons does not reach 10 completely due the atomic orbital basis used for projection not being complete, and would require additional, higher states (like Ge 5d- and Te 5d states etc.). Modified from [3].

GeTe Phase	Ge 4s	Ge 4p	Te 5s	Te 5p	$\sum$ **s states**	$\sum$ **p states**	$\sum$ **All states**
Cubic ($Fm\bar{3}m$)	1.90	2.18	1.82	4.03	**3.72**	**6.21**	**9.93**
Rhomb. ($R3m$)	1.88	2.15	1.79	4.12	**3.67**	**6.27**	**9.94**
Ortho. ($Pnma$)	1.86	2.18	1.78	4.10	**3.64**	**6.28**	**9.92**

unoccupied and could therefore contribute to bond formation. To investigate this matter further, more advanced methods have to be employed.

The Crystal Orbital Bond Index (COBI)

The bond order (BO) quantifies the number of electron pairs constituting a (covalent) bond. In classical chemistry, the diatomic oxygen molecule O_2 ($O=O$) would be assigned a bond order of $BO = 2$, while for diatomic nitrogen N_2 ($N \equiv N$) the bond order would be $BO = 3$. Wiberg and Mayer have developed a quantum mechanical framework of the bond order for molecules[36, 37], while Dronskowski et al. have further extended it to solids[38]. The quantity derived by Dronskowski et al. is called *Crystal Orbital Bond Index* (COBI), while the corresponding integrated property is called ICOBI. The ICOBI values for the three GeTe phases are shown in table 2.3. The ICOBI value of cubic GeTe is approximately $\text{ICOBI}_{\text{GeTe}}^{\text{cubic}} \approx 0.4$,

Table 2.3: Integrated crystal orbital bond index (ICOBI) for cubic ($Fm\bar{3}m$), rhombohedral ($R3m$), and orthorhombic ($Pnma$) GeTe. The number of bonding electrons is calculated as $2 \times \text{ICOBI} \times \text{ECoN}$, where ECoN is the effective coordination number (see chapter 7.4). The ICOBI values were calculated by Jakob Lötfering using the LOBSTER software package[39]. Adapted from [3].

GeTe Phase	**ICOBI**	**ECoN**	**Bonding Electrons** ($2 \times \text{ICOBI} \times \text{ECoN}$)
Cubic ($Fm\bar{3}m$)	**0.395**	**6.0**	**4.74**
Rhomb. ($R3m$)	**0.636**	**4.8**	**6.10**
Ortho. ($Pnma$)	**0.755**	**3.4**	**5.14**

Table 2.4: Electrons shared (ES) and electrons transferred (ET) values for cubic ($Fm\bar{3}m$), rhombohedral ($R3m$), orthorhombic ($Pnma$) GeTe, NaCl ($Fm\bar{3}m$) and diamond ($Fd\bar{3}m$). The total electron transfer has been renormalized by the formal oxidation state, i.e. 2 for GeTe and 1 for NaCl. As the charge transfer is symmetric, the absolute value is listed. The ES value of the shortest bond has been stated. The DFT and QTAIM calculations have been conducted using ABINIT/QUANTUM ESPRESSO and DGRID/CRITIC2 (see chapter 7). Adapted from [3].

	Cubic GeTe ($Fm\bar{3}m$)	**Rhomb. GeTe** ($R3m$)	**Ortho. GeTe** ($Pnma$)	**NaCl**	**Diamond**
ET	0.18	0.18	0.18	0.87	0.00
ES	0.88	1.30	1.36	0.13	1.82

which corresponds to a bond order of about $BO_{GeTe}^{cubic} \approx 0.8$. This result is compatible with the electron deficient metavalent bonding, where two atoms are bonded by half an electron pair, i.e. one electron, yet again incompatible with hypervalent bonding. While the ICOBI value is slightly higher for the orthorhombic and rhombohedral phases, the total number of bonding electrons remains in the range of 4.7 - 5.1 e^-, hence about one electron per bond, due to the change in coordination. In fact, calculating ES and ET of exemplary (electron-rich) hypervalent molecules like XeF_2 and SF_4 reveals that they are located on a different region of the bonding map, secluded from other compounds as well as metavalently bonded chalcogenides (see figure 8.3 in the appendix).

Electrons Shared, Electrons Transferred and the Domain Averaged Fermi Hole

As already introduced in chapter 1.4, the quantum chemical variables electrons transferred (ET) and electrons shared (ES) are useful tools to characterize chemical bonding (see also chapter 6). The respective values for the GeTe phases, as well as NaCl and diamond, are listed in table 2.4. The ET values for all three GeTe phases are virtually identical (ET = 0.18). This value is much smaller than the ET value of the archetypical ionicially bound NaCl, which indicates that the ionic character of the bonds in GeTe[4] is minor. The ES values of GeTe range between ES = 0.88 − 1.36 (for the shortest bond, respectively). For an archetypical covalent bond, an ES value of ES = 2, corresponding to one shared electron pair, is expected. While diamond features an ES value of only about ES ≈ 1.82 (see table 2.4), this value is sufficiently close to one shared electron pair to confirm diamond as a prime example of a covalent compound. With the ES values of GeTe being considerably lower at around ES ≈ 1, especially for the cubic phase, the bonding in GeTe deviates considerably from covalent bonding. GeTe furthermore opens a band gap $E_G > 0$ due to the charge transfer from Ge → Te (ET > 0). This renders it a semiconductor, ruling out metallic bonding as well.
Using the domain averaged Fermi hole (DAFH) orbitals (see chapters 6.2 and 7.2.3), the contribution of the atomic orbitals to bond formation can be assessed. The results are shown in tables 2.5, 2.6 and 2.7.

[4] If no phase is specified, "GeTe" refers to all phases.

Table 2.5: Domain averaged Fermi hole (DAFH) occupations of cubic GeTe ($Fm\bar{3}m$). The elements in brackets in the far right column denote the bonding partners. The delocalization index (DI) denotes the number of shared electron pairs: $\mathrm{ES} = 2 \times \mathrm{DI}$ (see chapters 6.2 and 7.2.3). Modified from [3].

$Fm\bar{3}m$	**Occupation (e^-)**	**Localization in Native Basin**	**Overlap to Partner Basin**	**DI Contribution**
		Germanium		
4p (3x)	0.51	27.5%	27.3% (2x)	0.28 (Ge-Te)
4s	1.74	87.7%	1.7% (6x)	0.06 (Ge-Te)
3d (5x)	2.00	>99.9%	-	-
		Tellurium		
5p (3x)	1.31	66.0%	11.4% (2x)	0.30 (Te-Ge)
5s	1.88	94.3%	0.8% (2x)	0.03 (Te-Ge)

Table 2.6: Domain averaged Fermi hole (DAFH) occupations of rhombohedral GeTe ($R3m$). The elements in brackets in the far right column denote the bonding partners. The delocalization index (DI) denotes the number of shared electron pairs: $\mathrm{ES} = 2 \times \mathrm{DI}$ (see chapters 6.2 and 7.2.3). Modified from [3].

$R3m$	**Occupation (e^-)**	**Localization in Native Basin**	**Overlap to Partner Basin**	**DI Contribution**
		Germanium		
4p (3x)	0.52	26.9%	39.2% 14.8%	0.41 (Ge-Te) 0.15 (Ge-Te)
4s	1.73	86.3%	2.0% (3x) 1.8% (3x)	0.07 (Ge-Te) 0.06 (Ge-Te)
3d (5x)	2.00	>99.9%	-	-
		Tellurium		
5p (3x)	1.30	65.1%	15.0% 8.1%	0.39 (Te-Ge) 0.21 (Te-Ge)
5s	1.87	93.8%	0.9% (3x) 0.8% (3x)	0.03 (Te-Ge) 0.03 (Te-Ge)

Table 2.7: Domain averaged Fermi hole (DAFH) occupations of orthorhombic GeTe ($Pnma$). The elements in brackets in the far right column denote the bonding partners. The delocalization index (DI) denotes the number of shared electron pairs: ES = 2 × DI (see chapters 6.2 and 7.2.3). Modified from [3].

$Pnma$	Occupation (e^-)	Localization in Native Basin	Overlap to Partner Basin	DI Contribution
		Germanium		
4p (2x)	0.51	26.3%	50.0% 4.4%	0.51 (Ge-Te) 0.04 (Ge-Te)
4p	0.57	28.7%	54.9% 3.5%	0.62 (Ge-Te) 0.04 (Ge-Te)
4s	1.77	88.6%	1.6% 0.9%	0.06 (Ge-Te) 0.03 (Ge-Te)
3d (5x)	2.00	>99.9%	-	-
		Tellurium		
5p	1.28	64.4%	25.2%	0.62 (Te-Ge)
5p (2x)	1.31	65.6%	18.9%	0.50 (Te-Ge)
5s	1.89	94.9%	0.65%	0.02 (Te-Ge)

The Ge 3d orbitals do not contribute at all, as they are completely localized in their native basin and do not overlap with any neighboring basins[5]. The overlap to neighboring basins of the Ge 4s orbitals is minor as well ($\leq 2\%$). They hence do not noticeably contribute to bonding, either. The Te 5s orbitals behave similarly. The absence of these s orbital contributions contradicts the concepts of hypervalent bonding (3-center - 4-electron), 10 electron solid and the lone pair picture that assume and require contributions from the Ge 4s orbitals, and possibly even the Te 5s orbitals. Tables 2.5, 2.6 and 2.7 dictate that the almost exclusive contributors to bond formation in GeTe are the p orbitals, each providing on average about 0.3 electron pairs.

Figure 2.4 b) visualizes the data shown in tables 2.5 and 2.6, i.e. that three equal p-like DAFH orbitals (and thus identical ES values) are formed for the cubic phase, where no Peierls distortion is present and all neighbors are equally spaced (left-most point). Increasing the Peierls distortion, as quantified by the long to short bond ratio, the p-like DAFH orbital contribution to the shorter bond becomes more pronounced (ES to short bond increases; ES to long bond decreases). The arithmetic mean of the long and short bond ES values remains virtually constant, however, indicating that the p orbitals are aligned along a chain of atomic sites. Figure 2.4 a) indicates that the energetic difference between the stable $R3m$

[5]The term "basin" refers to Bader basins. Bader basins can be computed to attribute each point in space to a specific atomic site. Quantitative statements regarding the overlap of orbitals with other basins (corresponding to neighboring sites) can be made in this manner. For more details see chapter 6.

and the cubic $Fm\bar{3}m$ phase is relatively low, i.e. $< 10\,\text{meV}$ per atom, while the properties change drastically upon this small energetic difference, as shown in panel c). This is one of the characteristic features of MVB, caused by the fragility of the competition between localization and delocalization of the bonding electrons[6]. Furthermore, while the transition from the MVB to the covalent region in figure 2.4 seems mostly continuous, it has been shown experimentally that metavalently bonded compounds tend to undergo a phase transition upon crossing the border of the bonding regimes[23], resulting in discontinuous property behavior. This is indicated here by the arrow to the horizontal broken lines, which denote the covalently bonded $Pnma$ phase of GeTe.
SnSe shows similar behavior to GeTe, i.e. large changes in (optical) properties upon little energetic investment, and is further discussed in section 2.5.3.

[6]It also implies that a phase transition is easily achieved.

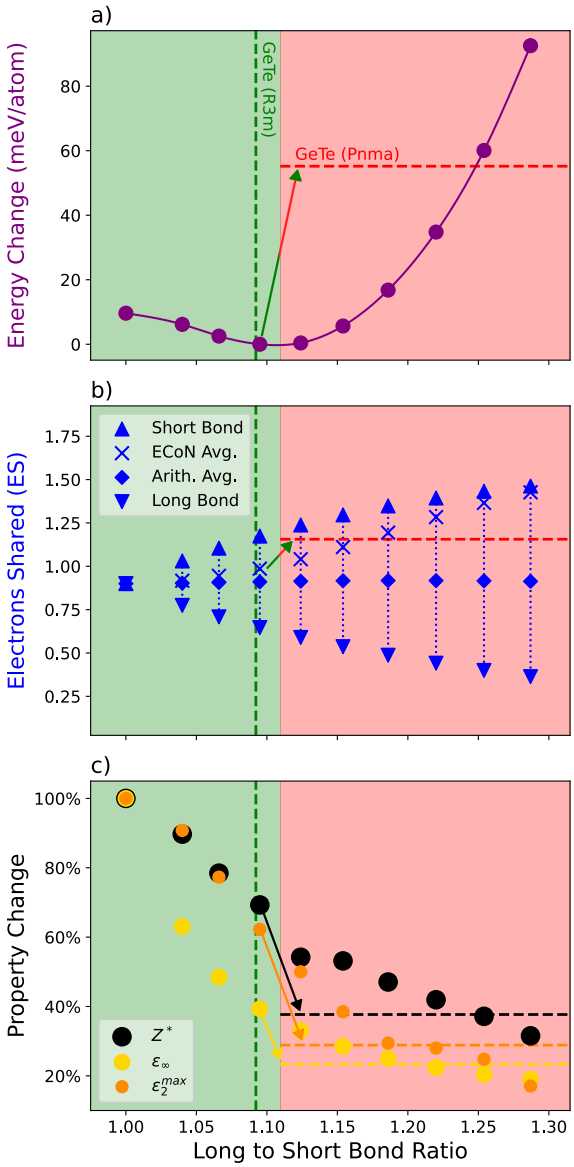


Figure 2.4: a): Energy cost per atom to deviate from the stable phase. Degree of Peierls distortion (PD) quantified by the long to short bond ratio. Point on extreme left: Cubic GeTe ($Fm\bar{3}m$); Broken green line: Rhombohedral ($R3m$) GeTe (stable phase); Broken red line: Energy per atom of the covalently bonded $Pnma$ GeTe (long to short bond ratio not applicable). b): ES of short bonds (triangles facing up), long bonds (triangles facing down), ECoN averaged ES (crosses) (see chapter 7.4) and arithmetic ES average (diamonds). The arithmetic ES average remains constant, while the ECoN ES average is similar to the arithmetic average for small distortions and approaches the short bond ES value for larger distortion. Broken red line: ECoN ES average of $Pnma$ GeTe. c): Relative changes of the Born effective charge Z^* (black), dielectric constant ϵ_∞ (red) and the maximum of the imaginary part of the dielectric function $\epsilon_2^{\max}$ (orange). All values relative to the cubic phase (100%), where $Z^*_{\text{cub.}} = 9.4$, $\epsilon_\infty^{\text{cub.}} = 144.9$ and $\epsilon_2^{\max,\text{cub.}} = 152.8$. Horizontal broken lines: Respective values of the $Pnma$ phase. Green background: Metavalently bonded phases; Red background: Covalently bonded phases. The properties appear to transition smoothly. However, it has been shown in experiments that upon the transition from metavalent to covalent bonding, compounds tend to conduct a phase transition as well (in this case $R3m \rightarrow Pnma$), causing discontinuous behavior[23], indicated by arrows. Adapted from [3].

Lone Pairs

While the approaches described above did not suggest that lone pairs may be involved in bonding of mono-chalcogenides, they also could not rule them out completely, either. Especially for the PbX systems, where X∈(Te, Se, S), the Pb $6s^2$ orbitals have been described to form lone pairs influencing properties. Such Pb $6s^2$ lone pair states should indeed play a prominent role if they are located energetically right below the Fermi energy E_F. This would imply that the p orbitals depopulate, and Pb assumes the oxidation state Pb^{2+}. Figure 2.5

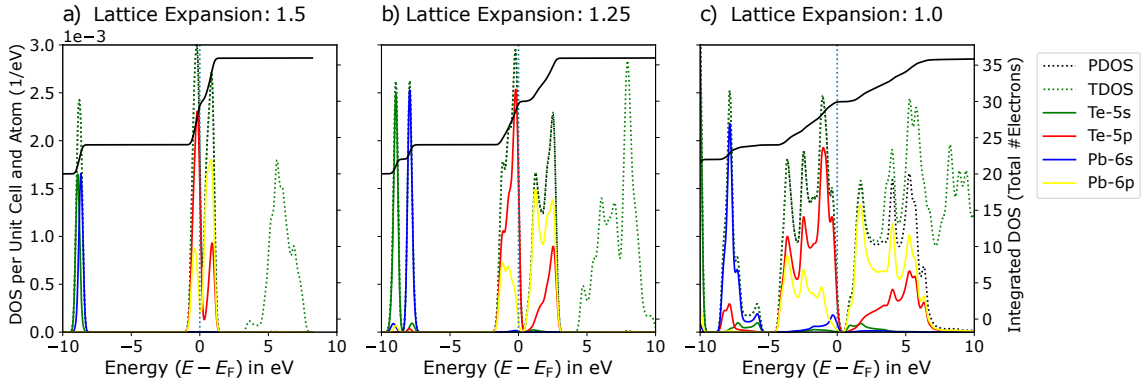


Figure 2.5: Orbital-resolved density of states (DOS) of PbTe in rock-salt structure ($Fm\bar{3}m$) for various lattice expansion factors: a) 1.50, b) 1.25, c) 1.00. By artificially increasing the lattice constant, the orbital overlap is decreased. The sum of all projected DOS values is denoted as PDOS (broken black line), while the total DOS is denoted as TDOS (broken green line). The total integrated DOS (IDOS) is depicted in black, using the right y-axis. Calculations conducted by Jakob Lötfering using the code described in chapter 7.5. Adapted from [3].

depicts the density of states for cubic PbTe ($Fm\bar{3}m$) for artificially expanded lattice constants to investigate the change in orbital overlap. Figure 2.5 a) expands the lattice by 50% (expansion factor of 1.5), resulting in a lattice constant of about $a_{1.5}^{PbTe} \approx 9.85\,\text{Å}$. This artificial (and unphysical) enlargement of the unit cell minimizes atomic overlap and causes the s states of Pb and Te to be completely filled far below the Fermi energy E_F. Minor Pb 6p and Te 5p orbital overlap remains, as can be seen from the corresponding peaks (red and yellow), right below the Fermi energy E_F, being slightly broadened. Figure 2.5 b) employs an expansion factor of 1.25, reducing the lattice constant to $a_{1.25}^{PbTe} \approx 8.21\,\text{Å}$. Due to the smaller distance between Pb and Te, orbital overlap increases for the p orbitals, indicated by the broadening of the corresponding peaks (red and yellow). Figure 2.5 c) shows the DOS of the equilibrium lattice constant, i.e. using an expansion factor of 1.0 and $a_{1.0}^{PbTe} \approx 6.57\,\text{Å}$. Orbital overlap between p orbitals significantly increases, while the Pb 6s orbitals also start to hybridize slightly with the p orbitals. The contribution of the Pb 6s orbitals remains miniscule, however, as confirmed by the DAFH calculations shown in table 2.8. The Pb 6s/Te 5s orbital overlap only contributes about 0.005/0.02 electron pairs[7] of the total of 0.4 electrons pairs, while the Pb 6p/Te 5p orbital

[7] Denoted in the table as delocalization index (DI): ES = 2 × DI, see chapter 6.

Table 2.8: Domain averaged Fermi hole (DAFH) occupations of cubic PbTe ($Fm\bar{3}m$) at equilibrium lattice constant (expansion factor of 1.0 and $a_{1.0}^{\mathrm{PbTe}} \approx 6.57\,\text{Å}$). The elements in brackets in the extreme right column denote the bonding partners. The delocalization index (DI) denotes the number of shared electron pairs: $\mathrm{ES} = 2 \times \mathrm{DI}$. Modified from [3].

PbTe (1.0) cubic	Occupation (e^-)	Localization in Native Basin	Overlap to Partner Basin	DI Contribution
		Lead		
6p (3x)	0.4	20.6%	30.8%	0.250 (Pb-Te)
6s	1.8	91.2%	0.13%	0.005 (Pb-Te)
5d (5x)	2.0	99.6%	<0.02%	0.004 (Pb-Te)
		Tellurium		
5p (3x)	1.4	72.3%	9.4%	0.260 (Te-Pb)
5s	1.9	96.1%	0.55%	0.021 (Te-Pb)

overlap provides about 0.25[8]. Bond formation is hence dominated by the p-p orbital overlap, while the Pb s^2 state contribution remains minor. The Pb 6p orbitals being filled with 0.4 electrons each also contradicts the lone pair picture requiring the Pb 6p orbitals to be empty.

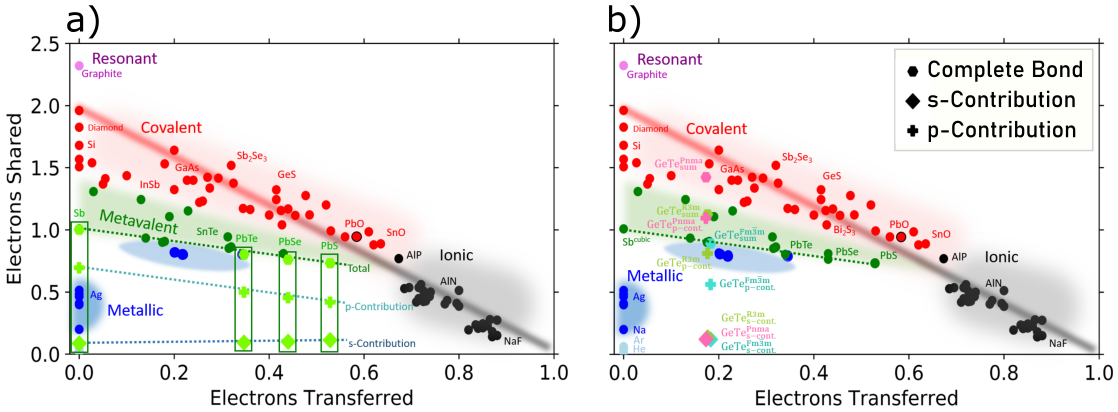


Figure 2.6: ES/ET map with orbital contributions for a): cubic Sb, PbTe, PbSe and PbS, b): GeTe in the rhombohedral ($R3m$, olive), cubic ($Fm\bar{3}m$, turquoise) and orthorhombic ($Pnma$, magenta) phase. The s-contributions are denoted by diamonds, the p-contributions by crosses and the complete bond by hexagons. Table 2.8 shows the DAFH values for PbTe, while the respective tables 11.1 and 11.2 in the appendix show the values for PbSe and PbS. Adapted from [3].

[8] The remaining 0.1 electron pair required to reach DI = 0.4 is provided by the sum of fractionally occupied higher orbitals that cannot be attributed to an atomic orbital, and/or are lost during projection unto the DAFH orbitals.

Figure 2.6 a) positions all PbX compounds along with cubic Sb on the ES/ET map, including the individual s and p orbital contributions. The slopes indicate that the s state contribution slightly increases with increasing ET, while the p state contribution decreases significantly. This decrease is the dominant cause for the reduction of optical absorption (see figure 2.7 c)), which will be discussed in more detail in section 2.1.2. Figure 2.6 b) shows the same orbital-resolved ES/ET map, but focusing on the three GeTe phases ($R3m$, $Fm\bar{3}m$, $Pnma$). The p orbital contributions as well as the total values differ drastically among the three phases, while the s orbital contributions remain minor and almost identical. Figure 8.4 in the appendix depicts the orbital contributions for different degrees of Peierls distortion of cubic GeTe. Again, the s-contributions remain minor and constant, while the p-contributions are responsible for the overall changes in ES. Consequently, the pronounced difference in properties of these phases cannot be attributed to the s orbitals and must originate from the p orbital contributions (see figure 2.4).

All approaches to describe bonding draw a coherent picture in favor of metavalent bonding[9]. The IDOS approach concluded that the s orbitals are almost completely filled and do not partake in bonding. This result is underlined by the DAFH analysis, stating that the contribution of the s orbitals is negligible to the DI/ES, as virtually no overlap exists to the basins of the bond partners. The bonding of GeTe is dominated by the p orbitals, which are about half-filled according to the DAFH and IDOS calculation. Hence about six p electrons are available to form six bonds to adjacent sites, resulting in an ES of about ES ≈ 1, in line with the results of the ICOBI calculations. The DAFH orbitals confirm the dominant role of the p orbital overlap (Overlaps of Ge p orbitals to neighboring basins of up to about 55% are reached in orthorhombic GeTe).

[9] Or classical covalent bonding in the case of orthorhombic ($Pnma$) GeTe.

2.1.2 Electrons Shared and Electrons Transferred as Property Predictors

> As soon as things become predictable,
> they become boring.
>
> *Hunter Parrish*

Figure 1.14 already indicated that ES and ET are good property predictors. In this chapter, arguments unraveling the origins of that predictive power are presented.

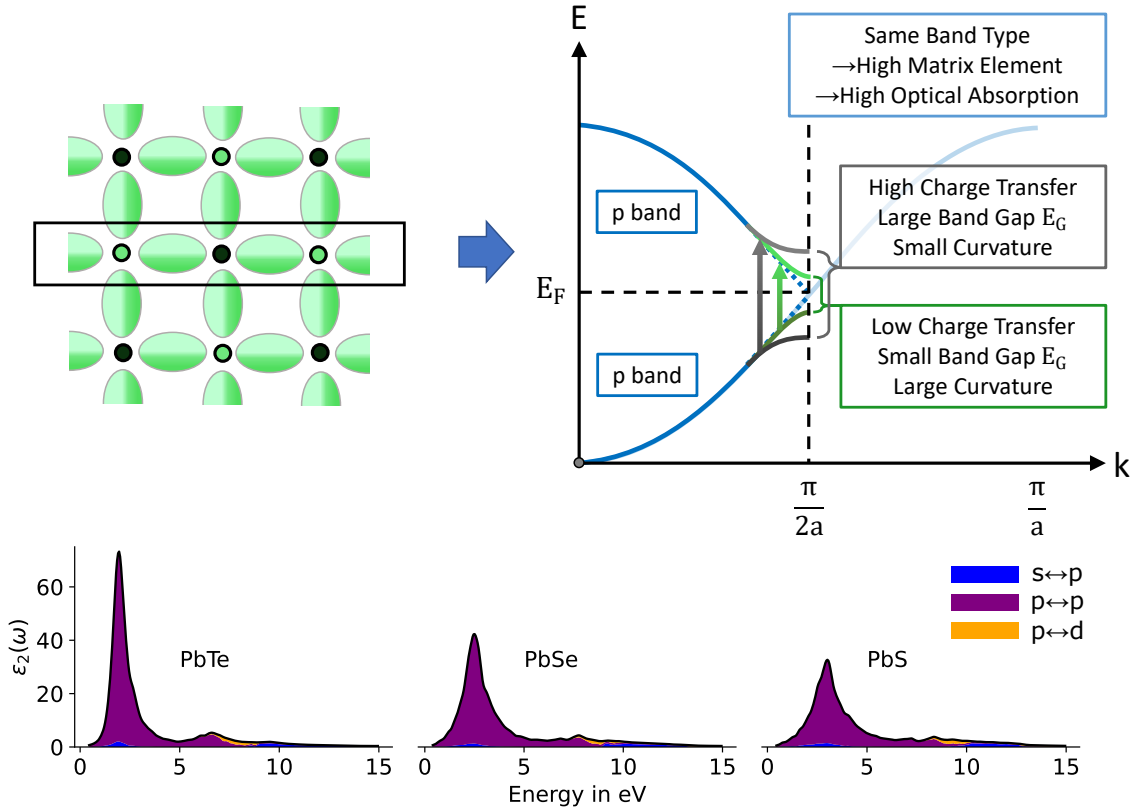


Figure 2.7: Sketch of bond formation in PbX systems, where X∈(Te, Se, S). The atomic arrangement of the (001) plane is depicted, along with the atomic 6p orbitals of Pb and the respective 3/4/5p orbitals of X responsible for bond formation. As the bonds are only half-filled by one electron or half an electron pair (ES = 1), they would typically form a metallic band (continuous, s-shaped blue band going from $0 \rightarrow \frac{\pi}{a}$). Due to the (rather small) electron transfer in these systems however, a band gap opens and the Brillouin zone is effectively halved as (charge) symmetry is broken. Bottom: Imaginary part of the dielectric functions $\epsilon_2(\omega)$ for the PbX systems. The p-p transitions are dominant for all systems and the general shape of all functions remains similar. The most prominent change is the decrease of $\epsilon_2^{\max}$. The imaginary part of the dielectric function for PbO is depicted in figure 8.5 in the appendix. For calculation details see chapters 5.7 and 7.3. Picture revised from [24] and also shown in [3].

Figure 2.7 shows a sketch of bond formation and band structure, as well as the imaginary part of the dielectric function $\epsilon_2(\omega)$ for the PbX compounds, where X∈(Te, Se, S). They feature an octahedral (-like) atomic arrangement and the chemical bonds formed between adjacent sites mainly consist of p orbital contributions from Pb and X, respectively, creating a σ bond. As there are again only three valence p electrons per atomic site (averaged over Pb and X), each bond to the six neighbors consists of only half an electron pair, i.e. one electron. Hence the system would be metallic, as the resulting band would be half-filled. Yet, due to the charge transfer between Pb and X[10], a band gap $E_G > 0\,\text{eV}$ opens up. In similar materials, albeit without charge transfer like cubic Sb, metallic behavior with $E_G = 0\,\text{eV}$ is observed instead. Figure 2.7 highlights the intercorrelation among band structure, optical properties and chemical bonding. The shape of the band structure is linked directly to the extent of the overlap of the bond forming orbitals, i.e. the p orbitals. The higher the overlap, the smaller the band gap E_G and the larger the curvature of the bands at the band gap (L-point). Concomitant to the increase of curvature of the bands, there is a decrease in the effective masses m^*[11]. Hence band gap E_G and effective masses can be tuned by changing ET, for example by substituting one atomic site with an isoelectronic element from another period. Trends are also observed in the imaginary part of the dielectric function $\epsilon_2(\omega)$. By increasing ET, i.e. moving from PbTe → PbSe → PbS, the maximum of the imaginary part $\epsilon_2^{\max}$ decreases, while its position moves to slightly higher frequencies ω. This correlation can be motivated by considering Fermi's golden rule in the single particle picture[40]:

$$\epsilon_2(\omega) \propto \frac{1}{\omega^2} \underbrace{\sum_{c,v,\vec{k}} \delta\left(\hbar\omega - (E_c - E_v)\right)}_{JDOS} \underbrace{|\ \langle c|\ \nabla_{\vec{k}}\ |v\rangle\ |^2}_{\text{ME}}, \tag{2.1}$$

with $|c\rangle/|v\rangle$ being conduction/valence band states with the corresponding energies E_c/E_v (see also chapters 5.7 and 7.3). The main constituents of the imaginary part of the dielectric function $\epsilon_2(\omega)$ are the joint density of states (JDOS) and the transition matrix elements (ME). The orbital decomposition of figure 2.7 c) indicates that the p-p transitions are the dominant contributors. Figure 2.8 shows the decomposition of the imaginary part of the dielectric function $\epsilon_2(\omega)$ into the MEs and JDOS. While the MEs decrease significantly going from PbTe → PbS, the JDOS remains virtually identical, with only slight shifts to higher energies due to the increase of the band gap E_G. With ET increasing from PbTe → PbS, the orbital overlap of the bonding p states decreases, which constitute the initial and final states of the transition described by the MEs. Hence, as mentioned above, ET is a suitable descriptor to explain and understand property trends.

So far, ES was kept mostly constant[12] and property transitions were attributed to changes in ET. Within the PbX systems (excluding PbO), this was accommodated by the fact that

[10] $\text{ET}_{\text{PbTe}} = 0.34$, $\text{ET}_{\text{PbSe}} = 0.43$, $\text{ET}_{\text{PbS}} = 0.5$. The formal oxidation state of all PbX compounds is 2, the absolute charge transfer hence is TET = 2 × ET.

[11] Which is another trend incompatible with the lone pair ansatz discussed in the previous section.

[12] At least for the metavalently bonded PbX systems: $\text{ES}_{\text{PbTe}} = 0.80$, $\text{ES}_{\text{PbSe}} = 0.76$, $\text{ES}_{\text{PbS}} = 0.74$, as $\text{ES}_{\text{PbO}} = 0.94$.

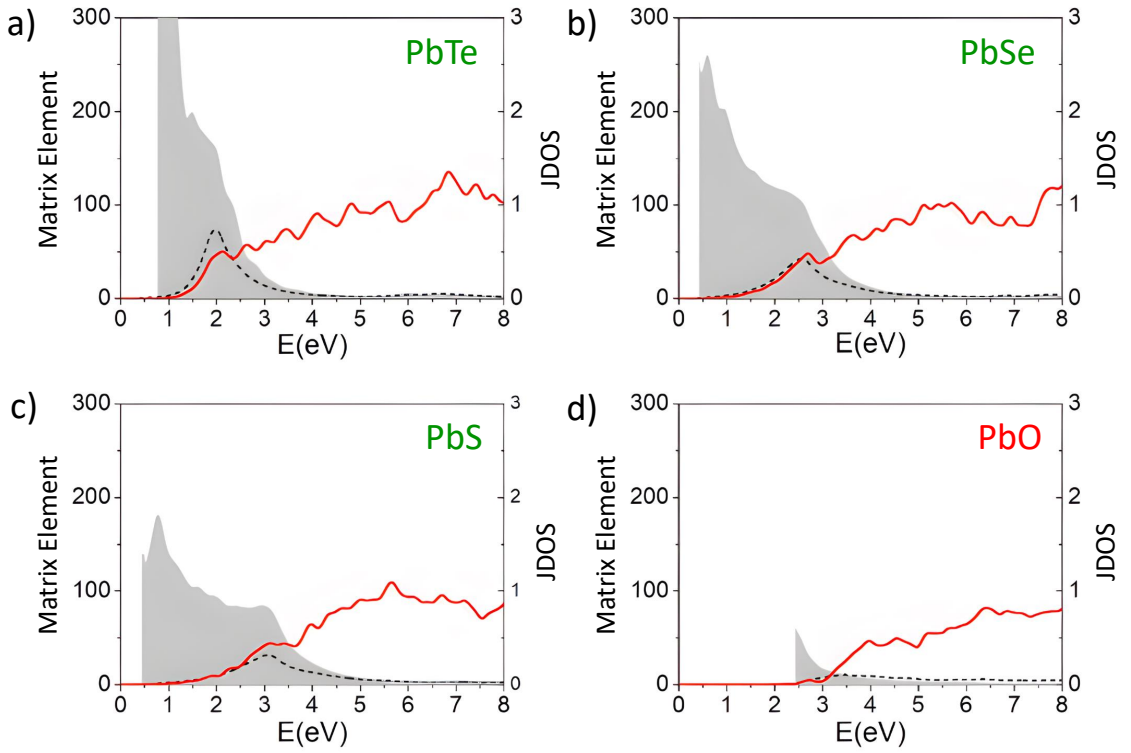


Figure 2.8: Decomposition of the imaginary part of the dielectric function $\epsilon_2(\omega)$ (broken line) into the matrix elements (ME, gray) and joint density of states (JDOS, red) for PbTe, PbSe, PbS and PbO (PbO included as well to illustrate the stark contrast between metavalently bonded system, i.e. PbTe, PbSe and PbS, and the covalent PbO). The (nominal) ME contribution is obtained via $\mathrm{ME}(\omega) = \frac{\epsilon_2(\omega)}{\mathrm{JDOS}(\omega)}$. Upon the transition from PbTe $\rightarrow$ PbO, the shape of the changes in the JDOS are small, hence the decrease of $\epsilon_2^{\max}$ and the general broadening of $\epsilon_2(\omega)$ can be attributed to alterations in the matrix elements, which in turn is caused by changes in the interband transition rates. With ET increasing, going from PbTe $\rightarrow$ PbO ($\mathrm{ET}_{\mathrm{PbO}} = 0.59$), the orbital overlap of the bonding p states decreases, which constitute the initial and final states of the transition described by the MEs. Picture and calculations by Jean-Yves Raty, taken from [3].

they exhibit a simple rock salt structure ($Fm\bar{3}m$) without (or with only little) distortion, while compounds like GeTe exhibit a pronounced distortion of their rock salt structure[41]. In contrast to figure 2.7, figure 2.9 depicts a sketch of bond formation, band structure and the imaginary part of the dielectric function $\epsilon_2(\omega)$ for various degrees of Peierls distortion (PD) in GeTe. The value of PD is calculated as the short to long bond ratio: $\mathrm{PD} = \frac{d_{\mathrm{long}}}{d_{\mathrm{short}}}$. While the link between ES and PD has already been shown in figures 1.14 and 2.4, the impact of PD (and consequently ES) is illustrated in figure 2.9 c). The changes originate from the decrease of the p orbital overlap due to stronger distortion, thereby effectively reducing the transition matrix elements. This shows that not only ET, but also ES, can be used to tailor properties, either by atomic substitution or modification of distortion (see chapter 2.4 for example).
Finally, it is insightful to look at property behavior under the combined change of ES and ET. Phase-change materials (PCMs) are a group of functional materials that feature a distinct

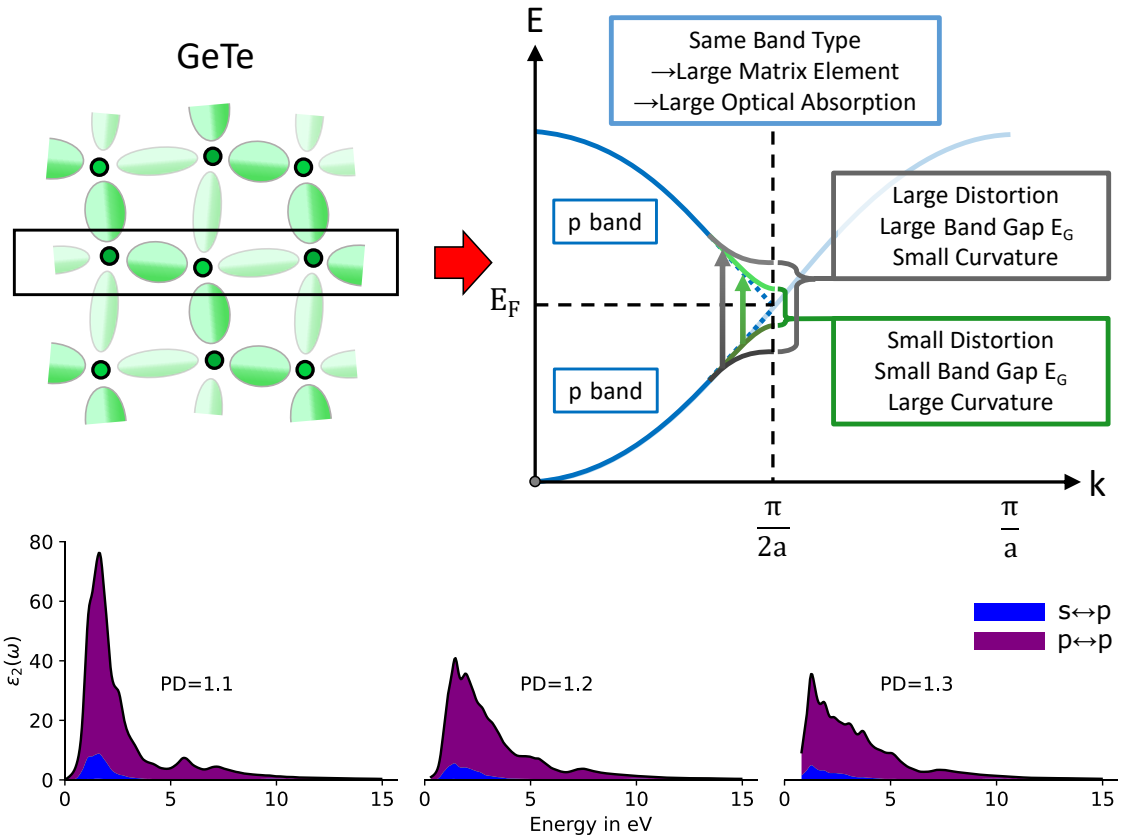


Figure 2.9: a): Sketch of bond formation in GeTe systems. The atomic arrangement of the (001) plane is depicted, along with the atomic 4p orbitals of Ge and the respective 5p orbitals of Te responsible for bond formation. b): Sketch of the band structure in the GeTe systems. As the bonds are only half-filled by one electron or half an electron pair ($\mathrm{ES} = 1$), they would typically form a metallic band (continuous, s-shaped blue band going from $0 \rightarrow \frac{\pi}{a}$). Due to the (rather small) electron transfer and the distortion, a band gap opens, and the Brillouin zone is effectively halved as symmetry is broken. By increasing the Peierls distortion (PD) the band gap E_G opens up further, reducing the curvature of the bands and consequently increasing the effective masses m^*. c): Imaginary part of the dielectric functions $\epsilon_2(\omega)$ for the GeTe systems for different degrees of PD. The value of PD is calculated as the short to long bond ratio: $\mathrm{PD} = \frac{d_\mathrm{long}}{d_\mathrm{short}}$. The p-p transitions are dominant for all systems. The most prominent change is the decrease of the maximum ϵ_2^max. For calculation details see chapters 5.7 and 7.3. Picture revised from [24] and also shown in [3].

contrast between the amorphous and crystalline phase in terms of electrical resistivity and optical reflectivity. For regular materials, these inter-phase contrasts tend to be low[42, 43]. PCMs are mainly employed in optical data storage devices such as Blu-Ray discs and DVDs, but have also been utilized for the Intel Optane 3D XPoint memory. While the optical and electrical property changes are the essential foundation, the time required to switch between the amorphous and crystalline phases, i.e. writing information to a bit, is the limiting factor in order to compete with alternative storage solutions. Generally, the process of crystallization tends to be slower than amorphization, which is why speeding up crystallization promises the largest performance gains[44, 45]. Figure 2.10 shows various stoichiometries of PCMs on

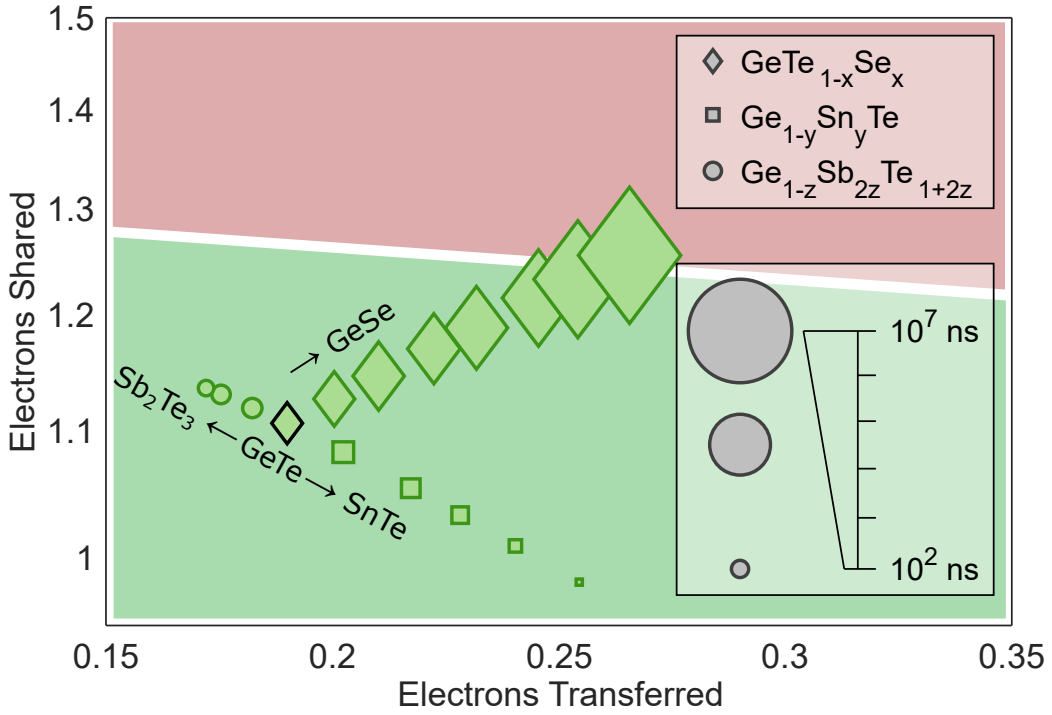


Figure 2.10: Minimum crystallization time τ positioned on the ES/ET map for different stoichiometries of phase-change materials. The marker size indicates the magnitude of τ. A decrease of τ can be observed in the metavalent region (green background), while τ increases towards the covalent border (red background). Crystallization time τ measured by Persch et al.[46]. Intermediate ES/ET values obtained by interpolation. Modified from [46], also shown in [3].

a reduced region of the ES/ET map, while the size of the markers indicate the minimum crystallization time τ. A reduction of τ can be observed by replacing Te with Se along the GeTe - GeSe line, increasing ET, while decreasing ES. Replacing Ge with Sn along the GeTe - SnTe line yields the opposite result, increasing the crystallization time τ along with ES and ET, approaching the covalent region. This trend indicates that the materials with the lowest crystallization time τ are presumably located at the bottom right-hand corner of the metavalent region. While ES and ET both change, it furthermore appears that the decrease in ES is the relevant parameter, as ET increases going from GeTe towards GeSe, where the crystallization time τ increases, as well as towards SnTe, where the crystallization time τ decreases. It should further be noted that crystallization starts by nucleation and subsequent growth in the amorphous phase[13], while the ES/ET values solely reflect the crystalline phase. It is quite peculiar that the trends in crystallization time, which one would expect to be governed by properties of the amorphous phases, are well described by the ES/ET values of the crystal phases.

[13]The amorphous phases are all covalently bonded.

2.2 Halide Perovskites: Third Generation Photovoltaic Materials

The paper "Halide Perovskites: Third Generation Photovoltaic Materials by Metavalent Bonding" was first-authored by Matthias Wuttig and published in *Advanced Functional Materials*[47]. Halide and oxide perovskites are investigated by employing the concepts of metavalent bonding and ES and ET as bonding descriptors and property descriptors to explain and understand the potential of these material classes as photovoltaics. It furthermore illustrates how more complex, i.e. non-binary compounds, can be described using ES and ET. All information is taken from [47] if not stated otherwise.

2.2.1 Properties of Halide Perovskites

The term *perovskite* refers to the mineral *calcium titanate* ($CaTiO_3$), but is nowadays used to refer to compounds of the same structure in general, i.e. materials with the chemical formula of type ABX_3, where A is a group XII, and B a group XI element. A *halide perovskite* is a compound where X belongs to the halogen elements (F, Cl, Br, I), while an *oxide perovskite* employs oxygen atoms instead. Within this study, the undistorted $Pm\bar{3}m$ structure is investigated, while the stable $Pnma$ phase is slightly distorted in comparison.
Oxide perovskites are widely used for various applications, e.g. $PbZr_XTi_{1-X}O_3$ as piezoelectric crystal in sonars[48], $LiNbO_3$ in photonics due to its electro-optic properties[49] and pyroelectric $LiTaO_3$ and $PbNbO_3$ in infrared detectors[50]. The interest in halide perovskites has grown over the last years, as they have been proposed as high-performance photovoltaic (PV) modules, as well as the basis for light-emitting and lasing devices for optoelectronic applications[51–59]. What makes halide perovskites suitable for these kind of applications is that there is a unique property portfolio, including inter alia strong optical absorption and (relatively) large charge carrier mobilities in conjunction with soft crystalline lattices with dynamic disorder[60]. This exclusive combination of properties is neither found in comparable oxide perovskites nor sp^3-bonded semiconductors. This set of properties is often explained as originating from lone pairs of the B atom, e.g. lone pairs formed by the $5s^2$ or $6s^2$ electrons of Sn or Pb, respectively[61]. However, this chapter will argue that metavalent bonding is the mechanism creating the unique property portfolio of halide perovskites. Table 2.9 briefly summarizes desirable properties of PV materials. PV materials should have the best optical absorption in the range (E_G = 1.1 eV to 1.4 eV), while keeping a high charge carrier mobility μ, which is supported by small effective masses m^*. The optical absorption and the effective masses are directly linked to the band structure (see chapters 5.6 and 5.7), which is shown for $BaTiO_3$ and $CsPbI_3$ in figure 2.11 as an exemplary comparison between oxide and halide perovskite. The conduction bands of halide and oxide perovskites show a different structure, e.g. the band minimum being located at different points in $\vec{k}$-space. This originates from different states being dominant in the conduction band. For the oxide perovskites, mainly d states of the B atom contribute (Ti in this case), while for halide perovskites the p states of the B atom (Pb) are dominant. The valence bands (VB) of both

Table 2.9: Overview over the general set of properties desirable for photovoltaic materials. Adapted from [47].

Optical Properties	Direct Band Gap of 1.1 eV to 1.4 eV Sharp and Steep Absorption Edge Low (Non-Radiative) Recombination Rate
Electrical Properties	Charge Carriers with High Mobility μ Low Defect Scattering Rate
Sample Preparation	Simple Manufacturing of High Quality Samples Low Temperature Preparation Route Inexpensive and Abundant Elements

systems are relatively similar, i.e. the VB minimum is located at the R-point, caused by strong contributions of the p orbitals of the respective X atoms, i.e. O for $BaTiO_3$ and I for $CsPbI_3$. One key difference between the valence band structure is the Pb s orbital contribution to the VB maximum for the halide perovskite, which is completely absent in the oxide perovskite, i.e. no Ti s state contribution. This Pb s (lone pair) state has been considered to play a prominent role in granting halide perovskites their favorable property set. A second difference of the VB structure can be observed when considering the dispersion of the respective bands. It is shaped by the overlap of O p states and transition metal (Ti, Zr, Hg in general, Ti in this case) d states for oxide perovskites. For halide perovskites, it is dominated by p states only, i.e. the overlap of halogen and metal p states (I and Pb in this example). The p orbitals also shape the conduction band, resulting in a mirror-like symmetry of conduction and valence bands in halide perovskites. This mirroring causes the hole and electron effect masses to be of comparable size for halide perovskites, while they differ more in oxide perovskites, as depicted in figure 2.12 as a function of ET. The effective masses m^* of the charge carriers near the Fermi level are a paramount property of PV materials. The effective masses are directly linked to the curvature of the bands and can be calculated via $m^* = \hbar^2 \left[d^2E/dk^2\right]^{-1}$, where d^2E/dk^2 denotes the curvature of the bands. It can be estimated by fitting a parabola to the conduction band minimum (electron effective masses m_e^*) or the valence band maximum (hole effective masses m_h^*). Hole and electron effective masses are mostly larger than the free electron mass $\{m_{ox,h}^* \,/\, m_{ox,e}^*\} > m_e$. This implies low mobilities, detrimental for opto-electronic applications. The opposite holds true for halide effective masses, for which the effective masses are always smaller than the free electron mass $\{m_{hal,h}^* \,/\, m_{hal,e}^*\} < m_e$. Furthermore, they follow a chemical trend, i.e. by successively substituting the halide X atom from I $\rightarrow$ Br $\rightarrow$ Cl $\rightarrow$ F, ET increases and the effective masses increase up to a factor of about 8 (see figure 2.12, right panel). This trend is not observed for the oxide perovskites. So far, the role of the lone pair states regarding low effective masses in halide perovskites remains uncertain. Table 2.10 shows the contribution of the s states for a larger variety of halide perovskites as well as the respective effective masses m^*. The data shows that both types of effective masses increase

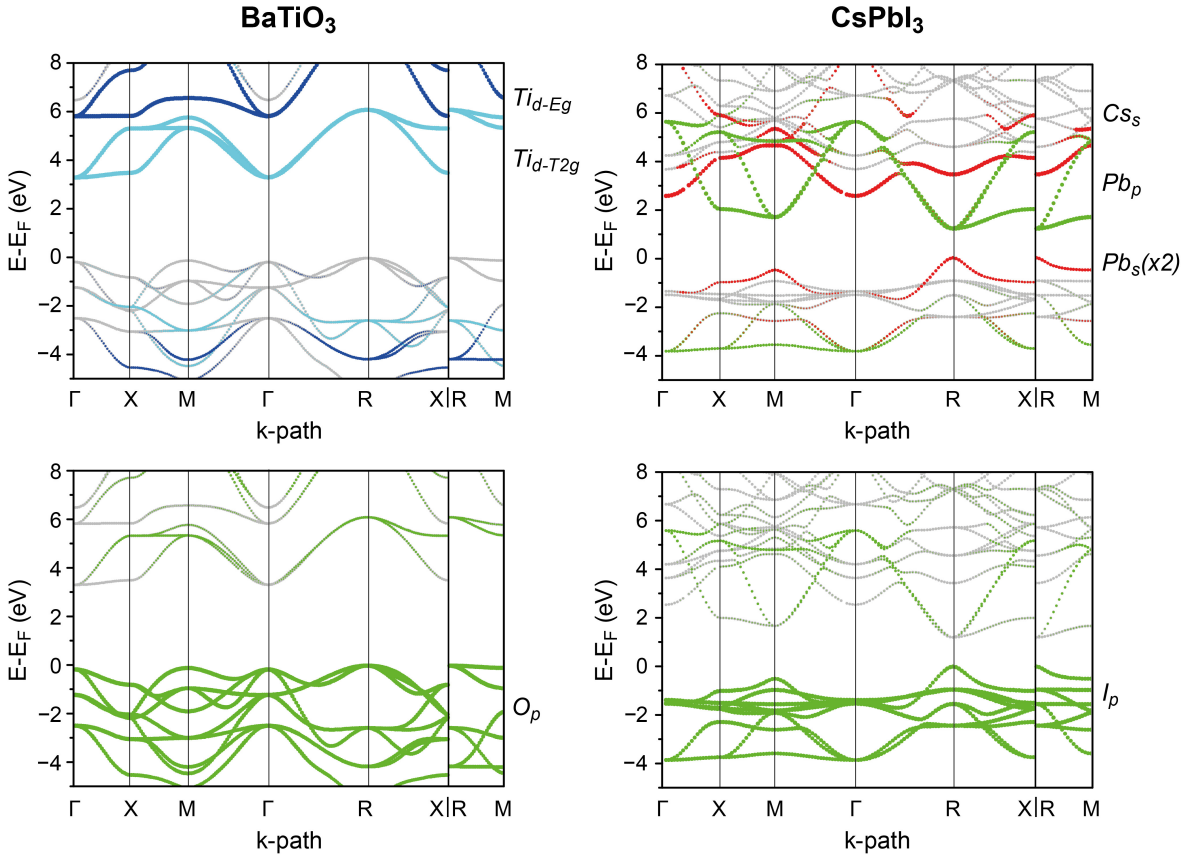


Figure 2.11: Orbital-resolved band structure of $BaTiO_3$ (left) and $CsPbI_3$ (right). The Fermi energy E_F is set to 0 eV (see chapter 5.6). The s states are depicted in red, p states in green and d states in blue (navy blue for the E_g and cyan for the T_{2g} states, respectively). The symbol size reflects the contribution of the respective state to a band at a given $\vec{k}$-point. The top panel of $BaTiO_3$ only shows the Ti T_{2g} states (subsuming the Ti d_{xy}, d_{yz} and d_{xz} orbitals) and the Ti E_g states (subsuming the Ti $d_{x^2-y^2}$ and the d_{z^2} orbitals), while the bottom panel depicts the O p states. The top panel for $CsPbI_3$ shows the Cs s states, as well as the Pb s and p states. The marker size of the Pb s states has been doubled to improve visibility. The bottom panel depicts the I p states only. All non-highlighted states are shown in gray. Figure and calculations (using the SCAN functional[62] without spin-orbit coupling) made by Jean-Yves Raty. Also shown in [47].

by substituting the X atom from I → Br → Cl → F along with the s state contributions[14]. Assuming the s states do form lone pairs responsible for the low effective masses, the increase in s state contribution (and thus the suggested lone pairs) should decrease the effective masses, while the opposite is observed. This finding hence speaks against the lone pair picture.
Besides the effective masses, the optical absorption is another deciding property of photovoltaic materials. Decomposing the imaginary part of the dielectric function into the orbital contributions (see chapter 5.7), as shown in figure 2.13 for $BaTiO_3$ and $CsPbI_3$, reveals that the oxide perovskites are dominated by oxygen p state → metal d state transitions.

[14] For $CsPbF_3$ this trend holds true for the s state contribution to the valence band maximum, while a small decrease of 2% points is observed for the s state contribution to the DOS. For $CsSnF_3$, both s state contributions decrease again to about the level of $CsSnBr_3$, while the effective masses further increase.

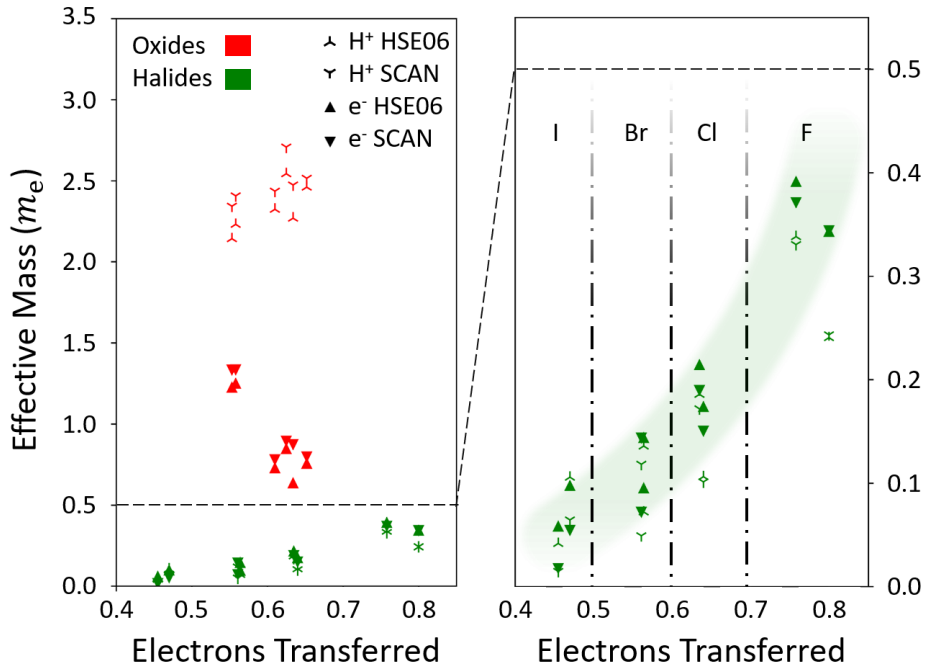


Figure 2.12: Effective masses m^* for halide (green) and oxide (red) perovskites. Hole effective masses are depicted as three-pointed stars, while electron effective masses are denoted as triangles. Different functionals were used to estimate the functionals impact on the results. Upward facing symbols are used for the HSE06 functional including spin-orbit coupling[63], while downward facing ones are utilized for the SCAN functional[62]. The obtained results do not depend strongly on the choice of functional. The oxide perovskites shown include $SrTiO_3$, $SrZrO_3$, $SrHfO_3$, $BaTiO_3$, $BaZrO_3$ and $BaHfO_3$. The halide perovskites shown include $CsPbI_3$, $CsSnI_3$, $CsPbBr_3$, $CsSnBr_3$, $CsPbCl_3$, $CsSnCl_3$, $CsPbF_3$ and $CsSnF_3$. The panel on the right depicts a zoomed-in version of the panel on the left for the halide perovskites, the X halide atom is denoted at the top. The hole and electron effective masses of halide perovskites are relatively similar, while they differ more for oxide perovskites. Both types of effective masses of halide perovskites furthermore increase smoothly with the ET value of the B atom, which is not observed for oxide perovskites. Effective mass calculations performed by Jean-Yves Raty. Taken from [47].

This is contrasted by the halide perovskites, which are almost exclusively governed by p state → p state transitions (I p state → Pb p state in this case)[15]. The total contribution of the p → p-transitions to the absorption $< 4\,\text{eV}$ amount to about 92%, while the s → p-transitions only account for about 8%[16]. The increase in B atom (Sn or Pb) s state contribution to the valence band maximum is accompanied by a decrease in the maximum value of the optical absorption, i.e. changing the halide X atom from I → Br → Cl → F (see table 2.10). Hence the s states (and potential lone pairs) seem to not only be detrimental regarding the effective masses m^*, but also regarding the optical properties.

[15] Figure 2.13 by itself does not specify the direction of the transition, i.e. whether a transition occurs from e.g. s → p state or from p → s state. Only in conjunction with figure 2.11 can such statements be made.

[16] This p-p domination is unusual at first glance, as the selection rules from atomic physics mandate a change of angular momentum of 1 for direct transitions, which is not fulfilled for the p-p transitions in halide perovskites. However, the atomic selection rules are not universally applicable in solids, where the symmetry of the wave function of initial and final state must be considered.

Table 2.10: Percentage of s state contribution to the density of states from $-5.5\,\mathrm{eV}$ to the Fermi energy E_F and at the valence band maximum, hole and electron effective masses m_h^* /m_e^*, band gap E_G, optical dielectric constant ϵ_∞ and height of the absorption maximum for various halide perovskites. Band gap E_G and effective masses m^* computed using the HSE06 functional[63] including spin-orbit coupling, remaining quantities computed using the SCAN functional[62]. All computations conducted by Jean-Yves Raty. Taken from [47].

	Contribution s state DOS	Contribution s state VBM	m_h^*	m_e^*	E_G (eV)	ϵ_∞	Max Abs.
$CsSnF_3$	33%	49%	0.24	0.34	1.99	3.45	3.40
$CsSnCl_3$	48%	65%	0.11	0.17	1.00	4.68	3.53
$CsSnBr_3$	35%	51%	0.07	0.10	0.49	6.23	4.06
$CsSnI_3$	29%	42%	0.04	0.06	0.28	8.59	7.26
$CsPbF_3$	20%	31%	0.34	0.39	2.94	2.88	2.87
$CsPbCl_3$	22%	26%	0.19	0.21	1.71	3.10	2.98
$CsPbBr_3$	16%	20%	0.14	0.14	1.13	4.38	4.28
$CsPbI_3$	12%	14%	0.11	0.10	0.76	5.56	7.59

2.2.2 Quantum Chemical Analysis

Using the QTAIM methods already discussed in the previous chapters offers a different ansatz to explore the origins of the low effective masses and high optical absorption of the halide perovskites. Figure 2.14 shows the halide and oxide perovskites on the ES/ET map. The A-X and B-X bonds are positioned separately, as their ES and ET values differ substantially, rendering averaging unjustified. This indicates that two different types of chemical bonding exist within perovskites. By separating the A-X and B-X bonds, the origin of the characteristic properties can be attributed to either of these bonds. The properties and ES/ET values are listed in tables 2.11 and 2.12 for the oxide and halide perovskites, respectively.

Figure 2.14 indicates that the A-X bond is of ionic character for both halide and oxide perovskites. This is corroborated by the value of the elevated Born effective charge $Z_+^* \approx 1$, which is typical for ionic bonds. The B-X (B-O) bond of the oxide perovskites is located between the ionic and covalent regions and can be considered iono-covalent, as considerable electron transfer and sharing occurs. The B-X bonds of the halide perovskites are located at the border triangle of metavalent, covalent and ionic regions. The number of electrons transferred is comparable to the B-X bonds of the oxide perovskites, but the ES is considerably lower. Additionally, halogen atoms usually show a coordination number of 1, as their valence shell is missing only one electron. The halogens of the halide perovskites (X atoms), however, exhibit an effective coordination number of 2 (see table 2.12), leading to a charge deprived A-X bond. The halide A-X bonds can hence be considered 2-center - 1-electron (2c-1e) bonds, which is incompatible with the 2-center - 2-electron (2c-2e) bond used to describe classical covalent bonds. This configuration is more akin to the mechanism behind metavalent bonding. That notion is further underlined by the number of electrons shared of about ES ≈ 1, which

Table 2.11: Properties of oxide perovskites ($Pm\bar{3}m$ structure) including Born effective charge Z^*, elevated Born effective charge $Z^*_+ = \frac{Z^*}{\text{Formal Oxidation State}}$, effective coordination number (ECoN), optical dielectric constant ϵ_∞, band gap E_G, electrons transferred (ET) and electrons shared (ES). Double asterisks (**) denote values obtained with the HSE06 functional with spin-orbit coupling[63]. Z^*, Z^*_+, ϵ_∞ and E_G values obtained by Jean-Yves Raty. Taken from [47].

Material	Element	Z^*	Z^*_+	ECoN	ϵ_∞	E_G (eV)	ET	ES
$CaHfO_3$	Ca	2.53	1.26	13.92	4.07	4.26	Ca-O	
	Hf	5.76	1.44	6			0.817	0.152
	O	-4.61	-2.30	2.04			Hf-O	
		-1.81	-0.90				0.651	0.982
		-1.81	-0.90					
$CaTiO_3$	Ca	2.61	1.31	13.92	6.56	2.03	Ca-O	
	Ti	6.82	1.71	6			0.8	0.168
	O	-5.54	-2.77	2.04			Ti-O	
		-1.98	-0.99				0.561	1.081
		-1.98	-0.99					
$CaZrO_3$	Ca	2.61	1.30	13.92	4.20	3.91	Ca-O	
	Zr	5.86	1.47	6			0.866	0.146
	O	-4.79	-2.40	2.039			Zr-O	
		-1.80	-0.90				0.642	1.024
		-1.80	-0.90					
$SrTiO_3$	Sr	2.50	1.25	13.92	5.66	3.74**	Sr-O	
	Ti	7.12	1.78	6			0.799	0.212
	O	-5.65	-2.82	2.04			Ti-O	
		-1.99	-0.99				0.558	1.072
		-1.99	-0.99					
$SrZrO_3$	Sr	2.52	1.26	13.92	4.19	4.88**	Sr-O	
	Zr	5.93	1.48	6			0.808	0.184
	O	-4.78	-2.39	2.04			Zr-O	
		-1.84	-0.92				0.625	1.020
		-1.84	-0.92					
$SrHfO_3$	Sr	2.54	1.27	13.92	4.03	5.30**	Sr-O	
	Hf	5.73	1.43	6			0.802	0.194
	O	-4.55	-2.28	2.04			Hf-O	
		-1.85	-0.93				0.652	0.972
		-1.85	-0.93					
$BaTiO_3$	Ba	2.78	1.39	13.92	6.00	3.66**	Ba-O	
	Ti	7.17	1.79	6			0.776	0.288
	O	-5.57	-2.79	2.04			Ti-O	
		-2.12	-1.06				0.553	1.060
		-2.12	-1.06					
$BaZrO_3$	Ba	2.70	1.35	13.92	4.49	4.66**	Ba-O	
	Zr	6.03	1.51	6			0.757	0.256
	O	-4.76	-2.38	2.04			Zr-O	
		-1.84	-0.92				0.634	1.010
		-1.84	-0.92					
$BaHfO_3$	Ba	2.71	1.36	13.92	4.34	5.03**	Ba-O	
	Hf	5.79	1.45	6			0.800	0.262
	O	-4.53	-2.26	2.04			Hf-O	
		-2.01	-1.00				0.610	0.996
		-2.01	-1.00					

Table 2.12: Properties of halide perovskites ($Pm\bar{3}m$ structure) including Born effective charge Z^*, elevated Born effective charge $Z^*_+ = \frac{Z^*}{\text{Formal Oxidation State}}$, effective coordination number (ECoN), optical dielectric constant ϵ_∞, band gap E_G, electrons transferred (ET) and electrons shared (ES). Z^*_+ is not stated for Sn and Pb, as the formal oxidation state is not clearly defined. Double asterisks (**) denote values obtained with the HSE06 functional with spin-orbit coupling[63]. Z^*, Z^*_+, ϵ_∞ and E_G values obtained by Jean-Yves Raty. Taken from [47].

Material	Element	Z^*	Z^*_+	ECoN	ϵ_∞	E_G (eV)	ET	ES
$CsSnCl_3$	Cs	1.30	1.30	13.92	3.53	1.00**	Cs-Cl	
	Sn	4.63		6			0.895	0.096
	Cl	-4.86	4.86	2.04			Sn-Cl	
		-0.58	-0.58				0.640	0.740
		-0.58	-0.58					
$CsPbCl_3$	Cs	1.29	1.29	13.92	3.10	1.71**	Cs-Cl	
	Pb	4.01		6			0.900	0.087
	Cl	-3.86	-3.86	2.04			Pb-Cl	
		-0.71	-0.71				0.635	0.710
		-0.71	-0.71					
$CsSnBr_3$	Cs	1.31	1.31	13.92	6.23	0.49**	Cs-Br	
	Sn	5.32	1.33	6			0.983	0.094
	Br	-5.33	-5.33	2.04			Sn-Br	
		-0.60	-0.60				0.561	0.788
		-0.60	-0.60					
$CsPbBr_3$	Cs	1.30	1.30	13.92	4.38	1.13**	Cs-Br	
	Pb	4.36	1.09	6			0.983	0.094
	Br	-4.25	-4.25	2.04			Pb-Br	
		-0.72	-0.72				0.564	0.741
		-0.72	-0.72					
$CsSnI_3$	Cs	1.34	1.34	13.92	8.59	0.28**	Cs-I	
	Sn	6.13	1.53	6			0.870	0.088
	I	-6.09	-6.09	2.04			Sn-I	
		-0.69	-0.69				0.455	0.844
		-0.69	-0.69					
$CsPbI_3$	Cs	1.34	1.34	13.92	5.56	0.76**	Cs-I	
	Pb	5.08	1.27	6			0.880	0.078
	I	-4.93	-4.93	2.04			Pb-I	
		-0.75	-0.75				0.470	0.794
		-0.75	-0.75					

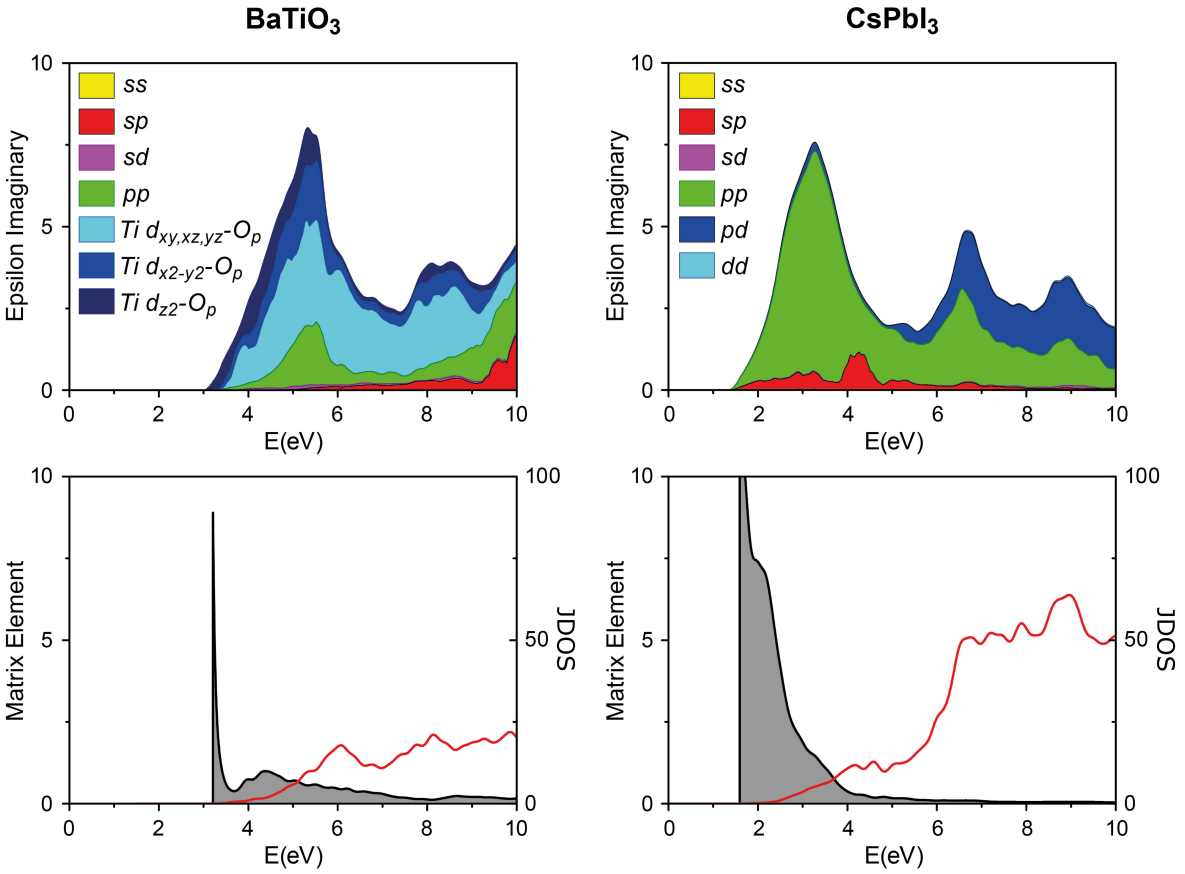


Figure 2.13: Orbital-resolved imaginary part of the dielectric function $\epsilon_2(\omega)$ (top), joint density of states (JDOS) in 1/eV and effective matrix elements $\mathrm{ME}(\omega)$ in eV (bottom) of $BaTiO_3$ (left) and $CsPbI_3$ (right). The effective matrix elements $\mathrm{ME}(\omega)$ are obtained by dividing the imaginary part of the dielectric function $\epsilon_2(\omega)$ by the JDOS: $\mathrm{ME}(\omega) = \frac{\epsilon_2(\omega)}{\mathrm{JDOS}(\omega)}$. If no element is stated, the transition encompasses contributions from all atomic sites. The absorption maximum of the halide perovskite $CsPbI_3$ is located at around $3\,\mathrm{eV}$ and dominated by p↔p transitions. It benefits from a larger matrix element $\mathrm{ME}(3\,\mathrm{eV}) \approx 2$ compared to the oxide perovskite $BaTiO_3$, for which the absorption maximum is located at about $5\,\mathrm{eV}$ with a matrix element of about $\mathrm{ME}(5\,\mathrm{eV}) \approx 1$. Calculations conducted and plot made by Jean-Yves Raty using the SCAN functional. Taken from [47].

is comparable to the values found for the metavalently bonded lead chalcogenides (PbX systems, see also chapters 2.1.1 and 2.5.2). It can hence be assumed that the B-X bonds in halide perovskites are of metavalent nature.

Tables 11.10, 11.11 and 11.12 in the appendix show the DAFH analysis of $CsPbI_3$, $CsPbBr_3$ and $CsPbCl_3$ respectively. The DAFH calculations reveal that the p orbitals of the B atoms significantly overlap with the basin of the X atom and vice versa. The p orbitals provide the dominant contribution to the ES values of the B-X bonds, in line with the characteristics of a metavalent bond. The s orbitals, in contrast, are almost negligible, with virtually insignificant contributions to the ES value, as they overlap at most 1.5% with neighboring basins, albeit they are much more populated (about 1.85 e^-) than the p orbitals (0.3-1.7 e^-).

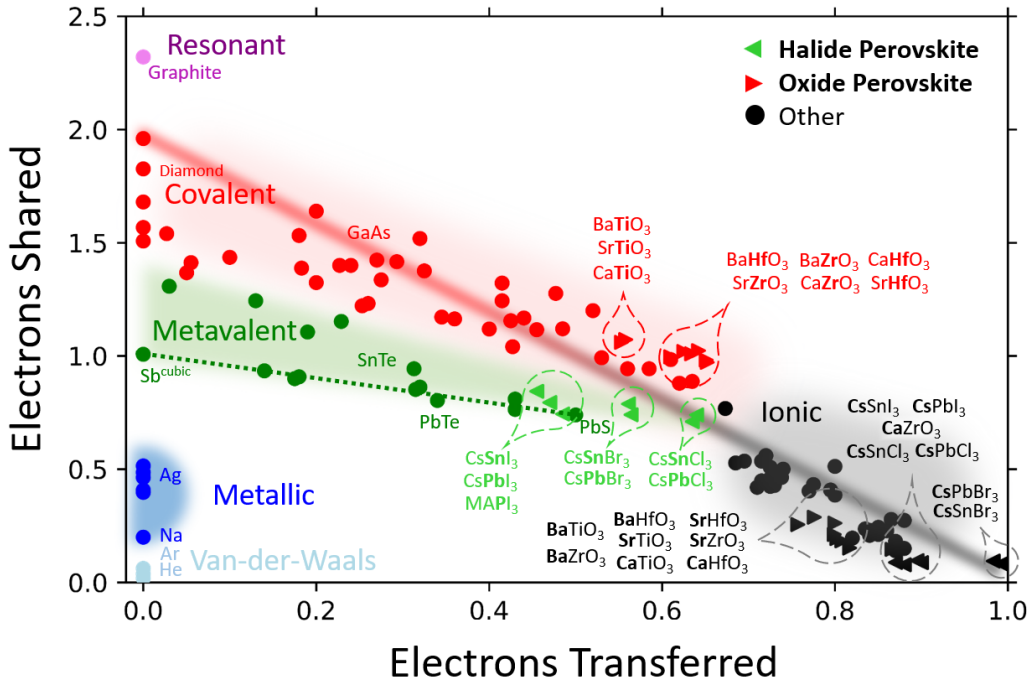


Figure 2.14: ES/ET map showing the position of the A-X and B-X bonds of various halide (left-pointing triangles) and oxide perovskites (right-pointing triangles). For the labels of the halide perovskites, the element in bold font signifies the bond partner considered. The A-X bond is located in the ionic region for both halide and oxide perovskites. The B-X bonds of the oxide perovskites are positioned in the iono-covalent area, while the B-X bond of the halide perovskites are located in or close to the metavalent region. The broken green line denotes the position of metavalent compounds with perfect (cubic) octahedral arrangement. MAPI refers to $CH_3NH_3PbI_3$, which is an organometallic halide perovskites. Modified from [47].

The orbital-resolved band structures shown in figure 2.11 and dielectric functions in figure 2.13 already implied that the p-p overlap and transitions are crucial for the unique properties of halide perovskites. The DAFH data in conjunction with figure 2.14 underlines this assumption and reveals that the metavalently bonded B-X bond provides this p-p overlap, while the contributions of the lone pairs on the B and X sites are minor.

Figure 2.15 depicts a sketch of bond formation in $CsPbI_3$. As shown above, the p orbitals are the most relevant contributors, forming a σ bond. It contains about 0.8 electrons (ES $\approx$ 0.8, see table 11.10), a typical value for a metavalent bond. Roughly half-filled, this σ bond would usually form a metallic band, but due to the non-zero charge transfer, a small band gap opens. Halide perovskites are therefore "incipient metals", i.e. semiconductors on the verge of becoming a metal[5]. The incipient metals described by Wuttig et al.[5] also share a multitude of properties with halide perovskites, e.g. a high Born effective charge Z^*, an effective coordination number ECoN violating the 8-N octet rule, elevated

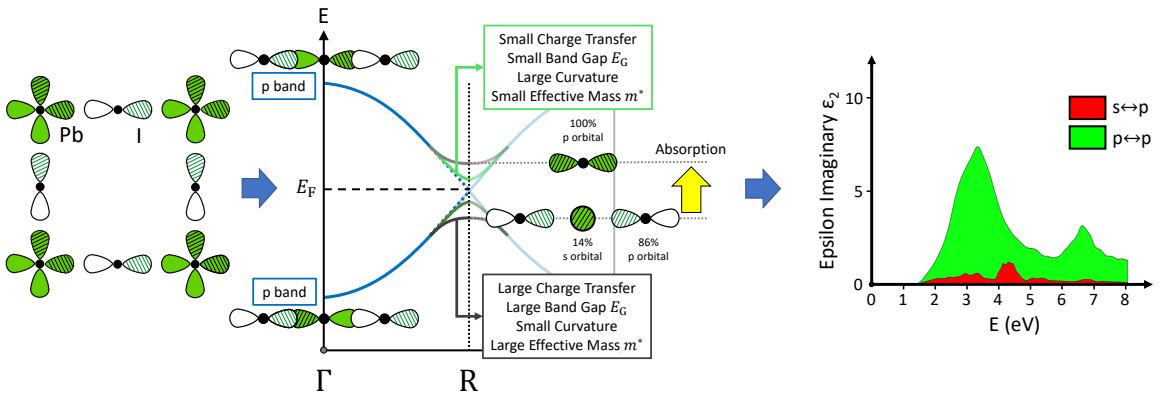


Figure 2.15: Sketch of bond formation and resulting band structure and optical properties of $CsPbI_3$. Left: 1D depiction of the Pb and I p orbitals forming bonds. The resulting σ bonds are half-filled, i.e. ES ≈ 1. Middle: A metallic band would form due to the half-filled σ bonds (blue band), but due to the electron transfer a small band gap opens (green and gray bands). At the Γ-point the σ bonds form a bonding and an antibonding state for the valence and conduction band respectively, equally incorporating Pb and I p orbitals. At the R-point this balance shifts towards the valence band consisting of 14% Pb s state and 86% I p state, while the conduction band now is 100% Pb p state (see figure 2.11). Dielectric function calculated by Jean-Yves Raty. Modified from [47].

optical dielectric constants ϵ_∞ and a soft crystalline lattice[17], which are related to their applicability in photovoltaics[64–66]. Figure 2.15 also motivates why an increase in band gap E_G increases the effective masses m^*, as confirmed by table 2.10. So far, only the perfect cubic structures ($Pm\bar{3}m$) have been discussed, while the stable structures are slightly distorted. It should therefore be noted that this distortion has to remain moderate in order to preserve metavalency. As has been shown before (e.g. figure 2.4), increasingly pronounced distortions can shift the bonding type towards covalency, altering the properties of the compound significantly[67]. For halide perovskites, the distortions are not strong enough to fundamentally change the properties, as discussed in detail in [47].

To summarize, the data presented within this chapter shows that the unique property portfolio of halide perovskites originates from the presence of metavalent bonding in the B-X bond. The p states near the Fermi energy E_F form half-filled σ bonds, while a band gap is opened due to charge transfer (and possible distortions in the stable, non-cubic phase). This configuration leads to small effective masses m^* and strong optical absorption, as is desired for PV applications. Regarding the objective of this thesis, it has been shown that metavalent bonding provides a powerful framework to explain, describe and understand the origins of certain properties, along with ES and ET to tailor them. To tune the band gap E_G, for example, strain engineering can be employed or the A atom of the ABX_3 structure can be modified.

[17] Considering the B-X bond, where applicable.

2.3 Investigation and Classification of Layered Systems

This section is based on the master thesis of Alexander Kiehn: "Analysis of Property Trends in Two-Dimensional Materials via Quantum-Chemical Bonding Descriptors"[68]. Two-dimensional materials refer in this case to layered materials, e.g. graphene. The characteristic feature of layered materials[18] is that the strong bonds, i.e. covalent, metallic, ionic (and metavalent), are confined to planes. These planes can vary in thickness, but for many compounds they only consist of a couple atomic layers. The planes in turn are coupled via the comparatively weak van der Waals (vdW) force (see figure 2.16). The interlayer interaction of layered chalcogenides, such as Bi_2Te_3, have been suggested to be stronger than that of regular vdW bonding, but still weaker than that of any of the strong bonds[69–73]. This interaction will be coined "pseudo-vdW bonding", to stress its phenomenological kinship to the standard vdW interaction. This section focuses on the description and identification of layered materials and their properties using QTAIM and DFT methods. The purpose of this section is to illustrate the applicability of bonding descriptors beyond the ES/ET map already presented. As the structure of layered compounds is more complex, a single ES value is no longer sufficient to describe the chemical environment of an atomic site (similar to the case of the halide perovskites in section 2.2). This section aims to provide another perspective on how ES can be employed to characterize material classes.
All DFT and QTAIM calculations have been performed using *Project DIO* (see chapter 7.5), either by Alexander Kiehn (ES) or the author of this thesis (translation energy ΔE_{Shift}).

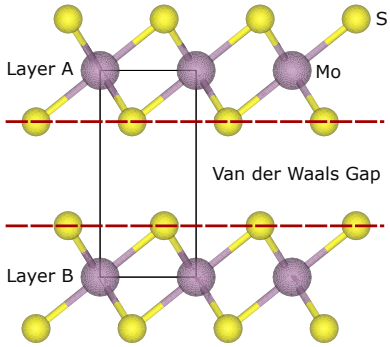


Figure 2.16: Sketch of the layered compound MoS_2. Strong bonds are employed within the layers, while the layers themselves adhere to each other employing van der Waals forces. The black rectangle denotes the unit cell.

2.3.1 Identification of Layered Materials

For the most prominent layered materials, like graphene, the weak vdW interaction has been investigated in great detail and can be considered confirmed. For many other compounds the existence of vdW interaction is often merely inferred, based on the layered structure and gap distance[74, 75]. To tackle a larger number of compounds, an algorithm was developed within the scope of this dissertation that predicts layeredness from the structure (unit cell) of a compound. This tool is called *Layeredness Estimator Algorithm* (*LEA*, pronounced like *layer*) and is based on the works of Stevanović et al.[76]. It is illustrated in figure 2.17.

[18] The term layered and two-dimensional will be used synonymously in this context.

The criterion of Stevanović et al. in itself measures the interlayer distance $d_{2D} = d_{\text{interlayer}} - d_0 - 0.3\,\text{Å}$, where d_0 is the nearest neighbor distance of an atom at the edge of the layer (to another atom within the same layer), and $d_{\text{interlayer}}$ the distance to the closest atom located in the neighboring layer (see figure 8.6 in the appendix). A compound is considered layered if $d_{2D} > 0$. *LEA* calculates the coordination number (CoN) by counting all atomic sites around a reference atom with distance $d_0 + 0.3\,\text{Å}$, as well as the effective coordination number (ECoN, see section 7.4) of that reference atom. Cutting planes are placed along probable layer gaps and the atomic sites beyond that plane are removed. If this process does not change the ECoN nor the CoN, the gap is a (true) vdW gap. For layered compounds that employ pseudo-vdW interactions, cuts seem to keep only the CoN constant, while the ECoN changes (as it is more susceptible to modifications in the atomic environment, while the CoN changes in discreet steps). For materials that are not layered, all cutting planes alter both the ECoN and the CoN.

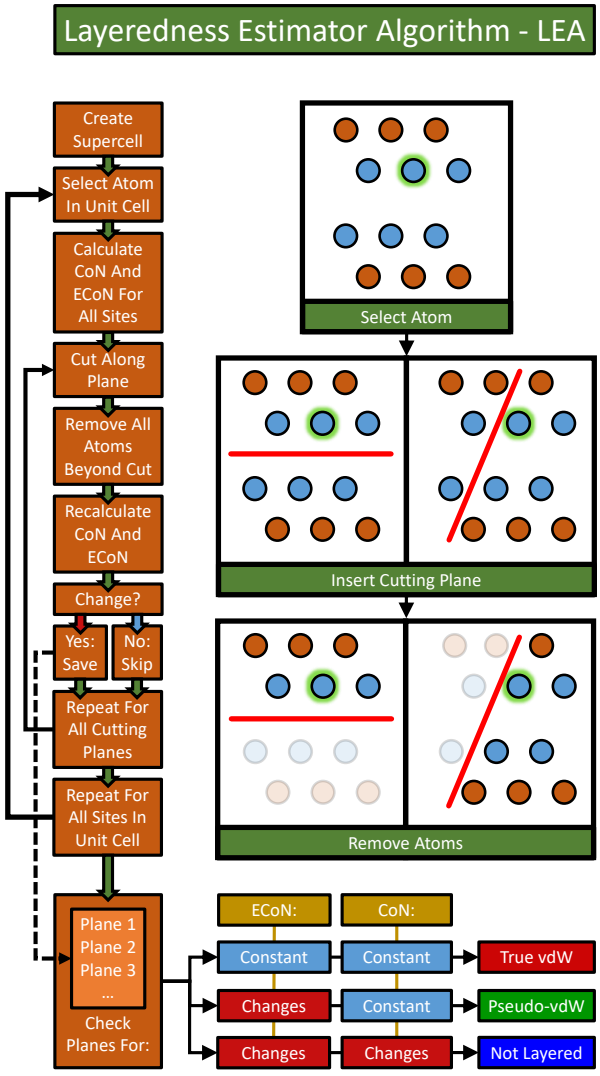


Figure 2.17: Flowchart illustrating the 2D/Layeredness detection algorithm. Cutting planes are chosen by sampling the spherical angles in sufficiently small steps around each atom. Additionally, a plane is fitted through the nearest neighbors of the selected atom to estimate a likely orientation of a cutting plane. Due to the finite spatial extent of the supercell, small deviations from the exact cutting plane orientation are acceptable.

In addition to the 2D criterion by Stevanović et al., alternative approaches were conceptualized amongst others by Ashton et al.[75], Cheon et al.[77] and Mounet et al.[78]. All these approaches are structure-based and differ in terms of which distances they consider and how the 2D threshold is defined. As will be shown hereafter, the approach by Stevanović et al.

Table 2.13: Comparison of the 2D criteria by Ashton et al.[75], Cheon et al.[77], Mounet et al.[78] and Stevanović et al.[76]. *False negative* refers to the case where a layered compound is incorrectly assigned to be 3D (not layered), *false positive* means a compound is incorrectly described as layered, while it is 3D (not layered). Criteria of Ashton, Cheon and Mounet assessed by Alexander Kiehn, Stevanović criterion tested using LEA. Material set A compiled by Alexander Kiehn, set B by Sebastian Gruner and Christian Stenz. Modified from [68].

Criterion by	Set A			Set B		
	False Negative	False Positive	Total Errors	False Negative	False Positive	Total Errors
Ashton et al.	1	4	5	32	1	33
Cheon et al.	3	2	5	3	3	6
Mounet et al.	28	0	28	7	2	9
Stevanović et al.	2	0	2	5	0	5

proves to be the most reliable, hence the exact working principle of the other criteria are omitted here, but are explained in detail in their respective publication as well as [68].
Two datasets have been compiled to evaluate the four criteria. Set A (see appendix tables 11.14, 11.15) contains about 100 general compounds, while set B consists of about 110 naturally occurring chalcogenide minerals (see appendix table 11.16)[19]. Table 2.13 summarizes the results. The method of Stevanović et al. clearly outperforms the alternative approaches and will hence be the one used in the further analysis. Within the datasets used for this thesis, the algorithm achieved a success rate of 97.1% in sum total with no false positives regarding whether a material is layered or non-layered. By employing the additional ECoN criterion to distinguish between true and pseudo-vdW systems, only 3 systems are misclassified as pseudo-vdW out of 36 true vdW systems (91.7% success rate). The misclassified compounds (mostly the minerals) often feature an intercalation layer, i.e. another row atoms that is inserted between the layer gap, making it unclear whether they should be considered layered, not-layered or a pseudo-vdW system. In any case, the reduced distances across the gap render these materials not-layered for the employed algorithmic approach.

[19] Set A compiled (including assessment of layeredness) by Alexander Kiehn, set B by Sebastian Gruner and Christian Stenz, RWTH Aachen University.

2.3.2 Characterization of Layered Materials

Besides the d_{2D} scale by Stevanović et al., as introduced in section 2.3.1, other parameters have been derived and employed to characterize 2D materials. These will briefly be described in the following.

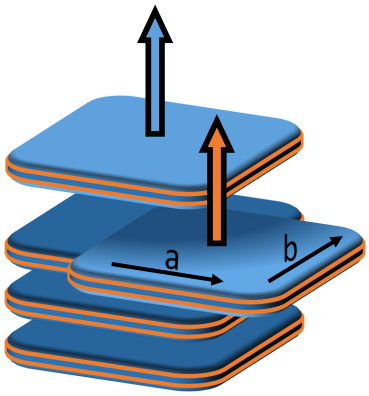


Figure 2.18: Shifted layer.

Translation Energy

In order to assess the strength of the layeredness of a given compound, the energy to laterally shift a layer ΔE_{Shift} can be calculated (see figure 2.18)[70]:

$$\Delta E_{\text{Shift}} = \frac{E_{\text{Shift}}}{N_{\text{UC}}} - \frac{E_{\text{Stable}}}{N_{\text{UC}}}, \tag{2.2}$$

where E_{Shift} and E_{Stable} are the total energy (per unit cell) of the shifted and stable configuration, respectively, and N_{UC} the number of atoms in the unit cell. While expression 2.2 is serviceable to compare structurally identical compounds, it fails when comparing differently coordinated materials. The increase of energy in the shifted configuration $E_{\text{Shift}} > E_{\text{Stable}}$ originates from the unfavorable coordination of atoms bordering the gap of the shifted layer. The number of bordering atoms N_{Gap} differs from the total number of atoms in the unit cell N_{UC}, so normalizing the energies by N_{UC} is not justified in the general case[20]. ΔE_{Shift} is hence better defined as:

$$\Delta E_{\text{Shift}} = \frac{E_{\text{Shift}} - R_{\text{Bulk}} E_{\text{Stable}}}{N_{\text{Gap}}} - \frac{E_{\text{Stable}}}{N_{\text{UC}}}, \quad R_{\text{Bulk}} = \frac{N_{\text{UC}} - N_{\text{Gap}}}{N_{\text{UC}}}, \tag{2.3}$$

where R_{Bulk} is the ratio of atoms within the unit cell that are not bordering the gap. The first term $\frac{E_{\text{Shift}} - R_{\text{Bulk}} E_{\text{Stable}}}{N_{\text{Gap}}}$ is hence the total energy of a gap atom, while ΔE_{Shift} denotes the total energy increase per gap atom caused by the translation. For vdW bonded layers, the energy to shift a layer ΔE_{Shift} should be small as the interaction is weak. If the layers are connected through a stronger interaction however, i.e. pseudo-vdW interaction of even strong bonds like covalent ones, this should be reflected in higher values of ΔE_{Shift}.

By probing an energy surface (within the scope of this thesis, a 20×20 translation grid was used), the maximum translation energy $\Delta E^{\text{Max}}_{\text{Shift}}$ for a given compound can be obtained. Figure 2.19 exemplifies this process for the example of Bi_2Te_3, where $\Delta E^{\text{Max}}_{\text{Shift}} \approx 485\,\text{meV}\,(\text{Atom at Gap})^{-1}$ at a relative translation of (0.33/0.66).

Definition 2.3 furthermore offers the advantage that the number of shifted layers is irrelevant,

[20] For example: If a compound A features a large unit cell with $N^A_{\text{UC}} = 36$, the total energy will be much larger compared to compound B with $N^B_{\text{UC}} = 18$. However, both compounds could require the same energy to shift the layer (as both feature $N_{\text{Gap}} = 6$). In this case the $\Delta E^A_{\text{Shift}}$ would be lower than $\Delta E^B_{\text{Shift}}$, which would be misleading in terms of layeredness.

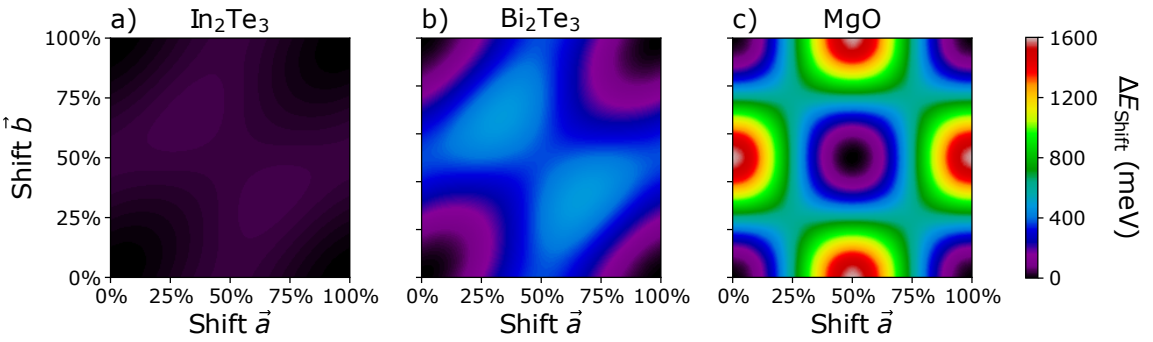


Figure 2.19: Energy surface of the translation energy ΔE_{Shift} of a): vdW bonded In_2Te_3, b): Pseudo-vdw bonded Bi_2Te_3; c): Ionically bonded MgO. The coordinates *Shift* $\vec{a}$ and *Shift* $\vec{b}$ indicate the relative translation of the layer. Calculations obtained with QUANTUM ESPRESSO (see chapter 7.1.1) using the Grimme-D3 van der Waals correction[79]. The grid of energy points contained 20×20 points and was then interpolated.

as long as the shifted layers are connected, i.e. they are neighboring layers. For example, $\Delta E^{\text{Max}}_{\text{Shift}}$ of $HfSe_2$ for 1, 3 and 5 shifted layers are virtually identical:

$$\Delta E^{\text{Max}}_{\text{Shift}} = \begin{cases} 103.22\,\text{meV}\,(\text{Atom at Gap})^{-1} & 1\ \text{Layer} \\ 103.47\,\text{meV}\,(\text{Atom at Gap})^{-1} & 3\ \text{Layer} \\ 103.48\,\text{meV}\,(\text{Atom at Gap})^{-1} & 5\ \text{Layer} \end{cases} \tag{2.4}$$

Interlayer ES

As already explained in the previous chapters, ES is a measure of (covalent) bond strength between two atoms. In the ES/ET maps shown so far, e.g. in figure 2.14, the majority of the materials shown are binary compounds with isotrope atomic arrangement, hence rendering ES and ET values identical for neighbors in all direction and leaving only one choice for ES and ET. The major exception is the perovskites (see chapter 2.2), where for each perovskites two points have been marked on the map. For the layered systems the ES values within the layer (intralayer ES) and between layers (interlayer ES) differ naturally as well. It can furthermore be assumed that the interlayer ES should also be a measure of the interlayer interaction strength, and will be used in conjunction with the maximum translation energy $\Delta E^{\text{Max}}_{\text{Shift}}$ and the d_{2D} scale by Stevanović et al.

2.3.3 Classification of Layered Materials

$\Delta E^{\text{Max}}_{\text{Shift}}$, d_{2D} and $\text{ES}_{\text{Interlayer}}$ for about 100 layered chalcogenides and 20 non-layered compounds have been computed. Figure 2.20 depicts d_{2D} plotted against $\Delta E^{\text{Max}}_{\text{Shift}}$. The resulting map separates the compounds into four groups. The true vdW systems are located at low maximum translation energies $\Delta E^{\text{Max}}_{\text{Shift}} < 200\,\text{meV}\,(\text{Atom at Gap})^{-1}$ and high values

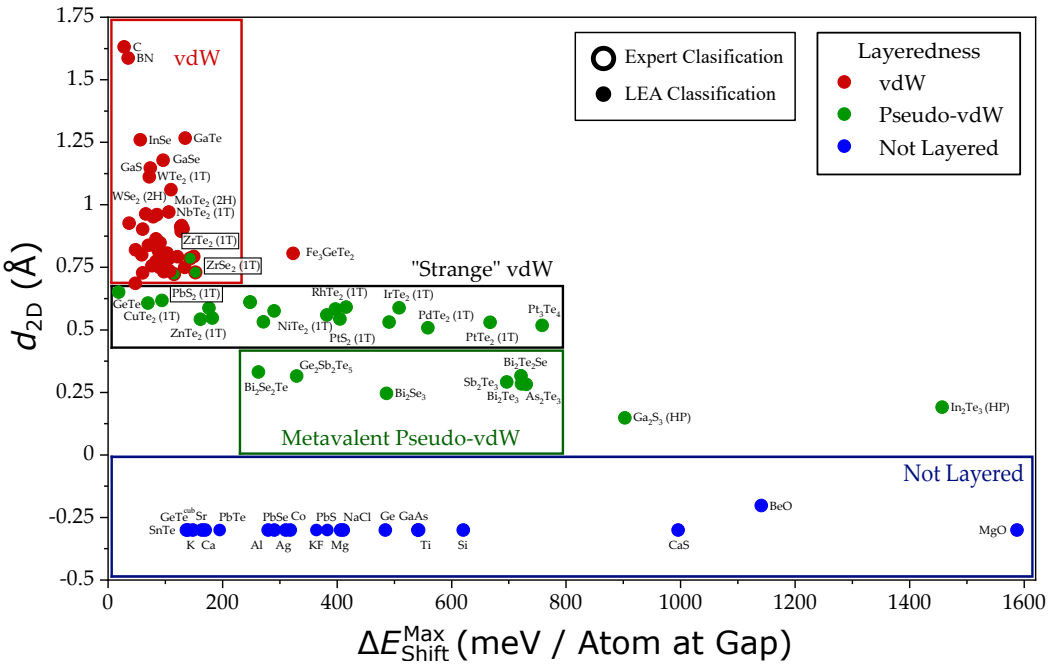


Figure 2.20: Maximum translation energy $\Delta E^{\mathrm{Max}}_{\mathrm{Shift}}$ plotted against the d_{2D} scale by Stevanović et al. (compiled by Alexander Kiehn). True vdW compounds are denoted in red, pseudo-vdW systems in green and non-layered materials in blue. The center color refers to the classification as determined by the algorithm described in section 2.3.1, while the edge color denotes the expert/human classification. Modified from [68].

of $d_{2D} > 0.65\,\text{Å}$. This reflects the expectation that vdW bonded layers will only be weakly coupled and relatively far apart. The non-layered systems are located at negative values of $d_{2D} < 0\,\text{Å}$. However, they cover the full range of maximum translation energies $\Delta E^{\mathrm{Max}}_{\mathrm{Shift}} = 100 - 1600\,\text{meV}\,(\text{Atom at Gap})^{-1}$, which is rather unexpected for a non-layered system. Still, most of the non-layered systems display maximum translation energies $\Delta E^{\mathrm{Max}}_{\mathrm{Shift}} > 250\,\text{meV}\,(\text{Atom at Gap})^{-1}$, clearly separating them from the true vdW compounds. Why compounds like K, Sr, GeTe and SnTe exhibit such low translation energies requires further investigation. It could be possible that this group of materials is susceptible to structural defect formation, which would justify why a dislocation of atoms is energetically relatively inexpensive. The remaining two groups are both located between the true vdW systems and the non-layered systems. Above the non-layered systems, however, within a narrower maximum translation energy range $\Delta E^{\mathrm{Max}}_{\mathrm{Shift}} = 200 - 800\,\text{meV}\,(\text{Atom at Gap})^{-1}$, are the metavalently bonded pseudo-vdw materials. This group mainly consists of sesquichalcogenides[21], such as Bi_2Se_3. The last group is labeled "strange" vdW, as it overlaps with the true vdW systems with regard to the maximum translation energy range $\Delta E^{\mathrm{Max}}_{\mathrm{Shift}}$, but not in terms of d_{2D}. Which strong intralayer bonding type is employed within this group is not clear for all members, but many of them are presumably metallic or semi-metallic[80, 81].

[21] The prefix "sesqui" originates from Latin and translates into "one and a half". In this case, it refers to compounds with a 3:2 ratio of atomic types.

Figure 2.21 depicts the maximum translation energy $\Delta E^{\mathrm{Max}}_{\mathrm{Shift}}$ plotted against $\mathrm{ES}_{\mathrm{Interlayer}}$. This

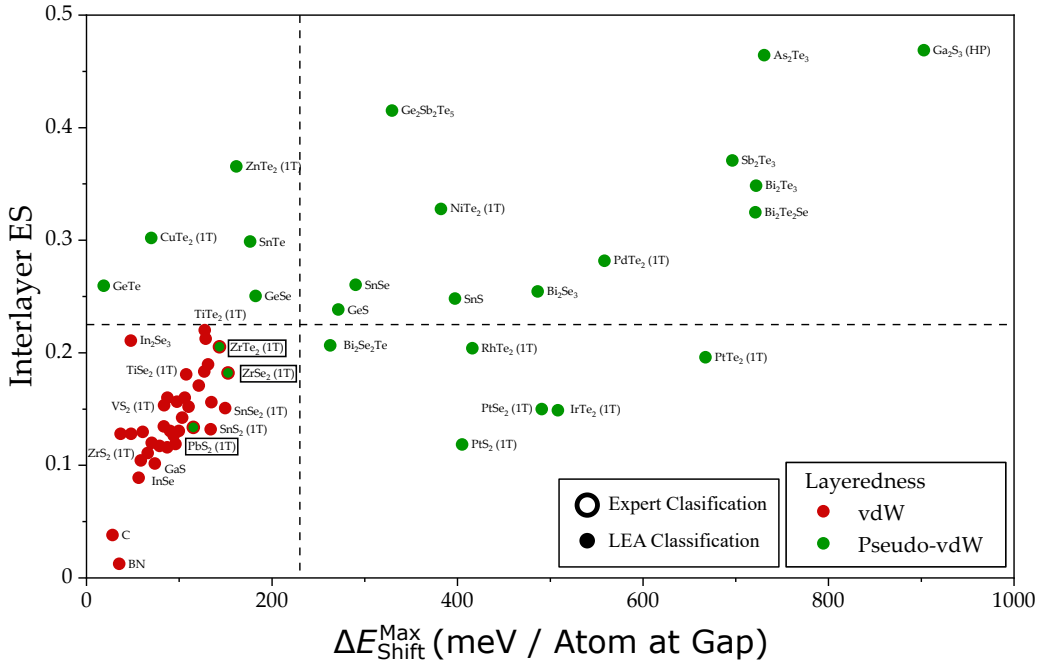


Figure 2.21: Maximum translation energy $\Delta E^{\mathrm{Max}}_{\mathrm{Shift}}$ plotted against the interlayer electrons shared value $\mathrm{ES}_{\mathrm{Interlayer}}$. True vdW compounds are denoted in red and pseudo-vdW systems in green. The center color refers to the classification as determined by the algorithm described in section 2.3.1, while the edge color denotes the expert/human classification. Modified from [68].

representation separates the true from the pseudo-vdW systems. The lower left quadrant only contains true vdW systems (which show a linear trend), while the remaining three quadrants are populated by the pseudo-vdW compounds (no obvious trend is visible). This split implies that neither the maximum translation energy $\Delta E^{\mathrm{Max}}_{\mathrm{Shift}}$ nor $\mathrm{ES}_{\mathrm{Interlayer}}$ by themselves, but both quantities are required in conjunction to achieve distinction between true and pseudo-vdW materials. As far as the author is concerned, maps as shown in figure 2.20 and 2.21 have not been reported in the literature thus far.

Within this chapter, it has been shown that structural (d_{2D}), DFT-based ($\Delta E^{\mathrm{Max}}_{\mathrm{Shift}}$) and QTAIM-based ($\mathrm{ES}_{\mathrm{Interlayer}}$) properties can be utilized to analyze layered systems and to detect subtle differences among them. While a more detailed analysis has been conducted in [68], the topic of layered compounds in conjunction with QTAIM bonding descriptors offers many open questions that promise novel perspectives on 2D material science. So far MVB has mainly been discussed in materials where all p orbitals formed metavalent bonds (e.g. GeTe). It could be interesting to investigate materials where only two p orbitals contribute, creating atomic layers of metavalent character, with the remaining p orbital forming e.g. regular covalent bonds to interconnect these layers. This could enable the tailoring of strongly anisotropic behavior, as MVB has been linked to unique property behavior (see chapter 1).

2.4 Scaling and Confinement in Ultrathin Chalcogenide Films

The paper "Scaling and Confinement in Ultrathin Chalcogenide Films as Exemplified by GeTe" was first-authored by Peter Kerres and published in *Small*[82]. As already mentioned in section 2.1, chalcogenides feature a unique property set that enables their application as functional materials such as thermoelectrics, topological insulators and photonic switches. Their properties include high optical absorption, high electron mobilities, large values of the Born effective charge, low thermal conductivities, and relatively small effective masses. The origin of this unusual property set has been attributed to metavalent bonding. As metavalent bonding emerges when electron localization and delocalization are competing with each other, confinement of thin film samples, as realized by reducing the film thickness, should alter this balance and change properties. By decreasing film thickness, the ratio of interface to octahedrally coordinated bulk atoms increases, where only the latter should be able to form metavalent bonds. While the original publication by Kerres et al.[82] provides an exhaustive report regarding experimental details and various properties, this section will focus on optical properties and quantum chemical descriptors, as they are computationally more readily available. It serves to underline that experimental results are in good agreement with the calculated properties and that the degree of distortion is a tangible parameter that can be tuned not only in simulation, but also in practice.

Confinement and Peierls Distortion

Kerres et al. employ reflection high-energy electron diffraction (RHEED) and X-ray diffraction (XRD) to experimentally verify that the Peierls distortion (i.e. long to short bond ratio, see figure 2.4) in GeTe[22] is thickness dependent. By decreasing the film thickness, the Peierls distortion (PD) increases and vice versa, while the influence of Ge vacancy states can be ruled out as a driving force of changes in PD. Three limiting cases of film thickness will be used to illustrate the effect of the PD, i.e. films with thicknesses of 2.8 nm, 7.3 nm and 62.5 nm. Using XRD, Kerres et al. calculate the value of the long to short bond ratio of these films as shown in table 2.14.

Table 2.14: Values of the short to long bond ratio for different thicknesses and corresponding ES and ET values of GeTe thin films, as obtained by Kerres et al.[82].

GeTe Film Thickness	Short to Long Bond Ratio	Electrons Transferred	Electrons Shared
2.8 nm	1.24	0.16	1.25
7.3 nm	1.20	0.16	1.13
62.5 nm	1.16	0.18	1.05

[22] The GeTe films feature rhombohedral ($R3m$) structures, which approach a perfect cubic arrangement ($Fm\bar{3}m$) for vanishing Peierls distortion.

Dielectric Function and Peierls Distortion

Based on the experimentally derived unit cells of the 2.8 nm, 7.3 nm and 62.5 nm thin films, the orbital-resolved imaginary part of the dielectric function $\epsilon_2(\omega)$ is calculated (see chapters 5.7 and 7.3), as shown in figure 2.22[23]. As already discussed in chapter 2.1.1 (see table 2.6),

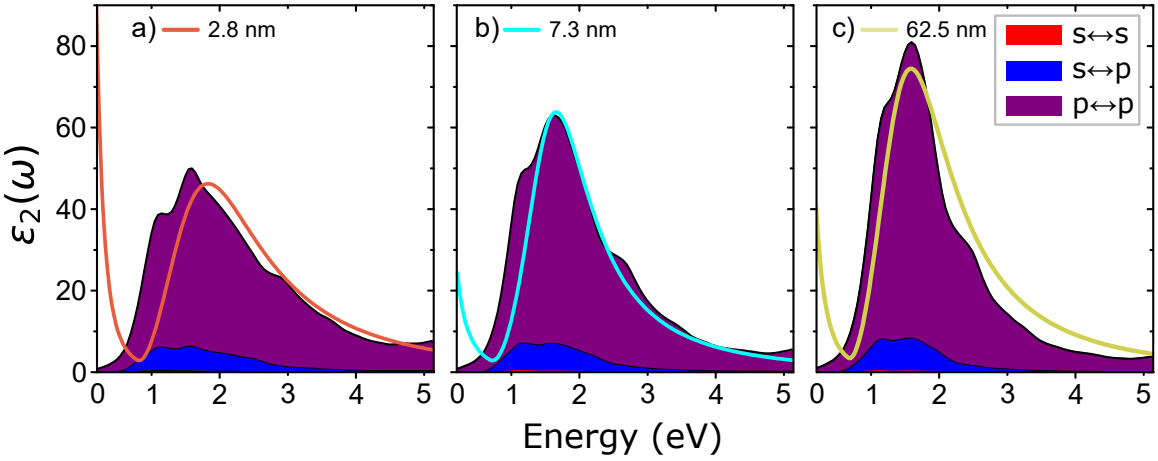


Figure 2.22: Orbital-resolved imaginary part of the dielectric function $\epsilon_2(\omega)$ for a): 2.8 nm, b): 7.3 nm and c): 62.5 nm thin films of GeTe. Experimental values denoted as orange (2.8 nm), turqouise (7.3 nm) and dark yellow (62.5 nm) lines. Simulated values are obtained by DFT calculation, as described in chapters 5.7 and 7.3. The shaded areas indicate the contribution of the respective orbital transitions. The s-p transition is approximately identical for all film thicknesses, while the p-p transitions along with the total are reduced for thinner films, as the p-p overlap is reduced due to larger Peierls distortion. Experimental data (Ellipsometry) obtained by Kerres et al. and modified from [82].

GeTe is metavalently bonded and the bonds are mainly p orbital dominated. In the limit of no distortion ($Fm\bar{3}m$ structure), a perfect octahedral arrangement is assumed, where each site features six nearest neighbors, forming half-filled σ bonds ($ES \approx 1$, see also figure 2.9). The optical transitions hence also mainly originate from p-p transitions (shaded areas in figure 2.22). With increasing PD, the alignment of the p orbitals is altered and their overlap decreases, which reduces the transition dipole matrix element and consequently the maximum of $\epsilon_2(\omega)$. It is noteworthy how well the simulated data (black line) fits the experimentally obtained imaginary part of the dielectric function (colored lines), especially for the 7.3 nm thin film. While DFT does have its weaknesses and caution regarding DFT results is always justified, figure 2.22 shows that experimental and DFT results are compatible and complement each other with additional information.
Figure 2.23 shows the position of the three GeTe thin films on the ES/ET map. As already illustrated in figure 1.14, increasing the PD increases the ES value and renders compounds more covalent, i.e. in this case they approach the metavalent to covalent border. This is the expected behavior for metavalently bonded systems, as the increase of PD tips off the balance

[23] While experimentally films of finite thicknesses are measured, the simulations assume bulk structures, while however employing the same structure and distortion as the thin films.

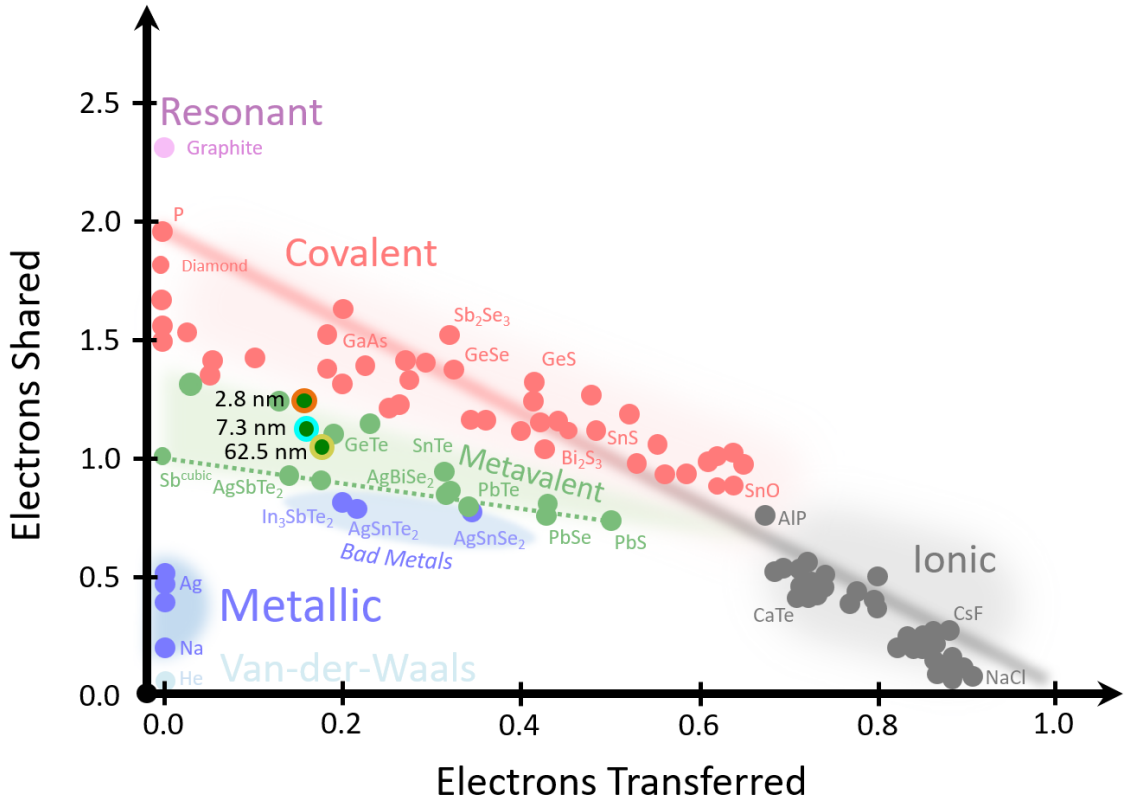


Figure 2.23: ES/ET map highlighting the positions of the $2.8\,\mathrm{nm}$ (orange outline), $7.3\,\mathrm{nm}$ (turquoise outline) and $62.5\,\mathrm{nm}$ (dark yellow outline) thin films of GeTe. With decreasing film thickness the Peierls distortion increases, and the ES value increases as well, i.e. the compounds become more covalent. Modified from [10, 82].

of the six bonds (in the cubic phase equivalent) to the nearest neighbors. With three shorter and three longer bonds, the electrons are increasingly more localized in the shorter bonds, weakening MVB and ultimately breaking it down when the competition between localization and delocalization favors localization too strongly.

Within this section, it has been shown that experimental and simulated (DFT) results are in good agreement with each other, at least regarding GeTe. This is reassuring as GeTe is one of the prime examples of metavalent bonding, implying that the theoretical concepts used to describe MVB are on solid foundations. Furthermore, reducing the sample thickness has been established as an experimentally available method to modify ES (keeping ET mostly constant). In conjunction with the option to modify ET via atomic substitution, methods to separately tailor ES and ET are established and experimentally tangible.

2.5 Further Landmarks of Property Space

Within this section, only short summaries of publications are presented, as the contributions of the author of this thesis to these papers was of a supportive nature. As each publication assumes a separate viewpoint, they are able to highlight features of property space from a different perspective. The first paper discusses the functional material In_3SbTe_2, which is located outside the green region of the ES/ET map. The publication shows that the bonding map provides insights beyond metavalently bonded systems as well. The second paper dives deeper into the already discussed lead chalcogenides, focussing on additional properties that change during ET-driven bond type switching. Moreover, the properties presented are obtained experimentally, confirming the predicted discontinuous behavior. The third publication discusses off-resonant light excitation to drive phase switching in SnSe, indicating that metavalent bonding could be linked to many other physical phenomena.

2.5.1 In_3SbTe_2 as a Programmable Nanophotonics Material Platform

The publication "In_3SbTe_2 as a Programmable Nanophotonics Material Platform for the Infrared" was first-authored by Andreas Heßler and published in *Nature Communications*[83]. It discusses the compound In_3SbTe_2 (IST312), a next-generation phase-change material (PCM). What sets it apart from the PCMs discussed so far, e.g. compounds in the GeTe family (see chapter 2.1.1), is that it is metallic in its crystalline phase, and a dielectric in the amorphous phase. The other PCMs that have already been discussed in the previous chapters were all metavalently bonded in their crystalline phase. As shown in figure 2.24, IST312 is located in the metallic region of the ES/ET map, more precisely in the region of the *bad* or *strange metals*, characterized amongst others by conductivities that are low for metals, but still higher than the conductivity values of metavalently bonded semiconductors. Heßler et al. use the high dielectrical optical contrast of IST312 to envision tunable nanophotonic devices by using laser pulses to write, modify and erase resonant metallic nanostructures, i.e. nanoantennae, on IST312 thin films. These nanoantennae are then proposed to enable programmable nanophotonic devices for various fields of application, such as telecommunication and reconfigurable holograms[83].
The work of Heßler et al. shows that the concept of the bonding map can be employed to describe non-metavalently bonded systems as well. In particular, the region around (in this case below) the metavalent-metallic borders seems to host many intriguing compounds, such as IST312, or superconductors like $AgSnTe_2$[84]. The transition zone between metavalent and metallic promises to be an abundant field of research for further investigation.

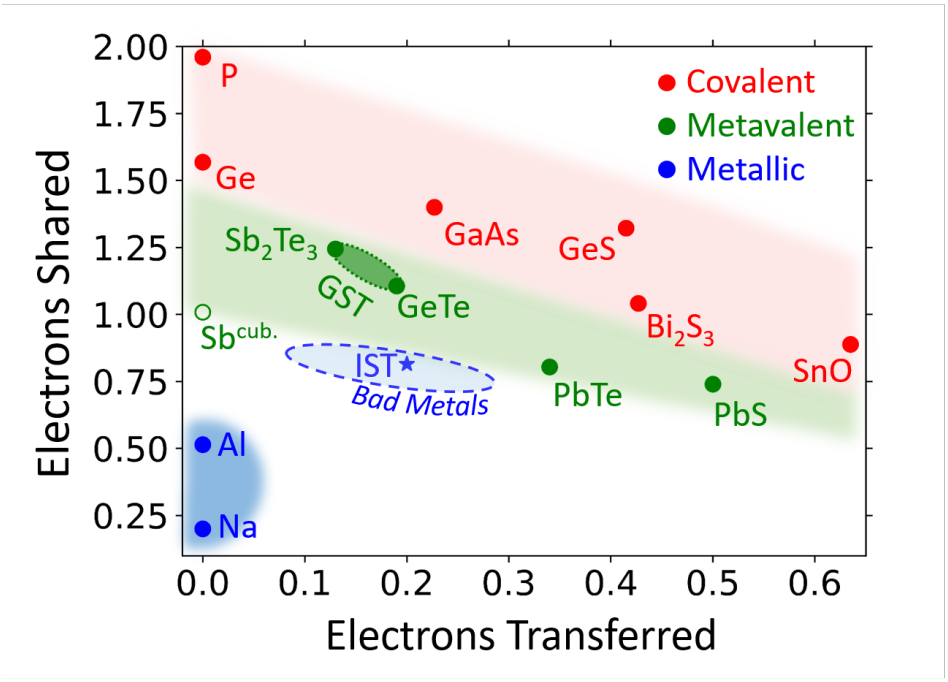


Figure 2.24: ES/ET map highlighting the position of In_3SbTe_2 (IST312). IST312 is located in the region of the so-called *bad* or *strange metals*, right below the metallic-metavalent border. This positioning is corroborated by the conductivity of IST312 being higher than that of regular PCMs, but relatively small for a metal. Modified from [83].

2.5.2 ET-Driven Changes of Chemical Bonding in Lead Chalcogenides

The publication "Discovering Electron-Transfer-Driven Changes in Chemical Bonding in Lead Chalcogenides" was first-authored by Stefan Maier and published in *Advanced Materials*[24]. It focuses on changes of the chemical bonding within lead chalcogenides (PbX, where X=Te, Se, S, O) by invoking quantum-chemical bonding descriptors, as well as property and bond-breaking descriptors. The PbX systems have already been discussed in the previous chapters to some extent (see section 2.1), and their position on the ES/ET map is depicted in figure 2.6. The interest in PbX systems stems from the fact that PbTe and PbSe are known to be good thermoelectrics[85–87], as well as PbO being located in the border region of metavalent to iono-covalent bonding, while PbTe, PbSe and PbS are located in the metavalent region (see figure 2.6). This is indeed reflected in a change of properties, as shown in figure 2.25. While the properties of PbTe, PbSe and PbS are generally relatively similar, discontinuous behavior is observed upon the transition from PbS $\rightarrow$ PbO, indicating the transition from metavalent to iono-covalent bonding. This property shift is, however, accompanied by a change of structure as well. While PbTe, PbSe and PbS are crystals having a rock salt structure (ECoN = 6), PbO assumes a massicot-type structure (ECoN = 4, also called β-PbO). To disentangle structure and chemical bonding, Maier et al. employ laser-assisted atom probe tomography (APT) to

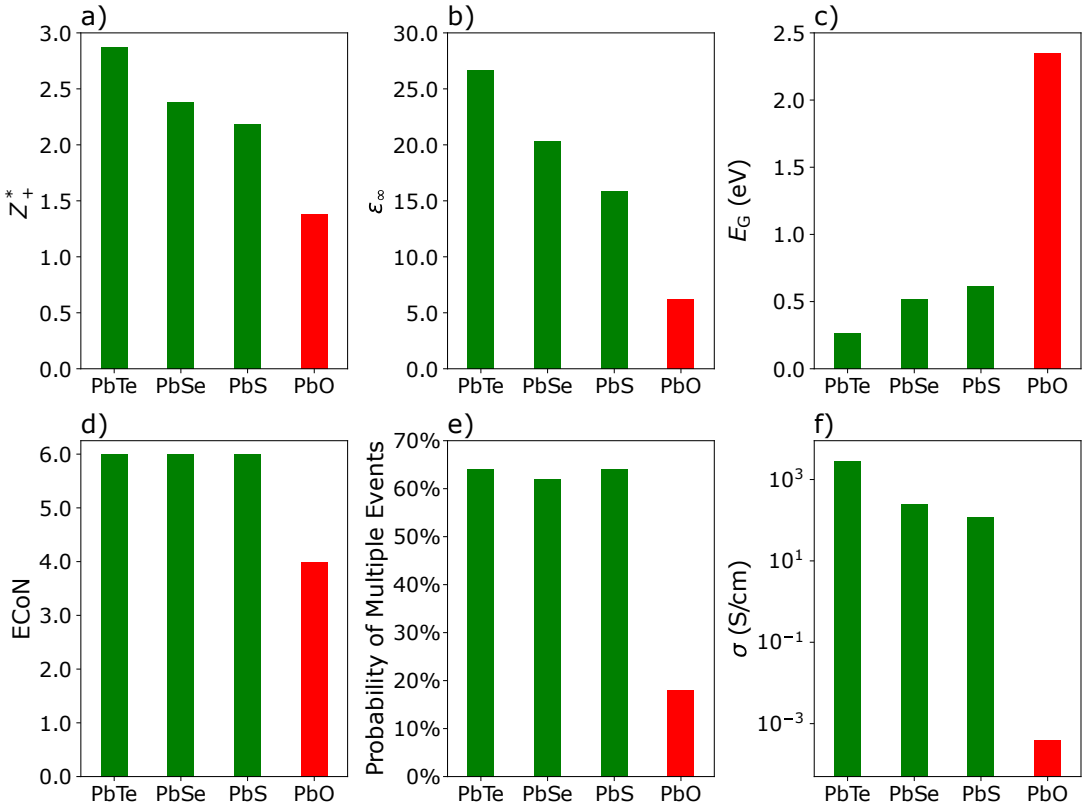


Figure 2.25: Properties of the lead chalcogenides PbTe, PbSe, PbS and PbO. a): Elevated Born effective charge $Z^*_+ = Z^*/[\mathrm{OxidationState}]$, b): Optical dielectric constant ϵ_∞, c): Optical band gap E_G, d): Effective coordination number ECoN, e): Probability of multiple events (PME), f): Electrical conductivity σ. The iono-covalent PbO (red) differs in terms of properties distinctly from the metavalently bonded PbTe, PbSe and PbS (green). Properties obtained by and graph modified from Maier et al.[24].

investigate the bond breaking behavior of the PbX systems. This is done by applying a DC voltage of 3-8 kV to a microscopic sample and exposing it to short laser pulses with energies of about 15 pJ in order to detach atomic fragments of the sample's tip. Previous studies have shown the probability of multiple events (PME), describing the probability that more than one atom is detached per laser pulse, is considerably higher for metavalently bonded systems (PME $> 50\%$), while only moderate for all other compounds (PME $\approx 20\%$)[88–90]. Employing APT, the difference in bonding is confirmed for the lead chalcogenides, as for PbTe, PbSe and PbS the PME $> 60\%$, while for PbO PME $\approx 18\%$. This sudden drop in PME accompanied by the property change is indicative for a change in chemical bonding. Maier et al. thus show experimentally that metavalent bonding can be impaired by not only ES (as shown in the foregoing), but also by increasing ET, up to the complete breakdown of MVB. Concomitantly, the properties of ET-tailored materials change as well, once again underlining the employability of ES and ET to tailor material properties. While Maier et al. discuss the property changes in minute detail, figure 2.7 exemplifies the bonding-property relation focusing on the imaginary part of the dielectric function.

2.5.3 Nonresonant Optomechanical Phase Control

The manuscript "Nonresonant Optomechanical Phase Control" was first-authored by Jiaojian Shi and is submitted for publication[91]. Shi et al. utilize Raman spectroscopy and X-ray diffraction (XRD) to investigate mid-infrared (MIR) excited SnSe, for which an off-resonant light excitation has been proposed to be able to drive phase switching, promising ultrafast switching speeds and reduced energy consumption for memory devices. While Shi et al. provide a detailed description of the experimental methods and their analysis, this section focuses on the link to metavalent bonding only. All information is taken from [91], if not stated otherwise.

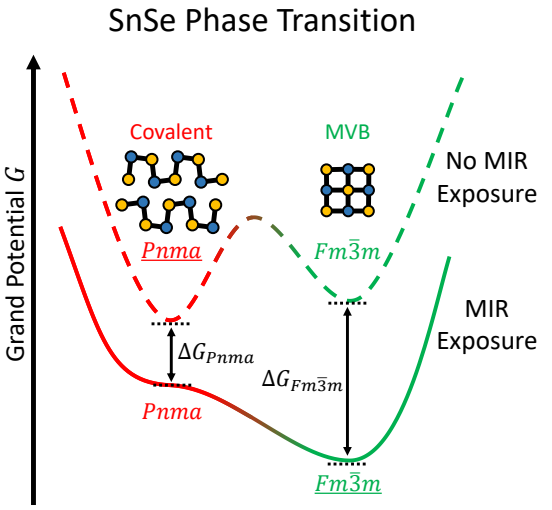


Figure 2.26: Sketch of the phase transition from $Pnma$ to $Fm\bar{3}m$ in SnSe under MIR (mid-infrared) exposure. Without MIR field (broken line), the stable configuration is the covalently bonded $Pnma$ phase, separated by an energy barrier from the only slightly less favorable metavalently bonded $Fm\bar{3}m$ phase. Under MIR exposure (continuous line), the thermodynamic grand potential G is altered (see equation 2.5) in such a way that the $Fm\bar{3}m$ phase becomes the stable configuration. Adapted from [91].

In the intrinsic case, the stable phase of SnSe is the covalently bonded $Pnma$ phase. With a difference of only about $2-3\,\mathrm{meV}$ per atom, the metavalent $Fm\bar{3}m$ phase is energetically only slightly less favored. Under MIR exposure, this situation changes and the $Fm\bar{3}m$ phase becomes the stable structure (see figure 2.26). This effect is caused by the light field inducing polarization, which couples to specific atomic coordinates and changes the thermodynamic equilibrium[20, 92, 93][24]. As shown in figure 2.27 c), the change in reflectivity R between the $Fm\bar{3}m$ and $Pnma$ phases averages to about 25% in the energy range between $1-4\,\mathrm{eV}$. The pronounced property change of SnSe upon small energy investment, as well as p orbital dominated bonding/properties (see figure 2.27) is characteristic of metavalent compounds due to the susceptibility to small perturbations of the half-filled σ bond, as already exemplified for GeTe in figure 2.4. The ES/ET values of the SnSe phases shown in table 2.15 further underscore that the $Fm\bar{3}m$ phase is metavalently bonded and the $Pnma$ phase is covalently bonded, as they are located in distinctively different regions of the ES/ET bonding map (compare e.g. figure 2.14). The significant role of MVB in enabling the phase transition under MIR exposure can further be motivated by considering that the free energy density G is modified by a term G_F proportional

[24] It should be noted that the distortions under MIR exposure measured by Shi et al. are not fully consistent with the expected $Fm\bar{3}m$ phase. The $Fm\bar{3}m$ phase will still be assumed in the further analysis, as the exact phase could not be determined.

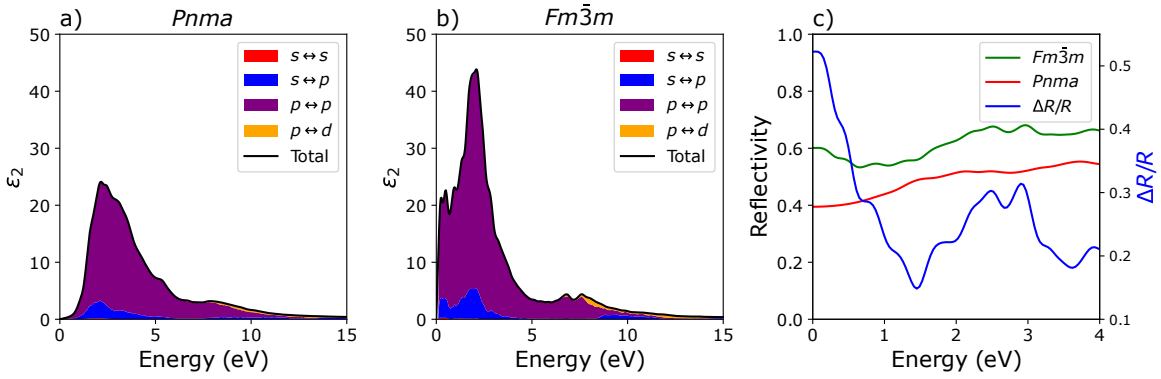


Figure 2.27: Orbital-resolved imaginary part of the dielectric function of: a) $Pnma$ SnSe, b) $Fm\bar{3}m$ SnSe, c) Reflectivity R and relative reflectivity change $\Delta R/R$ of $Pnma$ and $Fm\bar{3}m$ SnSe. Calculated as described in chapter 5.7 and 7.3. Modified from [91].

to the real part of the dielectric function:

$$G_{\mathrm{F}} = -\frac{1}{2}\vec{F}^{*}(\omega_0, t) \cdot \epsilon_1(\omega_0) \cdot \vec{F}(\omega_0, t), \tag{2.5}$$

where $\vec{F}(\omega_0, t)$ is the incident MIR field, ω_0 the MIR field frequency and $\epsilon_1(\omega_0)$ the real part of the dielectric function (tensor). It has already been shown for GeTe in figure 2.4 that the dielectric properties vary substantially between the covalent and metavalent phase. The same holds true for SnSe, where the ratio of the real part of the dielectric function amounts to about $\frac{\epsilon_1^{Fm\bar{3}m}(\omega_0)}{\epsilon_1^{Pnma}(\omega_0)} \approx 2.5$ at MIR wavelengths of about $\lambda_0 \approx 5\,\mu\mathrm{m}$. This pronounced constrast in ϵ_1 drives the phase transition in SnSe, as under MIR exposure of $F \gtrsim 0.5\,\mathrm{V\,nm^{-1}} \sim 5\,\mathrm{mJ\,cm^{-2}}$, the additional term G_{F} reduces the total free energy of the metavalent phase ($Fm\bar{3}m$) to be below the free energy density of the covalently bonded phase ($Pnma$). Within the MIR field, the metavalent phase hence becomes the thermodynamically stable phase, while without the MIR field the covalent phase is energetically favorable. Shi et al. estimate the required free-space incident energy density for the $Pnma$ phase switching efficiency of the non-resonant photoexcitation to be about $E_\nu = 0.5\,\mathrm{meV}$ per atom. The free-space incident energy density corresponds to the applied energy density and is independent of the optical properties of the material. It can be calculated via $E_\nu = \frac{F}{ct}$, where F is the excitation fluence, c the speed of light and t the duration of the laser pulse. Alternatively, the material switching efficiency, corresponding to the heat load, can be assessed via $E_{\mathrm{s}} = \frac{F}{d}$, where d is the penetration depth. Assuming the linear penetration depth of SnSe to be within the

Table 2.15: Electrons transferred (ET), electrons shared (ES) and the effective coordination number (ECoN) of $Pnma$ (covalent) and $Fm\bar{3}m$ (metavalent) SnSe. Calculated as described in chapters 6 and 7.1.

SnSe	$Pnma$	$Fm\bar{3}m$
ET	0.41	0.43
ES	1.09	0.80
ECoN	3.71	6.04

range of 1-1000 µm yields a material switching efficiency of 0.02 to 17 meV per atom. While values within this range can be achieved in other light-induced phase transitions, the estimated minimum of ca. 0.02 meV per atom is approximately one order of magnitude lower than what is reported for other compounds in the literature[91].

The manuscript by Shi et al. shows that MVB enables fascinating physical phenomena and mechanisms promising significantly improved functional performance, e.g. for ultrafast switching speeds in non-volatile rewritable memory devices.

Chapter 3

Summary - The Journey's End

"Is there structure in property space, and if yes, can it be navigated?"
This was the initial question posed by this thesis. By classifying a dataset of about 330 compounds, chapter 1 conclusively showed that such a structure indeed exists. An expectation maximization algorithm (EMA) assigned each compound to one of four clusters. Each of these clusters contained compounds of one of the classical chemical bonds: covalent, ionic, and metallic bonding as well as the recently proposed metavalent bonding. It could be shown that chemical bonding is the foundation of property space, upon which the properties of a material are built. Exploring the relation of the properties with each other and their distribution within the different bonding types showed that sets of correlated properties are required to attribute a material to a specific bonding type. The trends found and interconnections confirmed the notion of a structured property space.
To navigate the emerging landscape of properties, the quantities "Electrons Transferred (ET)" and "Electrons Shared (ES)" were shown to be excellent bonding descriptors that can be used to draft a map of property space, reproducing the separation of bonding types in agreement with the classification algorithm. While providing insight into the structure of property space, the *t*-distributed stochastic neighbor embedding (t-SNE) was shown to be unusable as a tool of navigation. Depicting properties on the z-axis of the ES/ET map revealed that these quantities are also excellent property predictors. The ES/ET map can be and already is being used to tailor properties and discover new materials. An interactive version of this topographic version of the map was developed to provide an easy access point to journey into property space.
Chapter 2 utilized the framework of metavalent bonding and quantum-chemical calculations presented in chapter 1 to start exploring property space. To this end, various publications to which the author of this thesis contributed were revisited. The first excursion covered a review paper published in the *Hall of Fame* of *Advanced Materials*, highlighting the concept of chemical bonding in general and how it can be revisited using the latest quantum-chemical methods. Especially conflicts surrounding GeTe and the number of electrons partaking in bond formation have been addressed, concluding that metavalent bonding is the soundest explanation.
The second stop along the journey through property space focused on halide perovskites and how ES and ET can be used to explain and predict their properties. The underlying reason was shown to be that halide perovskites feature one metavalent (-like) and one ionic bond, while the structurally similar oxide perovskites feature covalent and ionic bonds. The metavalent bond shapes the band structure in such a way that the effective masses are low, while the optical absorption is high, which is desirable for photovoltaic applications.
At the third stopover, layered systems were investigated. An algorithm has been established

to automate the classification into true van der Waals layered systems, pseudo van der Waals systems, and bulk structures, based on atomic arrangement alone. A similar separation could be achieved by employing the translation energy in conjunction with interlayer ES values. A map for 2D materials has been developed by utilizing the translation energy as well as the layeredness scale by Stevanović et al., promising to unravel hidden links among layered compounds.
The fourth publication dealt with thin films of GeTe, their optical properties, and how ES and MVB can be tailored experimentally by modifying film thickness.
The journey was concluded with a bouquet of publications, highlighting the applicability of the ES/ET map to non-MVB materials, the tailoring of ET by atomic substitution in lead chalcogenides and the potential of border crossing for applications exemplified by ultrafast switching in SnSe.
Property space is vast and many regions remain uncharted. This thesis aims to be an invitation to everyone to become an adventurer, start exploring and to discover the uncountable treasures property space provides. The tools provided in this thesis should be helpful for the first trip alone, but to reach the highest property peaks in the lush green mountainsides of MVB and the lowest resistances in the trenches of the metallic blue sea, additional methods will surely become necessary. Some daunting territories to conquer next are presented in the following outlook.

Chapter 4

Outlook - New Frontiers

The concept of metavalent bonding has gained more and more traction in the last few years. It is being employed by an increasing number of research groups to describe phenomena of functional materials and others. Arora et al.[94] for example investigate the emergence of MVB along continuous transitions of structure and chemical composition. Arora et al. conclude that MVB results from *weakly* breaking symmetry within metalloids, which is in accordance with figures 2.7, 2.9 and 2.15.

Still, many challenges remain. The most imminent one is probably the positioning of d electron bonded systems in the ES/ET map. The ES values of such systems tend to be much higher than the ones of s or p bonded compounds. FeN for example features an electrons shared value of ES ≈ 1.06 with an electron transfer of ET $\approx \frac{1.18}{3} = 0.39$ (using the formal oxidation state of nitrogen). These values would place the metallic FeN around the metavalent/covalent border in the center of the ES/ET map, inconsistent with the separation of bonding types achieved thus far. Arguably, compounds featuring d electrons require some sort of renormalization to be compatible with the ES/ET map in its current form. The necessity to renormalize could originate from the increased total number of electrons that can occupy d-shells (up to 10) compared to p-shells (up to 6). Another likely reason could be the more complex and pronounced hybridization of the d orbitals with the remaining orbitals, while the compounds discussed so far were mostly p orbital dominated with minor s orbital contributions.

Another challenge is the incorporation of more complex compounds into the map concept. Thus far, almost exclusively monoatomic and binary compounds have been positioned on the map, as they only feature one ET and one (dominant) ES value, rendering their location on the map unambiguous. The exception are compounds like Sb_2Te_3, where the ES and ET values are similar enough to justify averaging, and the halide perovskites discussed in chapter 2.2, where two bonds per compound were marked on the map. Yet, these compounds are structurally still relatively simple and the general approach to positioning arbitrarily complex compounds needs to be defined.

While the ES/ET map is actively used to design properties, e.g. thermoelectrics[95] and halide perovskites[47] (see also section 2.2), the fully automated navigation system for material properties advertised in the introduction of this thesis has not yet been achieved. One possible approach could be to train a machine learning algorithm on predicting the position on the map for a given property portfolio, and then to find a compound, by alloying neighboring compounds or atomic substitution, located at that point. However, this concept is once more limited by the two-dimensional character of the map. Attempting to reduce highly complex functional materials to only two parameters is probably a slightly too optimistic endeavor. One could imagine a multi-dimensional ES/ET map, e.g. by assigning each ES and ET value real-space

coordinates originating from the atomic arrangement (resulting in a somewhat condensed form of the electron density on which DFT is based in the first place). While this kind of map cannot be visualized anymore, machine learning algorithms are perfectly suited for such complex data formats. The following section presents a proof of concept linking such a spacial ES/ET arrangement with properties.

The Amadeus Code

The *Amadeus (**A**utomated **M**aterial **A**ttribute **De**crypting **U**niversal **S**ystem) Code* aims to predict the band gap E_G from ES/ET and structural data to show that the information contained in the input used is sufficient to extract property data. The general layout of the network is sketched in figure 4.1. Each input compound is processed as follows: A 3D cube is

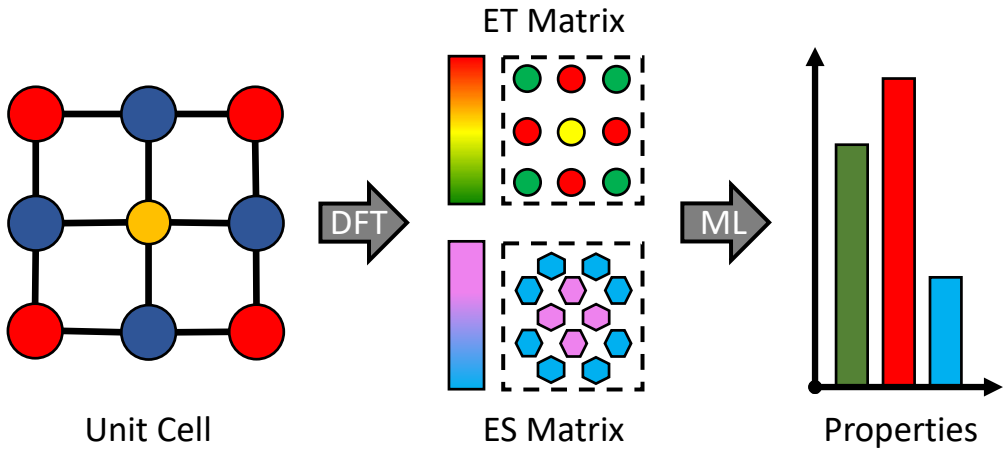


Figure 4.1: 2D sketch of the *Amadeus* neural network design. The unit cell data is used to compute ES and TET values. ES and TET are then arranged according to their positioning in three dimensions, where the TET are located at each site, and the ES between corresponding sites (here only depicted in 2D). The actual network also uses the atomic radius, the number of valence electrons and the number of shells, arranged similarly to the TET matrix. This input is then forwarded into a 3D convolutional neural network (CNN), where each input (ES matrix, TET matrix etc.) is treated as a separate *channel* (comparable to color channels of a pixel graphic), to predict a property.

discretized into a $30 \times 30 \times 30$ grid of voxels, each voxel representing a real-space cubic volume with an edge length of 0.75 Å. This cube is used to contain and represent the unit cell (shape) as well as properties that can be assigned to a certain point within the unit cell, such as ET at atomic sites and ES between the corresponding atomic sites. To accommodate for differently sized and shaped unit cells, a masking channel is introduced. The unit cell is placed within the cube, comparable to a 3D stencil, and each voxel (of the masking channel) is then assigned a value of 1, if the voxel is located within the unit cell, and 0 if the voxel is located outside the unit cell. Then, ES and TET are arranged according to their positioning in the unit cell

in three dimensions, where the closest voxel to each atomic site is assigned the corresponding TET value, and the voxels between corresponding sites the ES values. Additionally, the atomic radius, the number of valence electrons and the number of shells are used as input as well, arranged similarly to the TET values at the corresponding atomic site. A separate channel is used for each property. This results in a total of six input channels (including the masking layer). The input is then forwarded into a three-dimensional convolutional neural network (CNN)[1], typically employed for image recognition and related tasks. About 4000 unique materials from *Materials Project*[80] have been computed to serve as input (see chapter 7.5). Augmenting the data by rotating the inputs[2] results in a total of about 21000 compounds to train on[3]. Figure 4.2 shows the predictions of the normalized band gap E_G^N on the test set, i.e. compounds the CNN did not train on, containing about 2100 compounds. While the

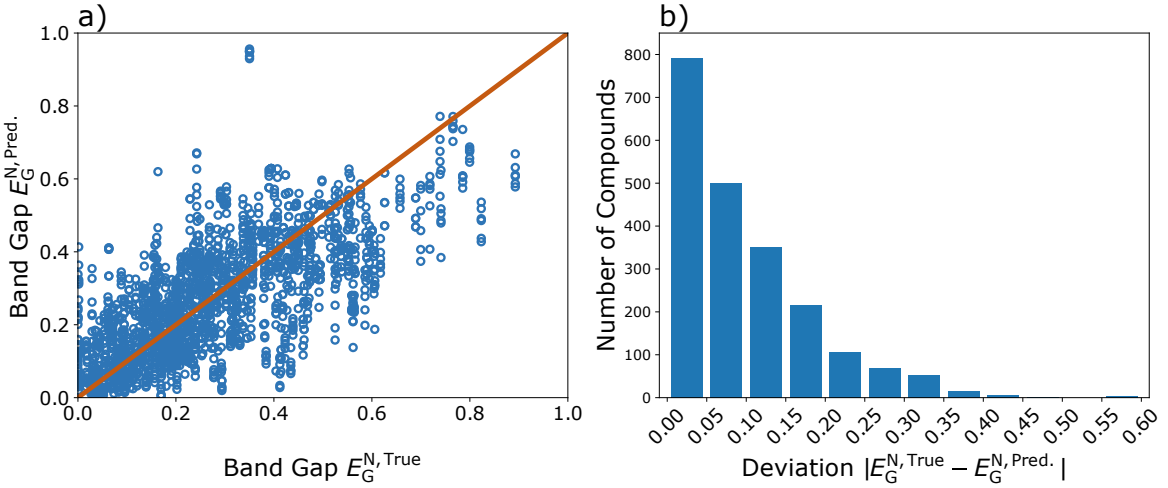


Figure 4.2: The *Amadeus Code* predicting the normalized band gap $E_\mathrm{G}^\mathrm{N} \in [0, 1]$ of about 2100 compounds. a): True band gap (as obtained from the *Materials Project*[80]) plotted against the predicted values. Optimally all datapoints should be located on the orange line. b) Deviation of true and predicted band gap $\Delta = |E_\mathrm{G}^{\mathrm{N,True}} - E_\mathrm{G}^{\mathrm{N,Pred.}}|$. For about 1640 compounds (78%) the predictions deviate $\Delta < 0.15$; for about 1290 compounds (61%) the predictions deviate $\Delta < 0.10$.

predictions are arguably not perfect, they show a general congruency and are thus a proof of concept. Especially considering that the input has been forced into a CNN-compatible format and that the hyperparameters and network design can barely be considered to be fully optimized, the predictive power achieved is promising. More sophisticated networks should improve significantly and enable a novel approach to material science.

[1] The property channels are therefore comparable to the color channels (RGB) of a regular pixel graphic. However, in this case more than 3 channels are utilized, in addition to operating in three instead of two dimensions. The masking layer acts as "outline/stencil", indicating the relevant part of the input volume.

[2] Considering a 2D example again: A club ♣ should by recognized as such, even when rotated: ♣ ♣ ♣. Hence, a single input file (♣), by rotating it, can be used multiple times to train the algorithm.

[3] Some compounds are dropped, as their unit cells do not fit into the 22.5 Å × 22.5 Å × 22.5 Å cube.

The secret to designing materials has come within reach. It can surely be found... somewhere in property space.

Part II

Chapter 5

Density Functional Theory

理論など、単なる言葉遊びに過ぎん。
いい加減、目の前で起きたことを認めるんだな。

岡部 倫太郎

There are two fundamental ways of determining the properties of a material. Any property can either be measured or calculated from first-principles. While both methods should yield identical results in an ideal world, reality imposes advantages and disadvantages to each respective approach. The experimental approach allows the measurement of real-world samples, yielding results that can be expected to occur in the application of said material. However, depending on the property in question, the required measurement setup can be financially expensive and is often limited to only being able to measure one single property. Each measurement requires labor time of skilled professionals, making it challenging to gather data for a large number of materials. Measurement noise and real-world conditions like air friction, gravity and sample contamination add to the challenges of the experimental approach. Determining properties by first-principle calculations is trivial in theory, as all information is contained in the (time-independent, nonrelativistic) Schrödinger equation:

$$H\Psi = E\Psi \tag{5.1}$$

While the Hamiltonian H is also trivial to construct, it can generally not be solved analytically. Furthermore, while numerically solving it is possible, it would be unfeasable even for the fastest supercomputers to date (especially for a system with a macroscopic amount of particles). Hence only by employing simplifications and approximations does it become possible to compute systems exceeding the complexity of a model system, which comes at the cost of introducing deviations from the exact values. Furthermore, even if solved sufficiently well, the results represent properties without any real-world influence, which are always present in application. In order to investigate a material suitably, experimental and computational results should always complement each other.
In this chapter, Density-Functional Theory (DFT) is presented, as DFT enables one to compute the properties of even complex systems numerically. DFT is a field of research in itself, so this chapter can only present a functional overview over the methods used and the theory behind them. The reader is referred to the respective literature for more detailed explanations. All information, if not stated otherwise, is taken from sources [96–102].

5.1 Mathematical Description of Crystals

This chapter aims to introduce the unit cell, a fundamental building block of solid-state physics, paramount for both experimental and theoretical considerations.
There are two main configurations a solid can assume: It can either be *crystalline*, and thus be called *crystal*, or it can be an amorphous solid. The fundamental difference between an amorphous solid and a crystal is that a crystal features long-range periodicity, while an amorphous solid does not[103–105]. Due to this long-range order, a crystal can be described by a *Bravais lattice*, requiring the definition of an atomic basis for any given compound. The basis represents a lattice point in the Bravais lattice and contains several atoms with a set relative distance to each other. The Bravais lattice itself is then constructed by translating the lattice points in such a way that the environment of each lattice point is identical, hence creating translational invariance. Each lattice point $\vec{R}$ of a Bravais lattice can hence be described via:

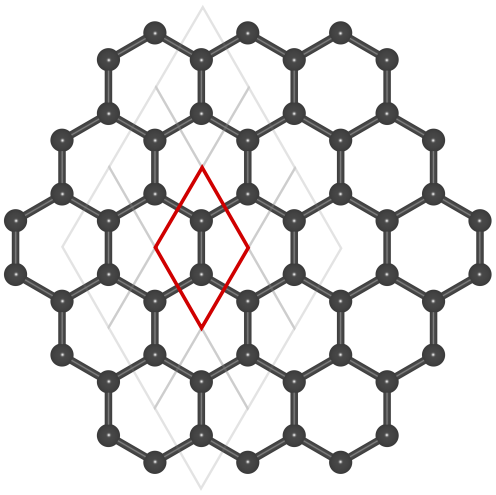


Figure 5.1: Graphene "honeycomb" lattice. This type of lattice is not a Bravais lattice, as the atomic sites do not have an identical environment, i.e. no translational symmetry exists. By defining a unit cell (red trapezoid), the resulting lattice becomes a Bravais lattice, as every lattice point is equivalent, as translating the lattice cell (gray trapezoids) reproduces the original lattice.

$$\vec{R} = \sum_{i=1}^{d} n_i \vec{a}_i \qquad n_i \in \mathbb{Z} \tag{5.2}$$

where d is the dimensionality of the lattice, and $\vec{a}_i$ are linearly independent vectors called lattice vectors. This relation signifies that each translation of any lattice point by any combination of multiples of lattice vectors will ultimately end up at another (identical) lattice point.
However, as not every lattice is necessarily a Bravais lattice, multiple lattice points can be subsumed into a *unit cell*, which then becomes the lattice point of a new lattice. For any given, infinitely repeating structure (with long-range order), a unit cell can be defined in such a way that a Bravais lattice is obtained (see figure 5.1). In practice, the basis is often incorporated into the definition of the unit cell, which is generally more practical. The unit cell itself is therefore defined by a set of lattice vectors (three in 3D) and the positions of the atoms within the unit cell. Each crystalline compound can then be fully described by a single unit cell and reconstructed by translational continuation (see figure 5.1, gray trapezoids). It should be noted that the unit cell is not uniquely defined for a given system. In fact, infinitely many representations exist. In most cases either a unit cell with the smallest number of atoms

possible (primitive unit cell), or one with orthogonal axes (conventional unit cell) is chosen.
Related to the description of a crystal in real-space is the description in momentum-space. Momentum-space is especially convenient, as it not only ties in directly with the periodic nature of the crystal in real-space, but also with the description of electrons as waves in quantum mechanics and density functional theory, where the wave vector $\vec{k}$ is the quantum number of choice. The wave function of a free electron can be expressed as a plane wave: $e^{i\vec{k}\vec{r}}$, where $\hbar\vec{k}$ is the momentum of the electron (which is why momentum-space is also called $\vec{k}$-space). The real-space (Bravais) lattice of a crystal can be transformed into the momentum-space reciprocal lattice by means of a Fourier transform:

$$e^{i\vec{G}\vec{R}} = 1 \tag{5.3}$$

where $\vec{R}$ is any lattice vector of the real-space lattice, and $\vec{G}$ the corresponding reciprocal lattice vector in $\vec{k}$-space. Hence the reciprocal lattice vectors are defined similarly to equation 5.2 as:

$$\vec{G} = \sum_{i=1}^{d} m_i \vec{b}_i \tag{5.4}$$

where $\vec{b}_i$ are linearly independent reciprocal lattice vectors. $m_i \in \mathbb{Z}$ holds true assuming the orthogonality:

$$\vec{a}_i \cdot \vec{b}_i = 2\pi\delta_{ij} \tag{5.5}$$

This implies that the reciprocal lattice is a Bravais lattice as well.
The reciprocal lattice vectors $\vec{b}_i$ can be constructed from a set of real-space lattice vectors $\vec{a}_i$ (in 3D) via:

$$\vec{b}_1 = \frac{2\pi(\vec{a}_2 \times \vec{a}_3)}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}, \quad \vec{b}_2 = \frac{2\pi(\vec{a}_3 \times \vec{a}_1)}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}, \quad \vec{b}_3 = \frac{2\pi(\vec{a}_1 \times \vec{a}_2)}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \tag{5.6}$$

From this relation follows that the unit cell volume in reciprocal space $V_{\mathrm{UC}}^{\mathrm{rec.}}$ is inversely proportional to the volume of the corresponding unit cell in real-space $V_{\mathrm{UC}}^{\mathrm{real}}$, as $V_{\mathrm{UC}}^{\mathrm{rec.}} = \vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3) \propto 1/V_{\mathrm{UC}}^{\mathrm{real}}$. This also holds true for the individual lattice vector pairs $\|\vec{b}_i\| \propto 1/\|\vec{a}_i\|$.

The Bloch theorem states that if electrons move in a periodic potential $V(\vec{r}) = V(\vec{r}+\vec{R})$, e.g. in a crystal lattice, each solution Ψ of the stationary Schrödinger equation:

$$H\Psi(\vec{r}) = \left[\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r})\right] = E\Psi(\vec{r}) \tag{5.7}$$

takes the form:

$$\Psi_{\vec{k}}(\vec{r}) = u_{\vec{k}}(\vec{r})e^{i\vec{k}\vec{r}} \tag{5.8}$$

where $u_{\vec{k}}(\vec{r}) = u_{\vec{k}}(\vec{r}+\vec{R})$ is a function with the same periodicity as the lattice. $\Psi_{\vec{k}}(\vec{r})$ is called *Bloch wave* and consequently a modulation of a plane wave ($e^{i\vec{k}\vec{r}}$) with the periodic function

$u_{\vec{k}}(\vec{r})$. As the Bloch wave is dependent on the wave vector $\vec{k}$, it can also be considered to be a function within the reciprocal space. Therefore, the Bloch waves $\Psi_{\vec{k}}(\vec{r})$ and the eigenenergies $E_n(\vec{r})$ (see equation 5.7) must also be periodic with the reciprocal lattice vector $\vec{G}$[102, 106]:

$$\Psi_{\vec{k}}(\vec{r}) = \Psi_{\vec{k}+\vec{G}}(\vec{r}) \rightarrow E_n(\vec{k}) = E_n(\vec{k}+\vec{G}) \tag{5.9}$$

Bloch waves are hence a convenient basis of electronic wave functions in crystalline solids.

5.2 Introduction to DFT

As stated in the introduction to this chapter, the solution of the (time-independent, non-relativistic) Schrödinger equation contains in principle all information about the corresponding many-body system. The Hamiltonian describes the kinetic energy T of the nuclei as well as the kinetic energy of the electrons t. Furthermore, all potential energy terms are included, i.e. electron-electron interaction $\nu_{\mathrm{el-el}}$, nuclei-electron interaction $\nu_{\mathrm{N-el}}$ and the nuclei-nuclei interaction $\nu_{\mathrm{N-N}}$. The full many-body Hamiltonian in Hartree atomic units thus reads:

$$H = \underbrace{-\frac{1}{2}\sum_n \frac{m_\mathrm{e}}{M_n}\nabla^2_{R_n}}_{T} \underbrace{-\frac{1}{2}\sum_i \nabla^2_{r_i}}_{t} + \underbrace{\frac{1}{2}\sum_{\substack{i,j \\ i\neq j}} \frac{1}{|r_i - r_j|}}_{\nu_{\mathrm{el-el}}} + \underbrace{\frac{1}{2}\sum_{\substack{n,n' \\ n\neq n'}} \frac{Z_n Z_{n'}}{|R_n - R_{n'}|}}_{\nu_{\mathrm{N-N}}} \underbrace{-\sum_{n,i} \frac{Z_n}{|R_n - r_i|}}_{\nu_{\mathrm{N-el}}} \tag{5.10}$$

with Z_n being the atomic number, M_n the mass, and R_n the position of nucleus n. The electron mass is indicated as m_e, while the position of electron i is denoted as r_i. R_n and r_i are vectors, for the sake of readability a vector sign is omitted, however. The Hamiltonian in equation 5.10 is written in Hartree atomic units to reduce the number of constants appearing in equations and is a common convention within the scope of DFT. This unit system results in all lengths being given in units of the Bohr radius $a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_\mathrm{e}e^2} \approx 0.53\,\text{Å}$, as well as all energies being given in units of Hartree $E_\mathrm{Htr} = \frac{e^2}{4\pi\epsilon_0} \approx 27.21\,\mathrm{eV}$.

The Born-Oppenheimer Approximation

Solving equation 5.1 with Hamiltonian 5.10 for a macroscopic object with a number of nuclei N of the order $\mathcal{O}(N) \approx (10^{23})$ is analytically impossible. Approximations need to be employed to make this task (at least numerically) feasible. The first reasonable assumption to make is that the dynamics of nuclei and electrons act on different timescales, i.e. for electrons the nuclei appear to be static, while for the nuclei the response of the electrons to changes in nuclear position can be assumed to be instantaneous. This simplification is called the *Born-Oppenheimer approximation* and is valid due to the mass difference between nuclei and electrons. Even for the lightest nucleus (the hydrogen core, i.e. a single proton), the mass

ratio amounts to $\frac{M_\mathrm{H}}{m_\mathrm{e}} \approx 1836$. This implies that electrons move much faster compared to the nuclei. The Born-Oppenheimer approximation thus allows the Hamiltonian 5.10 to be split into one describing solely the atomic cores (H_A) and another one dealing with the electronic contribution (H_el). Due to the substantial mass difference between nuclei and electrons, the atomic core can be considered static in the electronic Hamiltonian H_el and therefore the nuclei positions R_n become constant parameters. This reduces the Hamiltonian to be solved to the following form:

$$H_\mathrm{el} = -\frac{1}{2}\sum_i \nabla_{r_i}^2 + \frac{1}{2}\sum_{\substack{i,j \\ i\neq j}} \frac{1}{|r_i - r_j|} - \sum_i \underbrace{\sum_n \frac{Z_n}{|R_n - r_i|}}_{\nu(r_i)} = T + W + V \tag{5.11}$$

where $\nu(r_i)$ is the (constant) external potential acting on the electrons created by the nuclei. T denotes the kinetic energy term, W the electron-electron interaction term and V the potential energy term (created by electron-nuclei interaction).

5.3 The Hohenberg-Kohn Theorem

The solution of the Schrödinger equation as stated in 5.1 with the Hamiltonian 5.11 would yield a set of eigen(wave-)functions Ψ_i with corresponding eigenenergies E_i of the system in state i. The Hohenberg-Kohn theorem (HKT) provides an approach to solve the Schrödinger equation and is the foundation upon which density functional theory (DFT) is based. The HKT states the following[96, 101, 107, 108]:

- ◇ The system's energy E is a unique functional of the electron charge density $\rho(r)$:

$$\rho(r) : E_0 = E[\rho(r)] \tag{5.12}$$

- ◇ The charge density that minimizes the energy functional E is the ground-state electron density $\rho_\mathrm{GS}(r)$:

$$E[\rho_\mathrm{GS}(r)] \leq E[\rho(r)] = \langle \Psi | \, H_{el} \, | \Psi \rangle \,, \qquad \forall \rho(r) \neq \rho_\mathrm{GS}(r) \tag{5.13}$$

Consequently, the ground state expectation value of any observable O must also be a (unique) functional of the ground state electron density $\rho_\mathrm{GS}(r)$:

$$O[\rho_\mathrm{GS}(r)] = \langle \Psi_\mathrm{GS}[\rho_\mathrm{GS}(r)] | \, O \, | \Psi_\mathrm{GS}[\rho_\mathrm{GS}(r)] \rangle \tag{5.14}$$

This naturally includes the expectation value of H_{el}, i.e. the energy expectation value, for which a generic expression exists:

$$E_{\mathrm{el}}[\rho_{\mathrm{GS}}(r)] = \langle \Psi_{\mathrm{GS}}[\rho_{\mathrm{GS}}(r)] | \, H_{\mathrm{el}} \, | \Psi_{\mathrm{GS}}[\rho_{\mathrm{GS}}(r)] \rangle = T[\rho_{\mathrm{GS}}(r)] + W[\rho_{\mathrm{GS}}(r)] + \int \rho_{\mathrm{GS}}(r)\nu(r)dr \tag{5.15}$$

As $T[\rho]$ (kinetic part) and $W[\rho]$ (electron-electron interaction part) are independent of the external potential $V[\rho]$, their sum is addressed as $F[\rho]$ and assumed to be of identical form for each interacting system.

5.4 The Scheme of Kohn and Sham

The previous paragraphs showed that from the Hamiltonian 5.11 the electronic ground state density can by derived (in principle). In reality, the corresponding Schrödinger equation is still not solvable (neither analytically nor numerically) due to the term W, which represents the many-body electron-electron interaction. Fortunately, Kohn and Sham devised a scheme to mitigate this problem by replacing the Hamiltonian 5.11 with an equivalent, non-interacting Hamiltonian. This new Hamiltonian is equivalent in such a way that it replaces the real configuration of atoms and (interacting) electrons with an auxiliary system of non-interacting electrons that exhibits an identical ground state density $\rho_{\mathrm{GS}}(r)$[109]. The complex (and hard to compute) interaction term W is hence omitted in favor of the (effective) Kohn-Sham potential $V \rightarrow V_{\mathrm{KS}}$:

$$H_{\mathrm{el}} = T + W + V \rightarrow H_{\mathrm{KS}} = T + V_{\mathrm{KS}} \tag{5.16}$$

As the electron density $\rho_{\mathrm{GS}}(r)$ is periodic due to the lattice periodicity, V_{KS} is bound to be periodic as well. Moreover, the corresponding eigenstates of the Schrödinger equation utilizing the non-interacting H_{KS} are consequently one-electron (Bloch) wave functions φ_i[110]. Constructing Slater-determinants from the N_{e} one-electron wave functions φ_i, the total electronic wave function Ψ_{KS} can be reconstructed[101]:

$$\Psi_{\mathrm{KS}}(r_1, ..., r_{N_{\mathrm{e}}}) = \frac{1}{\sqrt{N_{\mathrm{e}}!}} det\left(\varphi_i(r_j)\right) \tag{5.17}$$

The corresponding ground state electron density $\rho_{\mathrm{GS}}(r)$ thus reads:

$$\rho_{\mathrm{GS}}(r) = \sum_{i=1}^{N_{\mathrm{e}}} |\varphi_i(r)|^2 \tag{5.18}$$

The Kohn-Sham energy functional can then be constructed as:

$$E_{\mathrm{KS}}[\rho(r)] = \langle \Psi_{\mathrm{KS}}[\rho(r)] | \, H_{\mathrm{KS}} \, | \Psi_{\mathrm{KS}}[\rho(r)] \rangle = T_{\mathrm{KS}}[\rho(r)] + V_{\mathrm{KS}}[\rho(r)] \tag{5.19}$$

While the kinetic part can be calculated as follows:

$$T_{\mathrm{KS}}[\rho(r)] = \sum_{i=1}^{N_{\mathrm{e}}} -\frac{1}{2} \int \varphi_i^*(r) \Delta \varphi_i(r) dr \tag{5.20}$$

The challenge consists in identifying an expression for $V_{\mathrm{KS}}[\rho(r)]$, which can be found by expanding E_{el} (equation 5.15) with a zero sum:

$$E_{\mathrm{el}} = T_{\mathrm{KS}}[\rho(r)] - T_{\mathrm{KS}}[\rho(r)] + T[\rho(r)] + W[\rho(r)] + \int \rho_{\mathrm{GS}}(r)\nu(r)dr \tag{5.21}$$

By comparison with 5.19, a general form of $V_{\mathrm{KS}}[\rho(r)]$ can be derived:

$$V_{\mathrm{KS}} = -T_{\mathrm{KS}}[\rho(r)] + T[\rho(r)] + W[\rho(r)] + \int \rho_{\mathrm{GS}}(r)\nu(r)dr \tag{5.22}$$

This enables the use of the one-electron eigenstates $\varphi_i(r)$ to compute $E_{\mathrm{el}}[\rho_{\mathrm{GS}}(r)]$ (while they are solutions of H_{KS}). Next, V_{KS} is further expanded by the so-called Hartree-exchange functional $V_{\mathrm{Htr}}[\rho(r)] = \frac{1}{2}\int\int \frac{\rho(r)\rho(r')}{|r-r'|} dr dr'$ and all terms encompassing many-body effects beyond the Hartree-exchange functional are collapsed into a single exchange-correlation functional $V_{\mathrm{XC}}[\rho(r)]$:

$$\begin{aligned} E_{\mathrm{el}}[\rho(r)] = {} & T_{\mathrm{KS}}[\rho(r)] - T_{\mathrm{KS}}[\rho(r)] + T[\rho(r)] + W[\rho(r)] \\ & + V_{\mathrm{Htr}}[\rho(r)] - V_{\mathrm{Htr}}[\rho(r)] + \int \rho_{\mathrm{GS}}(r)\nu(r)dr \end{aligned} \tag{5.23}$$

$$E_{\mathrm{el}}[\rho(r)] = T_{\mathrm{KS}}[\rho(r)] + V_{\mathrm{XC}}[\rho(r)] + V_{\mathrm{Htr}}[\rho(r)] + \int \rho_{\mathrm{GS}}(r)\nu(r)dr \tag{5.24}$$

$$E_{\mathrm{el}}[\rho(r)] = T_{\mathrm{KS}}[\rho(r)] + V_{\mathrm{XC}}[\rho(r)] + V_{\mathrm{Htr}}[\rho(r)] + V_{\mathrm{ext}}[\rho(r)] \tag{5.25}$$

Therefore:

$$V_{\mathrm{XC}}[\rho(r)] = T[\rho(r)] - T_{\mathrm{KS}}[\rho(r)] + W[\rho(r)] - V_{\mathrm{Htr}}[\rho(r)] \tag{5.26}$$

Unfortunately the correct form of the exchange-correlation functional $V_{\mathrm{XC}}[\rho(r)]$ is unknown. Finding appropriate approximations that yield reasonable results for any observable in comparison with the experiment remains the most daunting challenge of density functional theory. The variants of $V_{\mathrm{XC}}[\rho(r)]$ used within the scope of this thesis will be presented in chapter 5.5.4.

Existence of the Exchange-Correlation Density Functional

The existence of the exchange-correlation functional as a functional of the electron density ρ was proven by Mel Levy[111–113]. The proof is sketched for ground-state densities in figure 5.2. Assuming all fermionic many-particle wave functions[1] $|\Psi_\alpha\rangle$ could be constructed, the corresponding electron densities ρ_α can be calculated as:

$$\rho_\alpha(r) = N \int dr_2 ... \int dr_N |\Psi(r, r_2, ..., r_N)|^2 \tag{5.27}$$

All wave functions $|\Psi_\alpha\rangle$ are then sorted by their corresponding electron density ρ_α, i.e. wave functions with identical resulting electron densities are grouped together (see figure 5.2). For each of these sets of wave functions with identical electron densities, the wave function with the lowest energy is determined:

$$F^W[\rho_\alpha] = \min \langle\Psi| T + W |\Psi\rangle, \quad |\Psi\rangle \in M[\rho_\alpha] \tag{5.28}$$

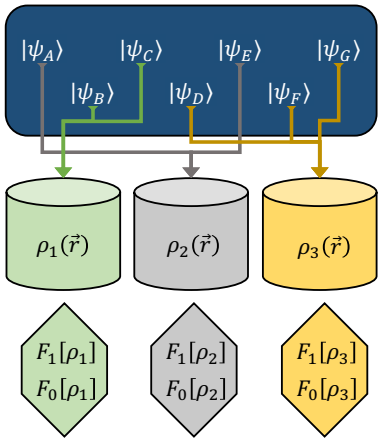


Figure 5.2: Outline of Levy's proof for the existence of an (exchange-correlation) density functional. For each fermionic wave function $|\Psi_\alpha\rangle$ (top part) an electron density ρ_α can be constructed. Recreated from [111].

with T the kinetic energy operator and W the electron-electron interaction operator. The external contribution V_{ext} is identical for all identical electron densities and can hence be omitted as it only results in a constant offset in energy. $F^W[\rho_\alpha]$ is hence independent of the external potential and can be considered a universal density functional. In the same fashion as $F^W[\rho_\alpha]$, another functional *only* considering the kinetic energy T is constructed:

$$F^0[\rho_\alpha] = \min \langle\Psi| T |\Psi\rangle, \quad |\Psi\rangle \in M[\rho_\alpha] \tag{5.29}$$

The total energy E therefore is a functional of the electron density ρ_α and can be expressed as:

$$E[\rho_\alpha] = F^W[\rho_\alpha] + \int \nu_{\text{ext}}(r)\, \rho_\alpha(r)\, dr \tag{5.30}$$

It follows that:

$$\underbrace{V}_{F^W + V_{\text{ext}}} = \underbrace{T}_{F^0} + V_{\text{ext}} + V_{\text{Htr}} + V_{\text{XC}} \tag{5.31}$$

$$V_{\text{XC}}[\rho_\alpha] = F^W[\rho_\alpha] - F^0[\rho_\alpha] - \underbrace{\frac{1}{2}\int\int \frac{\rho_\alpha(r)\rho_\alpha(r')}{|r - r'|} dr dr'}_{V_{\text{Htr}}} \tag{5.32}$$

Hence V_{XC} is indeed a functional of the electron density ρ.

[1] As electrons are fermions, their wave function are antisymmetric, i.e. $\Psi_{\text{F}}(r_1, r_2, ...) = -\Psi_{\text{F}}(r_2, r_1, ...)$. Bosonic wave functions are symmetric, i.e. $\Psi_{\text{B}}(r_1, r_2, ...) = \Psi_{\text{B}}(r_2, r_1, ...)$

Uniqueness of the Density Functional

The uniqueness of the functional required for the HKT can be motivated (for a non-degenerate ground-state) by proof by contradiction[107]: Given a system with

$$H_1 = V_1 + W + T, \tag{5.33}$$

where V_1 denotes the potential energy of the electrons originating from the potential $V_1'(r)$, T the kinetic energy of the electrons and W the electrostatic repulsion of the electrons. The total energy of the system then reads:

$$E_1 = \langle \Psi_1 | H_1 | \Psi_1 \rangle = \int V_1'(r)\rho_{GS}(r)dr + \langle \Psi_1 | (T + W) | \Psi_1 \rangle \tag{5.34}$$

It is then assumed that the same electron density $\rho_{\mathrm{GS}}(r)$ would also result from a different wave function Ψ_2 corresponding to the potential V_2 and therefore that $\rho_{\mathrm{GS}}(r)$ is not unique. In order for $\Psi_1 \neq \Psi_2$, V_1 and V_2 must differ by more than a constant offset. H_2 is then:

$$H_2 = V_2 + W + T \tag{5.35}$$

According to the Rayleigh-Ritz-Principle, the energy expectation value for a wave function that is not the ground state is always higher than the energy expectation value for the ground state wave function. Therefore, by assuming both Hamiltonians yield the identical electron density $\rho_{\mathrm{GS}}(r)$, it follows that:

$$E_1 < \langle \Psi_2 | H_1 | \Psi_2 \rangle = \langle \Psi_2 | H_2 | \Psi_2 \rangle + \langle \Psi_2 | H_1 - H_2 | \Psi_2 \rangle = E_2 + \int (V_1'(r) - V_2'(r))\rho_{\mathrm{GS}}(r)dr \tag{5.36}$$

as well as:

$$E_2 < \langle \Psi_1 | H_2 | \Psi_1 \rangle = \langle \Psi_1 | H_1 | \Psi_1 \rangle + \langle \Psi_1 | H_2 - H_1 | \Psi_1 \rangle = E_1 + \int (V_2'(r) - V_1'(r))\rho_{\mathrm{GS}}(r)dr \tag{5.37}$$

Adding inequalities 5.36 and 5.37 yields:

$$E_1 + E_2 < E_1 + E_2 \tag{5.38}$$

which is a contradiction. It follows that the assumption of $\rho_{\mathrm{GS}}(r)$ not being unique is false. While only the non-degenerate case was proven, this also holds true for degenerate ground states. The proof for the non-degenerate case is not presented here, but can be found in the literature, e.g. in [112].

The Kohn-Sham Equations

The previous sections have outlined the general approach on how to calculate the ground-state density $\rho_{\mathrm{GS}}(r)$, the actual method on how to do it is provided by the Kohn-Sham equations (KSEs). They are derived by employing the variation principle to 5.25:

$$\left[\frac{1}{2}\nabla_i^2 + \nu_{\mathrm{ext}}(r) + \int \frac{\rho(r')}{|r-r'|}d\vec{r}' + \nu_{\mathrm{XC}}\left([\rho(r)], r\right)\right]\varphi_i(r) = H_{\mathrm{KS}}\varphi_i(r) = \epsilon_i\varphi_i(r) \tag{5.39}$$

In conjunction with the relation between the single-electron states φ_i and the electron density $\rho_{\mathrm{GS}}(r)$:

$$\rho_{\mathrm{GS}}(r) = \sum_i |\varphi_i(r)|^2, \tag{5.40}$$

the KSEs can be solved in a self-consistent fashion, as the single-electron states φ_i directly correspond to the electron charge density $\rho(r)$. The KSEs are formally exact, given the exchange-correlation functional E_{XC} (i.e. $\nu_{\mathrm{XC}} = \frac{\partial V_{\mathrm{XC}}}{\partial \rho(r)}$ in equation 5.39) is complete and describes all electron-electron interactions correctly and the electron density can be expressed using equation 5.40. Nevertheless, the single-electron wave functions φ_i and their eigenenergies ϵ_i correspond to an artificial system of independent electrons, substituting the actual system. Hence they are not identical to the real solutions by design. Especially the excited states obtained by this approach are not the excited states of the real system, but are used as such, as the solutions obtained by employing the KSEs are (generally) confirmed by experimental methods, giving validity to the KSEs[96].

5.5 The Self-Consistent Field Cycle

The iterative process of solving the Kohn-Sham equations is called the self-consistent field cycle (SCF-cycle). As already mentioned in the previous section, all single-electron states φ_i are required to calculate the electron density $\rho(\vec{r})$, resulting in a set of coupled differential equations (see equation 5.39), given e.g. a plane wave ansatz is used (see section 5.5.1). To start the SCF-cycle, an initial guess of the starting electron density $\rho_0(\vec{r})$ is required. Obtaining a suitable initial guess can be challenging in itself (e.g. by using the so-called Thomas-Fermi electron density), but can be considered a technical problem rather than a physical one and is hence not discussed within the scope of this thesis. The general steps of (numerically) solving the KSEs are outlined below (see also figure 6.6):

1. Choose a starting electron density $\rho_n(\vec{r}) = \rho_0(\vec{r})$ as an initial guess.
2. Evaluate all terms of the KSE 5.39 in order to compute the Hamiltonian H_{KS} and obtain a set of coupled differential equations.
3. Diagonalize this set of differential equations, which yields a set of single-electron states $\varphi_{i,n}$.

4. Utilizing equation 5.40, update the electron density $\rho_n(\vec{r}) \rightarrow \rho_{n+1}(\vec{r})$.
5. Check for self-consistency, i.e. convergence of $\rho_{n+1}(\vec{r})$ into $\rho_n(\vec{r})$. If self-consistency is reached, the cycle is completed and $\rho_{n+1}(\vec{r}) = \rho_{\mathrm{GS}}(\vec{r})$. Otherwise merge $\rho_{n+1}(\vec{r})$ and $\rho_n(\vec{r})$ and repeat the cycle from step 2[2].

In order to speed up convergence and reduce computation time, several approximations are generally utilized within the SCF-cycle. Some of the most important ones are briefly described hereafter.

5.5.1 The Cut-Off Energy and K-Point Mesh

The single electron states φ_i can be expressed as plane-waves, as for free electrons, plane-waves satisfy the Bloch condition. The general shape of a Bloch wave $\Psi_{\vec{k}}(\vec{r})$ is given by the function:

$$\Psi_{\vec{k}}(\vec{r}) \propto e^{i\vec{k}\vec{r}} u_{\vec{k}}(\vec{r}) \tag{5.41}$$

with $\vec{k}$ being any wave-vector in reciprocal space and $u_k(r)$ being a periodic function with the same periodicity as the respective crystal lattice. The factor $u_k(r)$ can also be expanded into a Fourier-series with the Fourier-coefficients $c_{\vec{k}+\vec{G}}$:

$$u_{\vec{k}}(\vec{r}) = \sum_{\vec{G}} c_{\vec{k}+\vec{G}} \tag{5.42}$$

where $\vec{G}$ are the reciprocal lattice vectors and their multiples. The single electron states can hence be expressed as:

$$\varphi_{i,\vec{k}} = \sum_{\vec{G}} c_{\vec{k}+\vec{G}} e^{i(\vec{k}+\vec{G})\vec{r}} \tag{5.43}$$

In principle infinitely many coefficients need to be considered to exactly reproduce the original single-electron state $\varphi_{i,\vec{k}}$, which is, however, impossible to compute numerically. Luckily, due to the finite number of electrons in a given system, the number of occupied Bloch states i are also finite[3]. Unoccupied states get physically less important with increasing energy, i.e. the Fourier-coefficients $c_{\vec{k}+\vec{G}}$ become smaller for larger $|\vec{G}|^2$. It is hence legitimate to introduce and set a cut-off energy E_{cut}, above which the corresponding Fourier-coefficients are omitted:

$$E_{\mathrm{cut}} = \frac{1}{2}|\vec{G}_{\mathrm{max}}|^2 \tag{5.44}$$

The value of E_{cut} is a parameter of the given DFT implementation and should be chosen in such a way that convergence is achieved, while keeping the computational load manageable. Similar

[2] While sophisticated mixing algorithms exist that accelerate convergence, the so-called Pratt method is the simplest approach: $\rho^*_{n+1}(\vec{r}) = \alpha\rho_{n+1}(\vec{r}) + (1-\alpha)\rho_n(\vec{r})$, where $\alpha \in (0,1)$ is called the mixing parameter.

[3] To be precise: The number of significantly occupied Bloch states is usually finite. For particularly localized states, Bloch states might only be partially occupied, technically requiring infinitely many Bloch states to describe said localized state, while the total number of electrons is still finite.

to how the number of $\vec{G}$-vectors has to be restricted to enable computation, the Brillouin zone has to be sampled as well, as it is not possible to treat it numerically in its continuous form. Hence, the $\vec{k}$-vectors are discretized to create a $\vec{k}$-point mesh. The sufficient resolution of said $\vec{k}$-point mesh depends on the system at hand and is determined and bound by smooth convergence of the SCF-cycle and the available computation power respectively.

5.5.2 Pseudopotentials

The wave function of the electrons can be separated into two areas: Near the nuclei, where the electrons can be considered bound, the wave function oscillates significantly and approaches the behavior of atomic orbitals, as the electrons are almost fully screened from the neighboring atoms and spherical symmetry is (almost fully) restored. In the regions between the nuclei, the wave function is smoother and varies considerably less[114–116]. This oscillatory behavior implies that a large number of $\vec{G}$-vectors are required to achieve sufficient convergence (see section 5.5.1), which increases the computational load of the SCF-cycle. It should further be noted that the KS single-electron states are constructed to be orthogonal to each other, which implies that the wave functions of the valence electrons must also oscillate around the nuclei, not only the wave functions of the core electrons[117]. The idea of pseudopotentials (PPs) is to leverage the screening on the valence electrons, caused by the electrons close to the core, and to substitute the nucleus and the core electrons with a softer, effective potential. Removing the core electrons is motivated by the fact that they are generally quite inert, i.e. most physical phenomena are dominated by the behavior of the valence electrons, which is kept (sufficiently) intact[117]. This principle is schematically depicted in figure 5.3. PPs are often constructed from radial functions and spheric harmonics to obtain an effective potential that reduces the oscillations near the nuclei while keeping the behavior of valence electrons scattering on the nuclei identical. While an in-depth discussion of PPs and their construction can be found in the literature (e.g. in [114]), a short overview of the different types of PPs is presented in the following:

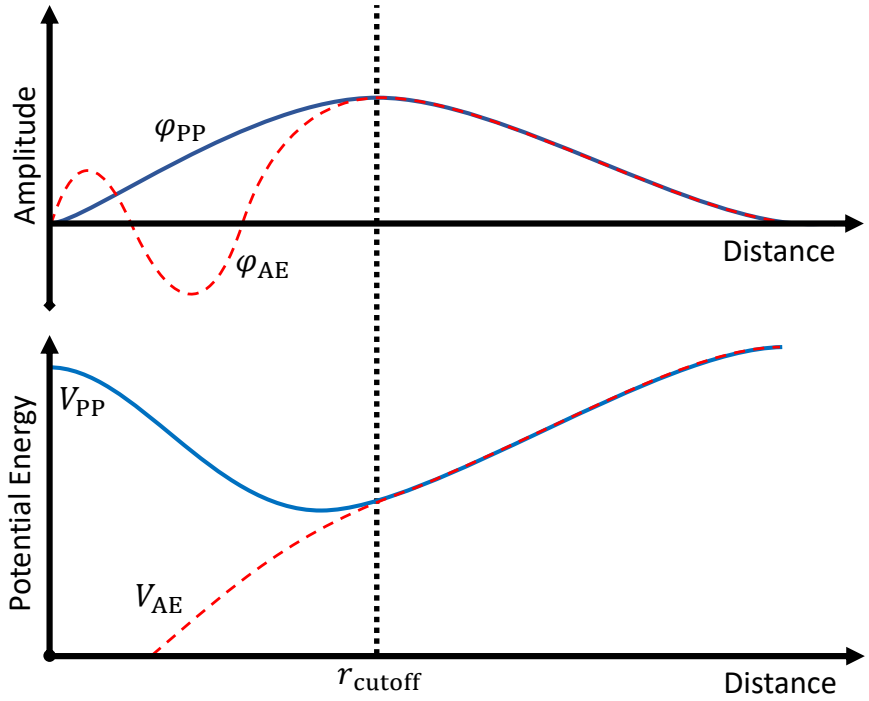


Figure 5.3: Schematic depiction of the different oscillatory behavior of the true, all-electron wave function φ_{AE} and potential V_{AE} and the pseudopotential V_{PP} together with the corresponding pseudo wave-function φ_{PP}. r_{cutoff} denotes the distance from the nucleus where φ_{AE} and φ_{PP} can be considered identical.

Norm-Conserving Pseudopotentials (NC-PPs)

This kind of pseudopotentials is constructed to adhere to the following conditions[118]:

- ◇ Within the cut-off radius r_{cutoff}, the norm of the pseudo-wave functions must equal the norm of the all-electron wave function:

$$\int_{r<r_{cutoff}} \Phi^{AE}_{\vec{R},i}(\vec{r})\,\Phi^{AE}_{\vec{R},j}(\vec{r})\,d\vec{r} = \int_{r<r_{cutoff}} \Phi^{PP}_{\vec{R},i}(\vec{r})\,\Phi^{PP}_{\vec{R},j}(\vec{r})\,d\vec{r} \tag{5.45}$$

  where $\Phi^{AE}_{\vec{R},i}(\vec{r})$ and $\Phi^{PP}_{\vec{R},i}(\vec{r})$ denote the all-electron and pseudo-reference states on the atom at position $\vec{R}$ respectively.

- ◇ Outside the cutoff radius r_{cutoff}, the all-electron wave function and the pseudo-wave function must be identical (also see figure 5.3):

$$\varphi_{PP}(\vec{r}) = \varphi_{AE}(\vec{r}), \qquad |\vec{r}| > r_{cutoff} \tag{5.46}$$

NC-PPs are generally considered to be accurate and transferrable. However, challenges arise for first-row elements and transition metals, as they do not feature any core-electrons that could be removed. Also, reducing the cutoff radius r_{cutoff} in order to increase accuracy can

lead to "harder" potentials, diminishing the speed-up of calculations that they were supposed to provide[115, 116].

Ultrasoft Pseudopotentials (US-PPs)

Ultrasoft pseudopotentials (US-PPs) were initially proposed by David Vanderbilt in 1990[119] and allow for larger cutoff radii r_{cutoff} and generally much "softer" potentials compared to NC-PPs. These benefits are achieved by discarding the conservation of the norm, adding numerical complexity. Using US-PPs, the computational load can be significantly reduced[120].

Projector Augmented Waves (PAWs)

Projector Augmented Waves (PAWs) are a slightly different approach, as they combine an (effectively) all-electron ansatz with the performance gain of PPs. The PAWs are hence not always classified as a PP-method, which results in sometimes inconsistent labeling and description of PP files in practice and in literature. However, as PAWs are widely used, they are discussed in detail in section 5.5.3.

5.5.3 Projector Augmented Wave Method (PAW-Method)

As explained in section 5.5.2, pseudopotentials reduce the computational load of carrying out an SCF-cycle by substituting the real potential created by nuclei and valence electrons by an effective, smoothened (pseudo-)potential. The PPs are constructed so that the true scattering potential is reproduced outside a certain radius. Hence, the KSEs are then solved for the valence electrons only. This of course comes with the downside that information of the wave function close to the nucleus is lost, which can be relevant for certain properties[121]. The projector augmented wave (PAW) method aims to mitigate these drawbacks by using an all-electron wave-function Ψ, which incorporates the core states. This method separates (real-) space into so-called augmentation spheres, which are centered around each atom, and a region outside said spheres, called bonding region. Within the augmentation spheres the wave functions are modeled as somewhat atom-like partial waves, while in the bonding region envelope functions are defined. At the boundary between augmentation spheres and bonding region, partial waves and envelope functions are matched to ensure a smooth transition. Up until this point, this modification is called the augmented-plane-wave method (APW), the PAW method generalizes this approach by combining the APW method with pseudopotentials. This is done employing a linear transformation T to the pseudo wave-function $\tilde{\Psi}$[122]:

$$\Psi(\vec{r}) = T\tilde{\Psi}(\vec{r}), \qquad T = 1 + \sum_a \sum_i \left(|\phi_{i,a}\rangle - |\tilde{\phi}_{i,a}\rangle \right) \langle \tilde{p}_{i,a}| \tag{5.47}$$

where $\tilde{\phi}_{i,a}$ are smoothened all-electron wave-functions that are located at atom i, with the index a denoting the quantum number (n,l,m) to the electron state. $\tilde{p}_{i,a}$ are fixed functions which are called smooth projector functions[121]. $\phi_{i,a}$ are all-electron partial waves that are orthogonal to the all-electron core states and furthermore constructed to be equal to the pseudo wave-functions $\tilde{\phi}_{i,a}$ outside the augmentation spheres. This means that deviations of the pseudo wave-function $\tilde{\phi}_{i,a}$ from the all-electron partial waves $\phi_{i,a}$ within the augmentation sphere are adjusted in the resulting all-electron wave-function. $\tilde{p}_{i,a}$ and $\tilde{\phi}_{i,a}$ additionally satisfy:

$$\sum_i |\tilde{\phi}_{i,a}\rangle \langle \tilde{p}_{i,a}| = 1 \tag{5.48}$$

In summary, three components are required for the transformation T:

- ◇ The partial waves $\phi_{i,a}$
- ◇ The smooth partial (pseudo-) waves $\tilde{\phi}_{i,a}$
- ◇ The smooth projector functions $\tilde{p}_{i,a}$

All functions required to define the transformation are fortunately independent of the system at hand. They can therefore be pre-calculated for all elements of the periodic table individually and the results can be stored for consequent calculations. For a more detailed description in general and on how to calculate the system-independent functions, the reader is referred to the literature, e.g. [123–125].
The standard Kohn-Sham wave functions oscillate (strongly) in some regions, while they are smooth in others. Be employing the PAW method, these wave functions are separated into auxiliary wave functions that are smooth everywhere and additional supporting wave functions that include the (intense) oscillations, although only in small spacial areas, where they are indispensable to describe the system accurately. The PAW method hence significantly reduces the required computational resources, while keeping the description of the system sufficiently accurate, even regarding properties that are influenced by the wave function close to the nucleus.

5.5.4 Exchange-Correlation Functional

As already discussed in chapter 5.2, density functional theory in itself contains the same information as the (solution of the) Schrödinger equation, as no approximations are contained in the approach itself. However, besides the technical approximations required to numerically compute a system with DFT, the absence of the "correct" exchange-correlation functional $V_{\mathrm{XC}}[\rho(r)]$ is the main source of error for DFT calculations. Assuming the correct $V_{\mathrm{XC}}^{\mathrm{Exact}}[\rho(r)]$ would be known and could be employed for calculations, DFT would yield the exact one-electron charge density of the ground state of a given system. Unfortunately, $V_{\mathrm{XC}}^{\mathrm{Exact}}[\rho(r)]$ has not been derived yet. However, approximations to $V_{\mathrm{XC}}^{\mathrm{Exact}}[\rho(r)]$ have been developed, which will briefly be described in this chapter. Motivating, developing, and testing an

exchange-correlation functional is a separate field of research and not the focus of this work. The different flavors of functionals will therefore be described briefly and qualitatively only, taking the perspective of application.
While a lot of specialized functionals have been tailored to best describe a specific (small) group of materials, the most commonly used functionals aim to provide a suitable accuracy for all compounds.

Local Density Approximation (LDA)

LDA describes the simplest exchange-correlation functional, as the exchange-correlation energy density is solely dependent on the particle density at point $\vec{r}$[126]. LDA functionals are exact for an infinite uniform electron gas (UEG). Since most non-model systems display inhomogeneous density distributions, however, these type of functionals fail to describe molecular properties accurately. The exchange part of the LDA functional(s) can be derived analytically[127]:

$$E_{\mathrm{X}}^{\mathrm{LDA}} = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}} \int \rho\left(\vec{r}\right)^{\frac{4}{3}} d\vec{r} \tag{5.49}$$

It should be noted that the different spin components can be considered separately, in which case LDA is more precisely called local spin-density approximation (LSDA)[127]. In contrast to the exchange part, no analytical form exists for the correlation functional. Popular parameterizations like VWN5[128], PW92[129] and PZ81[130] are obtained by employing Quantum Monte Carlo simulation results done by Ceperley and Alder[131], to which the parameters of the correlation functional are fitted[127]. LDA furthermore fulfills the exchange-hole exclusion principle exactly. The exchange-correlation hole is the region around any particle (in this case electrons) within which the probability of finding another identical particle is decreased. In sum total, the reduction of the particle density should amount to 1, meaning one particle (electron) should be removed, which corresponds to the Pauli exclusion principle. While the exact shape of the exchange-hole is inaccurate within LDA, the spherical mean remains correct[126].

Generalized Gradient Approximation (GGA)

To improve on the shortcomings of LDA, the generalized gradient approximation (GGA) also considers the gradient of the electron density $\nabla\rho$. This allows the GGA to also account for inhomogeneities in the electron density ρ and tends to provide a significant improvement compared to LDA[127]. While this improvement is achieved for most systems, it is not universal, as it stems from the cancellation of errors, which is system-dependent[126]. The exchange functional of the GGA can be expressed as:

$$E_{\mathrm{X}}^{\mathrm{GGA}} \propto \int e_{\mathrm{x}}^{\mathrm{UEG}} g_{\mathrm{x}}^{\mathrm{GGA}} d\vec{r} \tag{5.50}$$

where the function $g_{\mathrm{x}}^{\mathrm{GGA}}$ is an inhomogeneity correction factor (ICF), enhancing the UEG exchange energy density $e_{\mathrm{x}}^{\mathrm{UEG}}$[127]. Multiple correction functions $g_{\mathrm{x}}^{\mathrm{GGA}}$ have been developed over the years. A simple, but commonly used ICF takes the form of[132]:

$$g_{\mathrm{x}}^{\mathrm{GGA}} = 1 + c_{x,l}\frac{\gamma\xi^2}{1+\gamma\xi^2}, \qquad \xi = \frac{|\nabla\rho(\vec{r})|}{\rho(\vec{r})^{\frac{4}{3}}} \tag{5.51}$$

where γ is a non-linear parameter and $c_{x,l}$ a constant[127, 133]. One of the most commonly used GGA exchange-correlation functionals was proposed by John P. Perdew, Kieron Burke and Matthias Ernzerhof in 1996 and was hence coined PBE-GGA. It utilizes the ICF given in 5.51, as well as parameter-free correlation functionals, meaning no fitting to experimental results is required[134]:

$$E_{\mathrm{C}}^{\mathrm{GGA}} = \int \rho(\vec{r})\left[e_{\mathrm{c}}^{\mathrm{UEG}} + H^{\mathrm{GGA}}(t)\right]d\vec{r} \tag{5.52}$$

$$H^{\mathrm{GGA}}(t) = \frac{\beta^2}{2\alpha} ln\left[1 + \frac{2\alpha}{\beta}\frac{t^2+At^4}{1+At^2+A^2t^4}\right], \quad A = \frac{2\alpha}{\beta\cdot exp\left(-2\alpha e_{\mathrm{c}}^{\mathrm{UEG}}/\beta^2\right)-1} \tag{5.53}$$

where:

$$\alpha \approx 0.0716, \quad \beta \approx 0.0667, \quad t = \frac{|\nabla\rho(\vec{r})|}{2k_{\mathrm{s}}\rho(\vec{r})}, \quad k_{\mathrm{s}} = \sqrt{\frac{4k_{\mathrm{F}}}{\pi}}, \quad k_{\mathrm{F}} = \sqrt[3]{3\pi^2\rho(\vec{r})} \tag{5.54}$$

Its popularity originates from its overall good performance for most systems. Other popular GGA functionals include for example PBEsol[135] and PW91[136].

5.6 Electronic Structure Calculations

So far, DFT has only been described as a method to derive the ground-state charge density for a given system. In practice, the distribution of charge, i.e. electrons, in space is rarely the property of interest. The energetic distribution of electrons might be more insightful. This electronic structure information is frequently used to investigate a system in terms of its electronic conductivity or optical properties.

Band Structure

The electronic band structure is a paramount concept in solid-state physics, as it can be used to investigate optical transitions, the size of the band gap E_{G}, the type of band gap, i.e. direct or indirect and the effective masses of electrons and holes[137]. As shown in chapter 5.4, the single-electron wave functions $\varphi_{i,\vec{k}}$ along with the corresponding eigenenergies $\epsilon_{i,\vec{k}}$ can be obtained by means of DFT (see equations 5.39 and 5.43). This means that for each point $\vec{k}$ in reciprocal space, a set of eigenenergies $\epsilon_{i,\vec{k}}$ exists, yielding an energy dispersion relation $E(\vec{k})$. Technically speaking, as DFT calculations are conducted for a discretized reciprocal space

only, this dispersion relation is also discretized and a sufficiently dense $\vec{k}$-point mesh must be chosen to obtain smooth (and well converged) results. The dispersion relation $E(\vec{k})$ is furthermore a 4-dimensional property, which makes it challenging to visualize. Ultimately, $E(\vec{k})$ is only plotted for the high-symmetry points in the first Brillouin zone, as well as the connecting edges to mitigate this complication. The single-electron wave functions $\varphi_{i,\vec{k}}$, as computed by (standard) DFT, are only suited to obtain the ground state charge density of the system, and hence also only ground-state properties can be derived from it. In general, the single-particle wave functions $\varphi_{i,\vec{k}}$ and their eigenenergies $\epsilon_{i,\vec{k}}$ are still linked to a fictitious system with independent electrons, substituting for the real one. Hence, $\varphi_{i,\vec{k}}$ and $\epsilon_{i,\vec{k}}$ do not represent the real physical system directly. Still, the solutions generated by this approach reproduce experimental findings reasonably well. Yet, the obtained band structure should not be taken as exact by default. Its correctness should be evaluated on a case by case basis. For example, band structures obtained by DFT tend to underestimate the band gap, even closing it for some small bandgap semiconductors[96, 101, 138–140].

Density of States

The density of states (DOS) describes the number of electron states within a certain energy interval dE (at energy E). It is often evaluated in conjunction with the band structure, as it provides similar insights, but omits the spatial resolution in reciprocal space. The DOS can be computed by counting all electron states within a given energy range:

$$\mathrm{DOS}(E) = \sum_i g_i(E), \qquad g_i(E) = \int_{V_{\mathrm{UC}}} \frac{1}{4\pi^3} \delta(E - \epsilon_{i,\vec{k}}) d\vec{k} \tag{5.55}$$

where $g_i(E)$ is the DOS of the i-th state, or rather band in this context[141]. To numerically compute the DOS, (finite) energy intervals with width dE are constructed, within which the number of states are then counted. This makes the (numerically obtained) DOS susceptible to the choice of interval width dE and the $\vec{k}$-point mesh density, which must be chosen accordingly. Integrating the DOS up to a certain energy E' yields the IDOS, which denotes the number of electrons that are implemented into the system up to this energy:

$$\mathrm{IDOS}(E') = \int_{-\infty}^{E'} \mathrm{DOS}(E) dE \tag{5.56}$$

While the shapes of the band structure, DOS and the IDOS already provide insights into the electronic structure of a given system, resolving these properties by orbital contribution enhances the information gain substantially.

Orbital Projection

In order to investigate the orbital contributions to the properties listed above, the single-electron wave functions $\varphi_{i,\vec{k}}$, as obtained by the KS-equations, must be projected unto the atomic orbitals $\phi_{n,\vec{k},\lambda,i}$, or rather expressed in the basis of the atomic orbitals:

$$|\varphi_{i,\vec{k}}\rangle = \sum_{n,\lambda} \langle \phi_{n,\vec{k},\lambda} | \varphi_{i,\vec{k}} \rangle \cdot |\phi_{n,\vec{k},\lambda}\rangle \tag{5.57}$$

along with the corresponding energy contributions (see equation 5.39):

$$\epsilon_{n,\vec{k},\lambda,i} = \langle \phi_{n,\vec{k},\lambda} | \, H_{KS} \, | \varphi_{i,\vec{k}} \rangle \tag{5.58}$$

where $|\phi_{n,\vec{k},\lambda}\rangle$ denotes the atomic orbital of site n (λ=s, p_x, p_y, p_z, d_{xy}, d_{yz}, d_{xz}, $d_{x^2\text{-}y^2}$, d_{z^2} etc.). The atomic orbitals can be precalculated for each element of the periodic table, and are often provided within the pseudopotential used for the calculation. With these expressions, each energy value, either in the band structure, DOS, or IDOS, can be dispersed into the separate orbital contributions. While it would be possible to denote the contribution of each orbital of each site, the contributions of identical elements are usually summed up for reasons of clarity. The same holds true for the quantum number m_l, i.e. p_x, p_y and p_z for example are not shown separately, but rather as a sum. Nevertheless, which contributions are insightful depend on the system to be investigated and can be chosen as seen fit.
It should be noted that the orbital projections have to be checked for their validity. Especially between atoms (where bonding happens) the atomic orbitals are generally not orthogonal, meaning the projection can flawed. Some DFT implementations (e.g. LOBSTER, see section 7.3) employ additional methods to mitigate this issue.

5.7 Optical Properties

The dielectric function $\epsilon(\omega)$ is an important property, as it describes how a solid interacts with light (of frequency ω). This interaction can take the form of absorption, transmission, and reflection, all of them being present at the same time. The dielectric function $\epsilon(\omega)$ is a complex quantity:

$$\epsilon(\omega) = \epsilon_1(\omega) + i\epsilon_2(\omega) \tag{5.59}$$

Given that the full dielectric function $\epsilon(\omega)$ is known, optical properties can be derived, for example the reflectivity $R(\omega)$ (for normal incidence):

$$R(\omega) = \left| \frac{\sqrt{\epsilon(\omega)} - 1}{\sqrt{\epsilon(\omega)} + 1} \right|^2 \tag{5.60}$$

the energy-loss spectrum $L(\omega)$:

$$L(\omega) = \frac{\epsilon_2(\omega)}{\epsilon_1^2(\omega) + \epsilon_2^2(\omega)} \tag{5.61}$$

and the extinction coefficient $k(\omega)$ and refractive index $n(\omega)$, respectively:

$$k(\omega) = \left[\frac{\sqrt{\epsilon_1^2(\omega) + \epsilon_2^2(\omega)} - \epsilon_1(\omega)}{2} \right]^{\frac{1}{2}} \tag{5.62}$$

$$n(\omega) = \left[\frac{\sqrt{\epsilon_1^2(\omega) + \epsilon_2^2(\omega)} + \epsilon_1(\omega)}{2} \right]^{\frac{1}{2}} \tag{5.63}$$

With the extinction coefficient $k(\omega)$, the absorption $A(\omega)$ and transmission $T(\omega)$ can be expressed as:

$$A(\omega) = [1 - R(\omega)] \left(1 - e^{-\frac{4\pi k(\omega)}{\lambda_0} x} \right) \tag{5.64}$$

$$T(\omega) = [1 - R(\omega)]\, e^{-\frac{4\pi k(\omega)}{\lambda_0} x} \tag{5.65}$$

where λ_0 is the vacuum wavelength and x the penetration depth, fulfilling the relation $A + T + R = 1$.
Within this chapter, only a rudimentary motivation of the derivation of the dielectric function will be presented. The imaginary part of the dielectric function $\epsilon_2(\omega)$ is connected to the band structure, which can be understood by considering its connection to Fermi's golden rule within the linear-response theory (one-electron picture)[101, 142, 143]:

$$\epsilon_2(E) \propto \sum_{v,c} |\vec{\eta}_{\mathrm{P}} \langle c| \vec{r} |v\rangle|^2 \delta(E_c - E_v - E), \qquad E = \hbar\omega \tag{5.66}$$

where $\vec{\eta}_{\mathrm{P}}$ is the polarization vector of the impinging light and $\langle c| \vec{r} |v\rangle$ the transition dipole matrix element. It quantifies whether a transition is possible under electric dipole interaction, i.e. it incorporates the parity selection rule for dipole transitions into the equation. $|v\rangle$ is the initial valence band state with energy E_v before excitation, $|c\rangle$ the final conduction band state with energy E_c after excitation. $E = \hbar\omega$ describes the energy of the photon. Equation 5.66 can be understood as a summation over all possible pairs of states in the band structure, where the initial state $|v\rangle$ is occupied, while the final state $|c\rangle$ is empty, and their energy difference is equal to the energy of the photon $E = E_c - E_v$. An expression has been derived to obtain the imaginary part of the dielectric function in the context of DFT within full-potential codes (meaning all electrons are included, e.g. the PAW methodology (see section 5.5.3))[144]:

$$\epsilon_2^{(\alpha\beta)}(\omega) = \frac{4\pi^2 e^2}{\Omega} \lim_{q\to 0} \frac{1}{q^2} \sum_{c,v,\vec{k}} 2 w_{\vec{k}} \delta(E_{c,\vec{k}+\vec{q}} - E_{v,\vec{k}} - \hbar\omega) \cdot \langle u_{c,\vec{k}+e_\alpha q} | u_{v,\vec{k}} \rangle \langle u_{v,\vec{k}} | u_{c,\vec{k}+e_\beta q} \rangle \tag{5.67}$$

where $|u_{n,\vec{k}}\rangle$ is the cell periodic part of the total (pseudo)-wave-function $|\Psi_{n,\vec{k}}\rangle = e^{i\vec{k}\vec{r}}|u_{n,\vec{k}}\rangle$, Ω the volume of the primitive unit cell, $w_{\vec{k}}$ the $\vec{k}$-point weights with $\sum_{\vec{k}} w_{\vec{k}} = 1$ and e_α, e_β are unit vectors for the three cartesian directions. This way, $\epsilon_2^{(\alpha\beta)}(\omega)$ is defined as a 3x3 Cartesian tensor (as the dielectric function in real solids is not necessarily isotropic).
The real part is then obtained by employing the Kramers-Kronig transformation:

$$\epsilon_1^{(\alpha\beta)}(\omega) = 1 + \frac{2}{\pi} P \int_0^\infty \frac{\epsilon_2^{(\alpha\beta)}(\omega')\omega'}{\omega'^2 - \omega^2 + i\eta} d\omega' \tag{5.68}$$

where P is the Cauchy principal value of the integral and η a complex shift (that can be set to zero however)[144]. It should be noted that this approach only includes electronic contributions to the dielectric function. By employing the Born-Oppenheimer approximation (see section 5.2), nuclei are considered static, hence phonon excitations and subsequently phononic contributions to the dielectric function are omitted.
By calculating the orbital contributions of the initial and final state of equation 5.66 (and in extension equation 5.67), an orbital-resolved version of the imaginary part of the dielectric function $\epsilon_2(\omega)$ can be derived. For example, assuming at the Γ-point $\vec{k} = \vec{0}$ a transition occurs from initial state $|v_1\rangle$ to final state $|c_1\rangle$, with

$$|v_1\rangle = \begin{cases} 50\% \text{ s Orbital} \\ 50\% \text{ p Orbital} \\ 0\% \text{ d Orbital} \end{cases} \quad |c_1\rangle = \begin{cases} 0\% \text{ s Orbital} \\ 50\% \text{ p Orbital} \\ 50\% \text{ d Orbital} \end{cases} \tag{5.69}$$

the contribution to $\epsilon_2(\omega)$ can be calculated exactly, as a separate transition matrix element for each transition (e.g. s→p state) is obtained. However, this exact procedure increases the computation time and required storage exponentially. It is (without proof) reasonable to approximate this concept by calculating the full matrix element without orbital contribution $|v_1\rangle \rightarrow |c_1\rangle$ and subsequently assigning the orbital contributions based on the contributions to the initial and final states $|v_1\rangle$ and $|c_1\rangle$. In the case of the numbers stated in 5.69, this would result in the transition (matrix element) to be assigned being:

$$\text{Transition} = \begin{cases} 25\% \text{ s} \rightarrow \text{p Transition} \\ 25\% \text{ s} \rightarrow \text{d Transition} \\ 25\% \text{ p} \rightarrow \text{p Transition} \\ 25\% \text{ p} \rightarrow \text{d Transition} \end{cases} \tag{5.70}$$

Within the scope of this thesis, this method is the one employed (see also chapter 7.3).

Joint Density of States

The joint density of states (JDOS) describes the amount of available transitions for a given energy E_γ, i.e. the number of pairs of occupied valence and unoccupied conduction states for

a given $\vec{k}$-point separated by the energy E_γ[145]:

$$\text{JDOS}(\omega) = \sum_{\sigma} \sum_{n \in \text{V}} \sum_{n' \in \text{C}} \frac{\Omega}{(2\pi)^3} \int \delta(E_{\vec{k},n'} - E_{\vec{k},n} - E_\gamma) d\vec{k}, \quad E_\gamma = \hbar\omega \tag{5.71}$$

with Ω the volume of the lattice cell, σ the spin component and n / n' the summation indices for the valence / conduction states respectively.

5.8 Phase Information in DFT and Chemical Bonding

Chemical bonding is often described by the linear combination of atomic orbitals (LCAOs). Fig 5.4 illustrates this concept using the H_2 molecule, where two s orbitals either form a bonding σ, or an antibonding σ^* molecular orbital. Whether a bonding or antibonding orbital is formed depends on whether the atomic orbitals interfere constructively or destructively, which is decided by their phase difference[146]:

$$|\sigma\rangle \propto |s_\text{A}\rangle + |s_\text{B}\rangle \tag{5.72}$$

$$|\sigma^*\rangle \propto |s_\text{A}\rangle - |s_\text{B}\rangle \tag{5.73}$$

In the bonding case, i.e. when atomic orbitals or orbital lobes with the same sign interact, the electron density between the nuclei is increased, reducing the electrostatic repulsion between them. In the antibonding case, however, the internuclear electron density decreases, forming a node where the density becomes zero. Hence, the electrons are mainly located outside the internuclear region and the nuclei electrostatically repel each other[146].

In solids, a large number of orbitals are aligned and can interact with each other, as shown in figure 5.5, leading to energetic bands rather than discrete energy levels as found in molecules. Assuming s orbitals, the energetically lowest crystal orbital $|\Phi_0\rangle$ within such a band consists of constructively interfering atomic orbitals only (totally bonding, no sign changes), while the energetically highest crystal orbital $|\Phi_N\rangle$ features the strongest possible destructive interference (totally antibonding, maximum number of sign changes)[146]:

$$|\Phi_0\rangle \propto |s_1\rangle + |s_2\rangle + |s_3\rangle + ... |s_N\rangle \tag{5.74}$$

$$|\Phi_N\rangle \propto |s_1\rangle - |s_2\rangle + |s_3\rangle - |s_4\rangle + ... + |s_{N-1}\rangle - |s_N\rangle \tag{5.75}$$

where N is the number of sites (assuming N is even). The intermediate states can be expressed as:

$$|\Phi_n\rangle \propto \sum_{j=1}^{N} cos\left(\frac{n(j-1)\pi}{N}\right) |s_j\rangle \tag{5.76}$$

In general, s and p_π bonds are fully bonding and are energetically lowest for no sign changes and become antibonding with an increasing number of sign changes. The opposite holds true

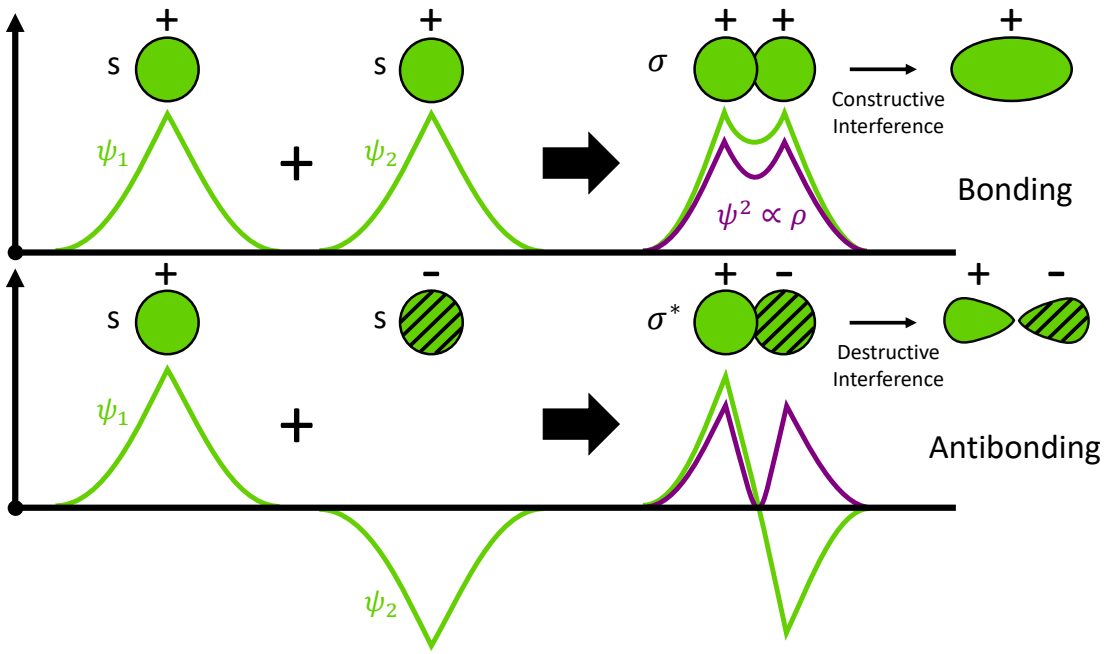


Figure 5.4: Linear combination of atomic orbitals (LCAOs) for the H_2 molecule. The wave functions Ψ are depicted in green, the electron density $\Psi^2 \propto \rho$ in purple. Top: The wave functions interfere constructively (identical sign of Ψ), increasing electron density between the nuclei which leads to a bonding configuration. Bottom: The wave functions interfere destructively (opposite sign of Ψ) and form a node between the nuclei, decreasing the electron density, which leads to an antibonding configuration. Adapted from [146].

for p_σ bonds. The ratio of sign changes can also be linked to the wave vector $\vec{k}$ in reciprocal space (see figure 5.5):

$$\vec{k} \propto \left(\frac{n}{N}, \frac{m}{M}, \frac{l}{L} \right), \tag{5.77}$$

where n, N, m, M and l, L are the number of sign flips and number of sites in the cardinal directions respectively. Assuming that N, M and L are sufficiently large, these ratios can be considered continuous parameters. Figure 5.6 exemplies this concept using the band structure of cubic antimony. At the Γ-point, corresponding to $\vec{k} = (0, 0, 0)$, the s band exhibits an energetic minimum, as it is in the optimal bonding configuration due to the absence of changes among the orbitals. Moving away from the Γ-point towards the R-point at $\vec{k} = (0.5, 0.5, 0.5)$, the energy of the s band increases. The opposite holds true for the p bands, which decrease in energy outside of the Γ-point, if they are forming σ bonds[146].

The concept of phase is crucial to explain bonding behavior within the LCAO framework presented above, as the bonding and antibonding states are constructed by knowing which kind of interference lowers the total energy. Within the framework of DFT, the phase is treated implicitely, as it is seemingly omitted when squaring the wave function to obtain the electron density ρ. However, as seen in figure 5.4, the electron density is directly related to how the states interfere. Using the energy functional $E[\rho]$, DFT determines the ground state density

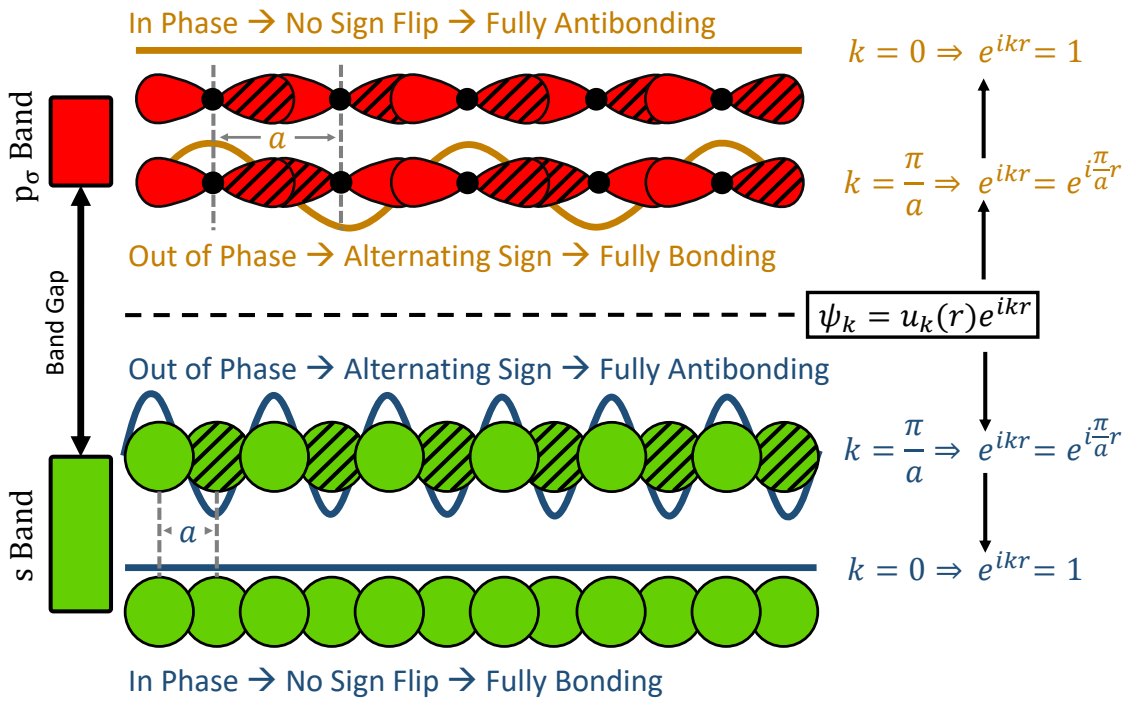


Figure 5.5: Bands formed by bonding and antibonding s and p_σ orbitals. For the s orbitals, identical signs lead to bonding behavior. For the p orbitals, when forming σ bonds, alternating signs lead to bonding behavior. For p orbitals forming π bonds, identical signs are again required for bonding behavior (assuming p orbitals shown would overlap vertically). As the electrons form Bloch states $\Psi_k(r) = u_k(r)e^{ikr}$ (1D case), the Γ-point at $k = 0 = 2\pi/\lambda$ corresponds to an infinite wavelength λ. Thus, all orbitals are in phase and no sign flip occurs. Conversely, for $k = \pi/a$ (corresponding to the border of the Brillouin zone), where a is the lattice constant, an alternating sign change is achieved. Adapted from [146].

ρ_{GS} with the lowest energy, corresponding to the bonding state (see section 5.5). As stated by the Hohenberg-Kohn theorem (see section 5.3), all properties are defined by the ground state density ρ_{GS}. Hence, as long as the correct density ρ_{GS} is obtained, the specific details of how it was determined is not relevant (as shown by the correct description of the band structure in figure 5.6). This can also be seen from the fact that the correct band structure can by obtained just by a simple tight binding approach, which is compatible with the DFT calculations[147]. While no absolute phase information can be restored from the electron density, all physical effects caused by interference, and thus the relative phase difference, are clearly contained within the DFT approach.

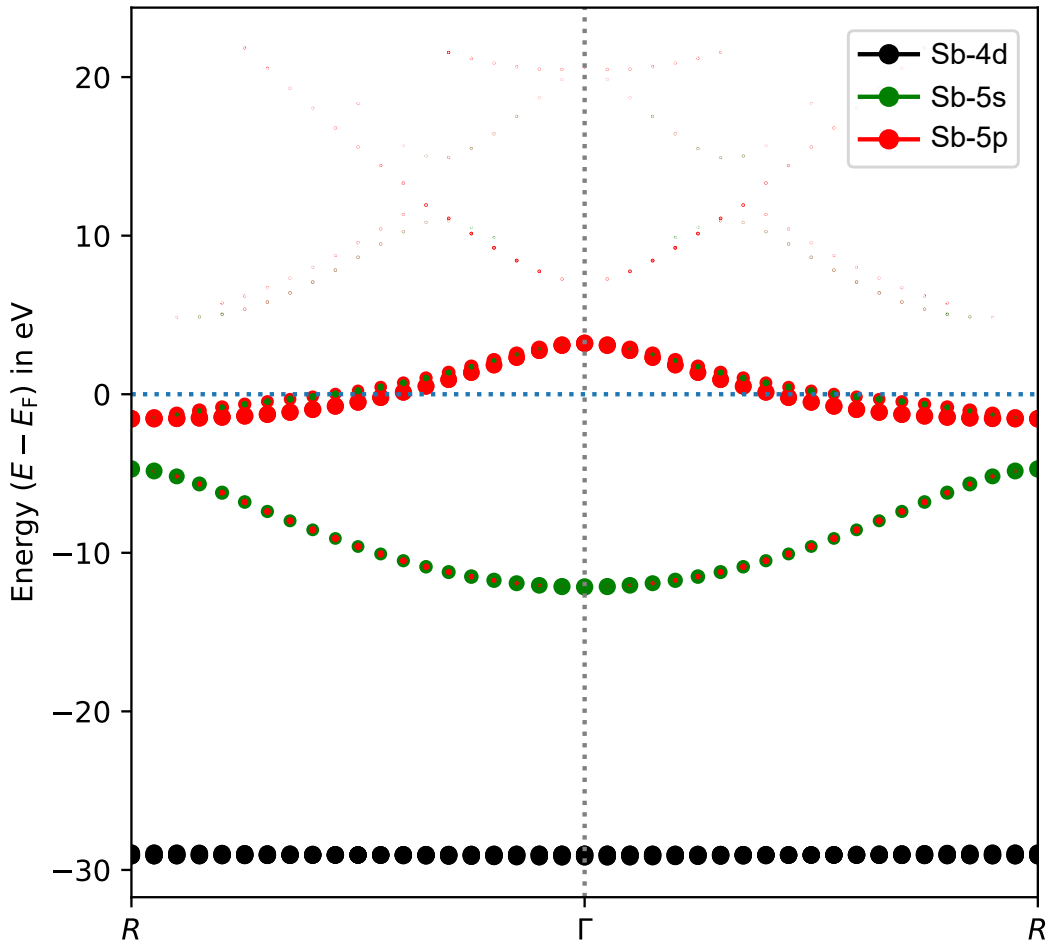


Figure 5.6: Orbital-resolved band structure of cubic Sb, obtained by DFT (see also chapter 7.1.3). The symbol size reflects the contribution of the respective state to a band at a given $\vec{k}$-point. At the Γ-point, where $\vec{k} = (0, 0, 0)$, the s band exhibits an energetic minimum, as it is in the optimal bonding configuration due to the absence of (sign) changes among the orbitals. Moving away from the Γ-point towards the R-point at $\vec{k} = (0.5, 0.5, 0.5)$, the energy of the s band increases. The opposite holds true for the p bands, which decrease in energy outside of the Γ-point, if they are forming σ bonds (see figure 5.5). The 4d orbitals do not overlap, forming a flat band.

Chapter 6

Quantum Theory of Atoms in Molecules

With the advent of quantum mechanics, the fields of physics and chemistry grew closer together. This novel description created a mathematical framework for both fields to explain and predict the behavior of matter, their common subject of interest. Originally, physics and chemistry approached the goal of understanding matter from a different angle. Chemistry focused on the study of substances, their interactions with other substances or energy and their synthesis and decomposition. Physics on the other hand was more concerned with unravelling the underlying principle of the behavior of matter, and how it relates to the fundamental laws of nature, i.e. its motion through space and time as well as its interaction with external and internal forces.
Quantum mechanics opened up a different perspective on matter, enabling both fields to look at it from a novel, but now shared perspective. However, as chemical models developed and utilized before the emergence of quantum mechanics, such as aromaticity, bonding type or order, do not have uniquely defined operators that can be applied to the wave function directly, an interpretation of quantum mechanics had to be developed in order to enable correlation with these fundamental properties of chemistry. One of these approaches is the quantum theory of atoms in molecules (QTAIM), developed by Richard Bader since the 1960s. QTAIM utilizes the electron density to define basins around atoms and consequently derives chemical properties from these so-called Bader basins[15–18]. Since the 1990s, significant advancements have been made to transfer the methods from the QTAIM framework towards the application in solid crystals, while it was originally developed for molecules and atoms, as the name suggests.

The Bader Basin

As already mentioned in the preceding paragraph, the fundamental building block of QTAIM is the quantum mechanical definition of the atom in a compound. At first glance, defining an atom seems to be trivial, as determining (or even setting) the position of each nucleus is generally relatively easy. However, assigning each point of the area between each nucleus to belong to a specific atom is not as straightforward, as no obvious criterion exists. While the geometric distance might appear to be a viable approach, the inequality of atoms in terms of size and chemical behavior renders this ansatz questionable at best. The approach of Richard Bader utilizes the topology of the electron density to define basins corresponding to atoms in the solid.

An atom is defined as the region that is enclosed by surfaces through which the flux in the gradient of the electron density $\vec{\nabla}\rho(\vec{r})$ is zero[17]:

$$\vec{\nabla}\rho(\vec{r}_s) \cdot \vec{n}(\vec{r}_s) = 0, \tag{6.1}$$

where $\vec{r}_s$ is a point on said zero-flux surface, and $\vec{n}(\vec{r}_s)$ is the unit vector normal to the surface at $\vec{r}_s$. The domain of an atom A is denoted by Ω_A in conjunction with the domain shape function $\omega^{\Omega_A}(\vec{r})$ with:

$$\omega^{\Omega_A}(\vec{r}) = \begin{cases} 1 & \vec{r} \in \Omega_A \\ 0 & \vec{r} \notin \Omega_A \end{cases} \tag{6.2}$$

Figure 6.1 depicts a 2D representation of the charge density along with zero-flux lines, while figure 6.2 visualizes the shape of Bader basins in 3D for gold and sodium bismuthide.

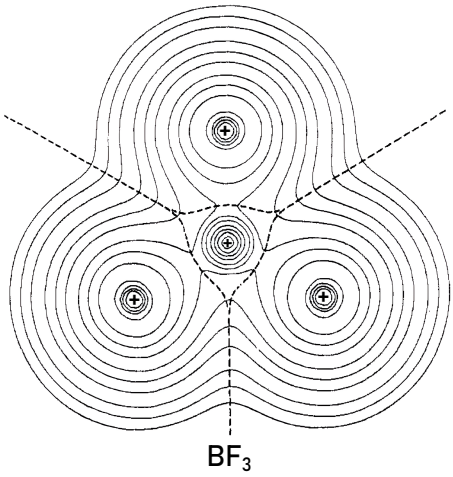


Figure 6.1: 2D contour of the charge density with zero-flux (dashed) lines of BF_3. Adapted with permission from [15]. Copyright 2023 American Chemical Society.

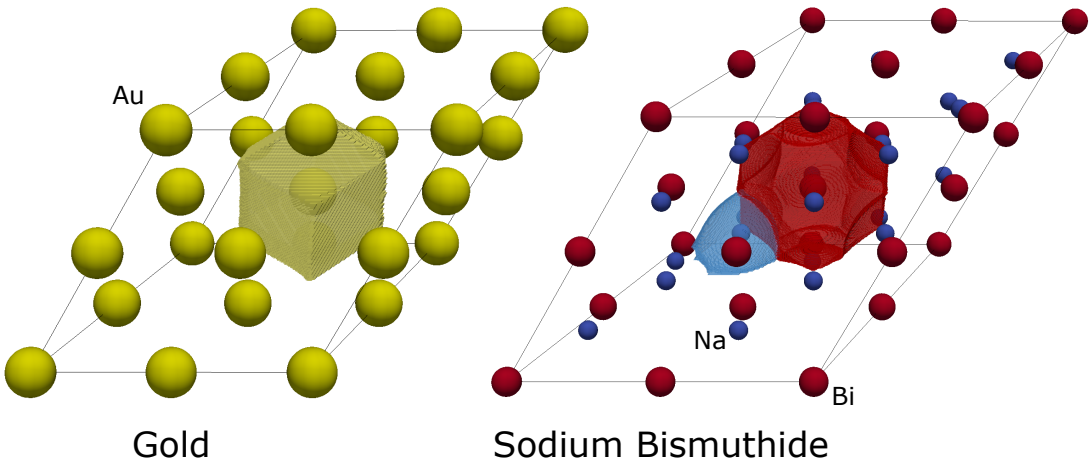


Figure 6.2: Bader basins of gold (Au) and sodium bismuthide (Na_3Bi). Computed using DGRID, as described in chapter 7.2.

6.1 Electron Population, Localization and Delocalization Index

Using the QTAIM approach, various (quantum-chemical) properties can be derived. A selection will be presented in this section. All information is taken from [3, 149], if not stated otherwise. Given the electron charge density (see chapter 5) as well as the spatial separation into Bader basins is known (as illustrated by figure 6.3, top), the easiest property to derive is the average electron population N_A, which is obtained by the integration of the electron density $\rho(\vec{r})$ over the respective basin Ω_A (also see figure 6.3, middle):

$$N_A = \langle n_A \rangle = \int_{\Omega_A} \rho(\vec{r}) d\vec{r} \tag{6.3}$$

where n_A is the electron distribution within basin Ω_A[150]. By subtracting the nominal charge Z_A of a free reference atom, the total number of electrons transferred (TET) to or from the atom can be obtained:

$$\mathrm{TET} = N_A - Z_A \tag{6.4}$$

The localization index (LI) λ_A and the delocalization index (DI) $\delta_{A,B}$ can be expressed by the variance and covariance of the atomic populations:

$$\lambda_A = N_A - \mathrm{Var}(n_A) = \langle n_A \rangle - \left(\langle n_A^2 \rangle - \langle n_A \rangle^2 \right) \tag{6.5}$$

$$\delta_{A,B} = -2 \cdot \mathrm{Cov}(n_A, n_B) = -2 \cdot \left(\langle n_A n_B \rangle - \langle n_A \rangle^2 \right) \tag{6.6}$$

Equations 6.5 and 6.6 imply that $N_A \geq \lambda_A \geq 0$ and $\delta_{A,B} \geq 0$, while for the limiting case of complete localization $\lambda_A = N_A$ and $\delta_{A,B} = 0$. Furthermore the relation:

$$N_A = \lambda_A + \frac{1}{2} \sum_{B \neq A} \delta_{A,B} \tag{6.7}$$

is fulfilled, which means that the average population of each atom (or basin) N_A can be divided into localized (first term) and shared electrons (second term). The average values of equations 6.5 and 6.6 are calculated by considering the one- and two-particle probability

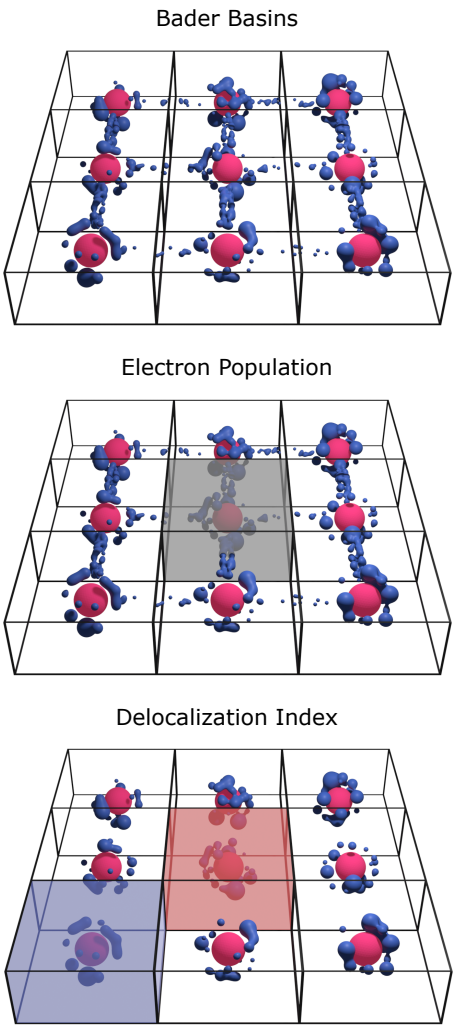


Figure 6.3: Simplified depiction of a solid. The red spheres depict the atom cores, while the blue clouds represent the electron charge. Top: Bader basins Ω represented by black boxes. Middle: The center basin is highlighted as the charge density within in is integrated to obtain the average electron population N_A of the basin (see equation 6.3). Bottom: Two basins highlighted as the charge densities are integrated to obtain the delocalization index (see equation 6.12). Illustration by Jean Felix Dushimineza, also shown in [148].

densities, which are the electron density $\rho(\vec{r})$ and the (electron) pair density $\pi(\vec{r}_1, \vec{r}_2)$ in this case. The pair density $\pi(\vec{r}_1, \vec{r}_2)$ can be understood as the probability of finding an electron at $\vec{r}_1$, given another electron is located at $\vec{r}_2$. It is normalized to the total number of possible electron pairs:

$$\int\int \pi(\vec{r}_1, \vec{r}_2) d\vec{r}_1 d\vec{r}_2 = N(N-1) \tag{6.8}$$

with N being the total number of electrons in the system. It can be expressed as:

$$\pi(\vec{r}_1, \vec{r}_2) = \rho(\vec{r}_1)\rho(\vec{r}_2) - \rho_{\mathrm{XC}}(\vec{r}_1, \vec{r}_2) = \int ... \int |\Psi_{\mathrm{KS}}(\vec{r}_1, ..., \vec{r}_{N_\mathrm{e}})|^2 d\vec{r}_3 ... d\vec{r}_{N_\mathrm{e}} \tag{6.9}$$

where $\rho_{XC}(\vec{r}_1, \vec{r}_2)$ is the exchange-correlation density, which denotes the deviation of the pair density from the independent-electron distribution. All non-classical correlations not contributing to the probability distribution of two electrons are subsumed in this property, e.g. the Pauli exclusion principle and the Coulomb correlation:

$$\rho_{\mathrm{XC}}(\vec{r}_1, \vec{r}_2) = \gamma(\vec{r}_1, \vec{r}_2)\gamma(\vec{r}_2, \vec{r}_1), \qquad \gamma(\vec{r}_1, \vec{r}_2) = \sum_{i_{\mathrm{occ.}}} \varphi_i^*(\vec{r}_1)\varphi_i(\vec{r}_2) \tag{6.10}$$

The DI and LI values can then be obtained by integration:

$$\mathrm{LI} = \lambda_A = \int_{\Omega_A}\int_{\Omega_A} \rho_{\mathrm{XC}}(\vec{r}_1, \vec{r}_2) d\vec{r}_1 d\vec{r}_2 = \int_{\Omega_A}\int_{\Omega_A} \left[\rho(\vec{r}_1)\rho(\vec{r}_2) - \pi(\vec{r}_1, \vec{r}_2)\right] d\vec{r}_1 d\vec{r}_2 \tag{6.11}$$

$$\mathrm{DI} = \delta_{A,B} = \int_{\Omega_A}\int_{\Omega_B} \rho_{\mathrm{XC}}(\vec{r}_1, \vec{r}_2) d\vec{r}_1 d\vec{r}_2 + \int_{\Omega_B}\int_{\Omega_A} \rho_{\mathrm{XC}}(\vec{r}_1, \vec{r}_2) d\vec{r}_1 d\vec{r}_2 \tag{6.12}$$

$$\mathrm{DI} = \delta_{A,B} = 2\int_{\Omega_A}\int_{\Omega_B} \rho_{\mathrm{XC}}(\vec{r}_1, \vec{r}_2) d\vec{r}_1 d\vec{r}_2 = 2\int_{\Omega_A}\int_{\Omega_B} \left[\rho(\vec{r}_1)\rho(\vec{r}_2) - \pi(\vec{r}_1, \vec{r}_2)\right] d\vec{r}_1 d\vec{r}_2 \tag{6.13}$$

utilizing the symmetry of the exchange-correlation density $\rho_{\mathrm{XC}}(\vec{r}_1, \vec{r}_2)$. In some sources $\rho_2(\vec{r}_1, \vec{r}_2) = \frac{1}{2}\pi(\vec{r}_1, \vec{r}_2)$ is used, due to a different choice of normalization[1]. The integration over two basins to obtain the DI is visualized in figure 6.3 (bottom). Alternatively, DI can be expressed by means of the domain overlap matrix (DOM) $S_{i,j}(\Omega)$, using the KS single-electrons wave functions $\varphi_i(\vec{r})$:

$$S_{i,j}(\Omega) = \int_{\Omega} \varphi_i^*(\vec{r})\varphi_j(\vec{r}) d\vec{r} \tag{6.14}$$

The delocalization is then given by:

$$\mathrm{DI} = \delta_{A,B} = 2\sum_{i,j} \Theta_i \Theta_j S_{i,j}(\Omega_A) S_{i,j}(\Omega_B) \tag{6.15}$$

where Θ_i is the occupation number of state i (Ángyán formulation)[151, 152][2].

[1] As will be mentioned in chapter 7.2.1, the DGrid code uses a different DI convention, where $\delta_{A,B}^{\mathrm{DGrid}} = 0.5 \times \delta_{A,B}$. This is due to DGrid carrying out only one (of the identical) integrals in equation 6.12, hence losing the factor 2 in equation 6.13.

[2] Another variant is the so-called Fulton formulation, where $\Theta_i^{\mathrm{Fulton}} = \sqrt{\Theta_i^{\mathrm{Ángyán}}}$.

6.2 The Domain Averaged Fermi Hole (DAFH)

Similar to the orbital projection for the band structure and density of states (see chapter 5.6), it is often desirable to analyze the respective orbital contribution to the delocalization index. This can be achieved by employing domain averaged Fermi holes (DAFHs), which were initially introduced by Robert Ponec and are rooted in the Fermi hole introduced by Eugene Paul Wigner[153, 154]. The Fermi hole can be expressed by:

$$\rho^{\text{hole}}(\vec{r}_2|\vec{r}_1) = \rho(\vec{r}_2) - 2\frac{\pi(\vec{r}_1,\vec{r}_2)}{\rho(\vec{r}_1)} = \frac{\rho_{\text{XC}}(\vec{r}_1,\vec{r}_2)}{\rho(\vec{r}_1)}, \qquad \frac{\pi(\vec{r}_1,\vec{r}_2)}{\rho(\vec{r}_1)} = \rho^{\text{cond.}}(\vec{r}_2|\vec{r}_1) \tag{6.16}$$

with $\rho^{\text{cond.}}(\vec{r}_2|\vec{r}_1)$ being the conditional probability density of finding an electron at $\vec{r}_2$ given the reference electron is located at $\vec{r}_1$. William L. Luken showed that $\rho^{hole}(\vec{r}_2|\vec{r}_1)$ tends to be localized in specific regions, especially around nuclei, which motivates the integration of the Fermi hole over a specific region (domain-averaging), i.e. a Bader basin Ω in this context[155]. (Additionally, fixing the reference electrons at $\vec{r}_1$ would be unphysical, considering the Heisenberg uncertainty principle.):

$$G^{\Omega}(\vec{r}_2) = N_{\Omega}\int_{\Omega}\rho^{\text{hole}}(\vec{r}_2|\vec{r}_1)d\vec{r}_1 = \int_{\Omega}\rho_{\text{XC}}(\vec{r}_1,\vec{r}_2)d\vec{r}_1 \tag{6.17}$$

where N_{Ω} is the number of electrons in the basin. $G^{\Omega}(\vec{r}_2)$ is the domain averaged Fermi hole, which has a direct correspondence to the DI, which can be seen by comparing equation 6.17 with 6.12. Both equations are almost identical, the only difference being that one positional dependence remains in the definition for the DAFH, while all are integrated out to obtain the DI. Eigenvalues and eigenvectors of the DAFH can be calculated by diagonalizing the matrix representation of the DAFH, i.e. the DOM $S_{i,j}(\Omega)$[156]. The eigenvalues then correspond to occupation numbers, while the eigenvectors are called DAFH orbitals. These orbitals can be further localized by subjecting them to the so-called isopycnic transformation[157]. While conserving the DAFH-density, the DAFH orbitals no longer remain orthogonal after transformation. However, the resulting DAFH orbitals and occupation numbers can be utilized to study the structure of the region they are localized in (i.e. the respective Bader basin Ω). It is often possible to assign the DAFH orbitals to atomic orbitals (s states, p states etc.) by visualizing the states in real space and matching symmetries. By calculating the overlap of a DAFH orbital α with a neighboring basin Ω', the orbital contribution to the DI can be calculated via:

$$\delta^{\alpha}_{A,B} = 2\sigma_{\alpha}\xi^{A\to B}_{\alpha}, \tag{6.18}$$

where σ_{α} is the occupation number of state α and $\xi^{A\to B}_{\alpha}$ the overlap of state α (in basin Ω_A) with basin Ω_B. As the states are real-space representations, the overlaps can be obtained by integration of the state over the respective basins. Besides the orbital contribution to the DI between two basins, other information on the valence states of an atom can be obtained, e.g. lone pairs, or dangling valences.

Lone Pairs and DAFHs

Table 6.1 depicts the DAFH values for trigonal Se ($P3_121$), which forms chains that are connected by covalent bonds, as illustrated in figure 6.4. Each Se atom features six valence electrons, two 4s and four 4p electrons. With each Se atom bonding to two adjacent Se sites in the chain, two electrons remain to form a lone pair[158]. This notion is confirmed by table 6.1: Two p orbitals are bonding, each featuring one p-lobe overlapping about 39% with an adjacent Se site. The remaining third p orbital does not show significant overlap with any of the adjacent Se sites (only about 3.5%) and is much more localized in its own basin (83%), while the bonding p orbitals are only about 52% localized in their own basins. This third p orbital represents the lone pair, matching the expectation of being almost completely filled with about 1.7 e^-. While the overlap of the lone pair to the neighboring Se sites is small, the DI contribution is not completely negligible, due to the orbital being almost filled. This example showcases how DAFH calculations can be employed to detect lone pairs in solids.

Table 6.1: Domain averaged Fermi hole (DAFH) table for trigonal Se ($P3_121$). The delocalization index (DI) denotes the number of shared electron pairs: $\mathrm{ES} = 2 \times \mathrm{DI}$.

Se	Occupation (e^-)	Localization in Native Basin	Overlap to Partner Basin	DI Contribution
4p (2x)	1.03	51.9%	38.9% (1x) 2.6% (1x)	0.802 (Se-Se) 0.055 (Se-Se)
4p (Lone Pair)	1.66	82.8%	3.5% (2x)	0.116 (Se-Se)
4s	1.95	97.4%	0.7% (2x)	0.026 (Se-Se)

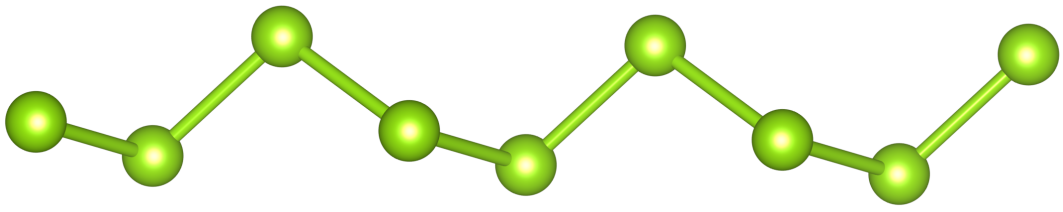


Figure 6.4: Chain of trigonal Se ($P3_121$). Covalent bonds are formed in (approximately) orthogonal directions.

6.3 Delocalization Indices via Maximally Localized Wannier Functions

An alternative way of calculating delocalization indices is employing so-called maximally localized Wannier functions (MLWFs), which allows one to omit the calculation of atomic overlaps that are centered too far apart to contribute, hence reducing the computational workload[149]. All information within this section is taken from [149], if not stated otherwise.

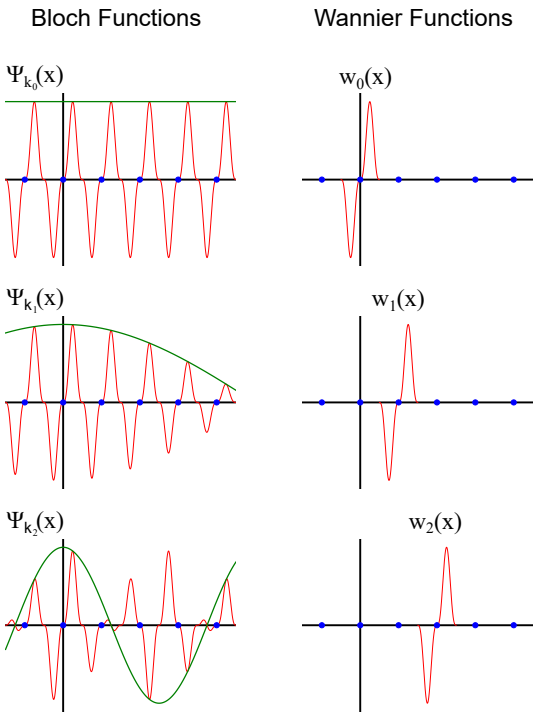


Figure 6.5: Comparison of Bloch (left) and the corresponding Wannier functions (right) in 1D real-space for a single band, i.e. $\Psi_{i,\vec{k}}(\vec{r}) \rightarrow \Psi_k(x)$ and $w_{i,\vec{R}}(\vec{r}) \rightarrow w_R(x)$. The blue circles denote lattice vectors, the green line the e^{ikx} envelope of the respective Bloch function. While Wannier functions contain information identical to the Bloch functions, they are localized in real-space which can offer advantages for specific applications. Redrawn with permission from [159]. Copyright 2023 by the American Physical Society.

Maximally Localized Wannier Functions

So far, Bloch states have been employed to represent the wave functions in a solid (see also equation 5.41):

$$\Psi_{i,\vec{k}}(\vec{r}) = e^{i\vec{k}\vec{r}} u_{i,\vec{k}}(\vec{r}) \tag{6.19}$$

where i in subscript denotes the band index, $\vec{k}$ a vector in reciprocal space within the first Brillouin zone and $u_{i,\vec{k}}(\vec{r})$ a lattice periodic function. Wannier functions can be obtained via the following transformation from the Bloch states[160, 161] (see also figure 6.5):

$$w_{i,\vec{R}}(\vec{r}) = \frac{V}{(2\pi)^2} \int \Psi_{i,\vec{k}}(\vec{r}) e^{i\vec{k}\vec{R}} d\vec{k} \tag{6.20}$$

with $\vec{R}$ being a real-space lattice vector, meaning the Wannier functions are a real-space representation of the wave functions, in contrast to the reciprocal space representation of the Bloch states. The Wannier functions $w_{i,\vec{R}}(\vec{r})$ are also complex and periodic in the supercell. In the reciprocal case, $\vec{k}$-space is sampled into a grid, with n_1, n_2 and n_3 being the number of grid points in each (reciprocal lattice vector) direction. This sampling is equivalent to assuming periodic boundary conditions in real-space for the one-electron states in a $n_1 \times n_2 \times n_3$ supercell, resulting in $N = n_1 \times n_2 \times n_3$ Wannier functions per band, represented by a corresponding lattice vector $\vec{R}_{i,m,n,l}$ $(m, n, l = 0, ..., n_{1,2,3} - 1)$ and the band index i. Wannier functions

$w_{i,\vec{R}}(\vec{r})$ furthermore satisfy the translation relation:

$$w_{i,\vec{R}}(\vec{r}) = w_{i,\vec{0}}(\vec{r} - \vec{R}) \tag{6.21}$$

as well as the normalization relation:

$$\int_{\text{supercell}} w_{i,\vec{R}}(\vec{r}) w_{i',\vec{R}'}(\vec{r}) d\vec{r} = \delta_{i,i'} \delta_{\vec{R}\vec{R}'} \tag{6.22}$$

It should be noted that the Wannier transformation (equation 6.20) is only defined for filled bands, meaning metallic compounds cannot be treated this way. A special flavor of Wannier functions are the so-called maximally localized Wannier functions (MLWFs), which employ the localization criterion of Marzari and Vanderbilt[159, 162, 163], minimizing the sum of the spread of the Wannier functions. This can be achieved by rotating the Bloch states $\Psi_{i,\vec{k}}(\vec{r})$ (see equation 6.19):

$$\widetilde{\Psi}_{i,\vec{k}}(\vec{r}) = \sum_{\alpha} U^{\vec{k}}_{i,\alpha} \Psi_{\alpha,\vec{k}} = e^{i\vec{k}\vec{r}} \sum_{\alpha} U^{\vec{k}}_{i,\alpha} u_{\alpha,\vec{k}} \tag{6.23}$$

where $U^{\vec{k}}_{i,\alpha}$ is an (in principal) arbitrary set of unitary matrices, that are chosen in such a way that the (quadratic) spread Υ of the Wannier functions is minimized:

$$\Upsilon(\Phi) = \sum_{i} \sigma^2(\Phi_i), \qquad \sigma^2(\Phi_i) = \langle \Phi_i | \vec{r}^{\,2} | \Phi_i \rangle - \langle \Phi_i | \vec{r} | \Phi_i \rangle^2 , \tag{6.24}$$

where Φ is a set of electronic states (e.g. Wannier states), and $\sigma^2(\Phi_i)$ the (real-space) variance of a state Φ_i[164]. This minimization (typically done numerically) ensures that each Wannier function $w_{i,\vec{R}}(\vec{r})$ is localized around $\vec{R}$ and approaches zero away from $\vec{R}$. This behavior is convenient for the calculation of the delocalization indices as the overlap between Wannier functions that are separated too far apart can be omitted to reduce computational load.

The delocalization indices can be obtained via the overlap matrices, similar to the equations 6.14 and 6.15, just by means of the Wannier functions:

$$S_{i\vec{R},j\vec{R}'}(\Omega + \vec{R}'') = \int_{\Omega+\vec{R}'} w^*_{i,\vec{R}}(\vec{r}) w_{j,\vec{R}'}(\vec{r}) d\vec{r} \tag{6.25}$$

with i and j being band indices, $\vec{R}$, $\vec{R}'$ and $\vec{R}''$ lattice vectors and $\Omega + \vec{R}''$ an atom (or Bader basin for that matter) translated by the lattice vector $\vec{R}''$.
In itself, computing this property (and consequently the DIs) is no trivial task. Otera-de-la-Roza et al. state that for a relatively small system of 16 atoms sampled by a $4 \times 4 \times 4$ $\vec{k}$-point grid and 24 occupied bands, a total of about 2.4 billion overlaps need to be calculated[149]. To make the computation feasible, a number of sophisticated methods can be employed. These methods will not be discussed in this thesis, but are described in detail in [149].

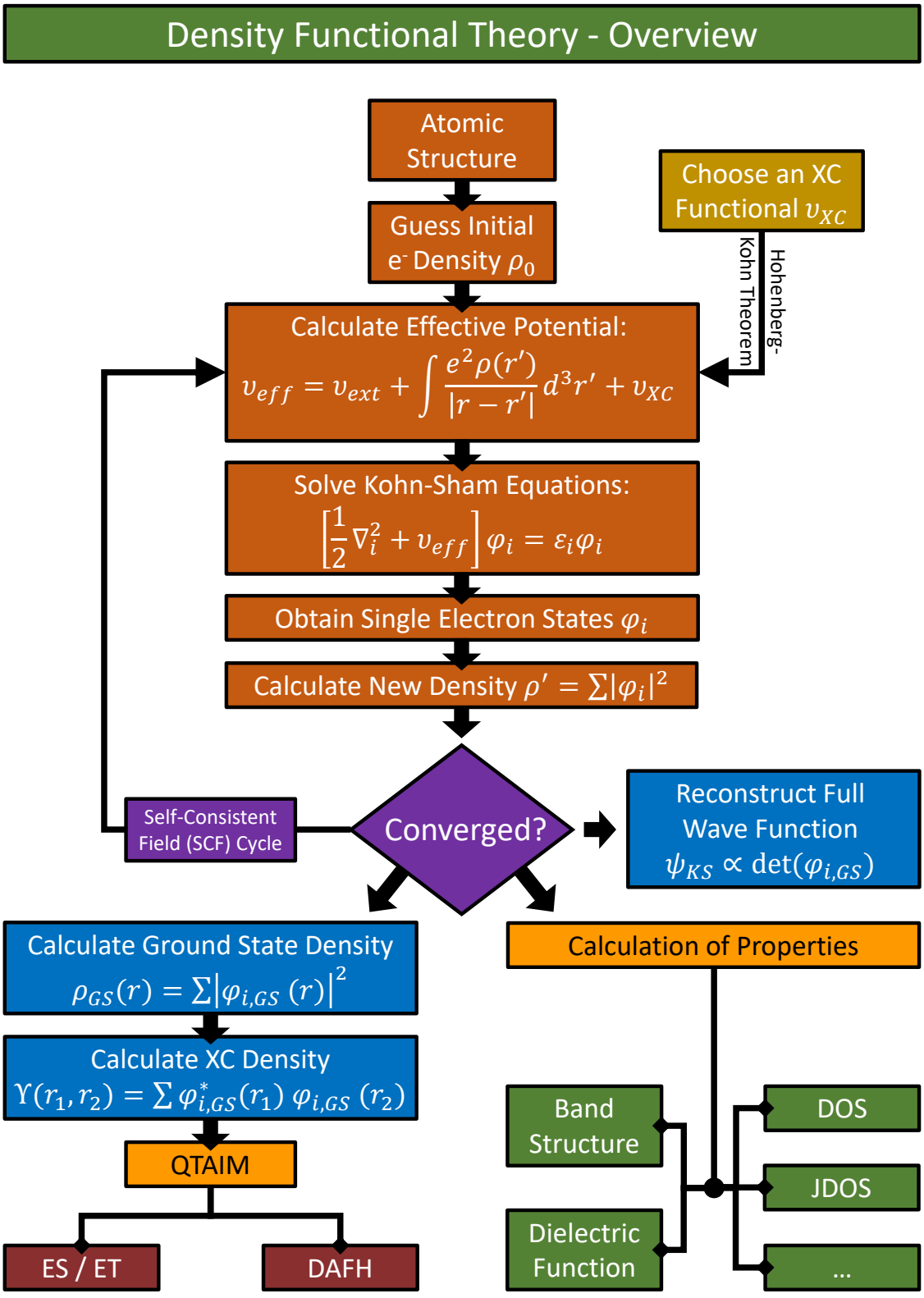


Figure 6.6: Overview over the different computation steps of DFT.

Part III

Chapter 7

Code Implementation

> Five hours of coding can save you five minutes of reading documentation.
>
> *Unknown*

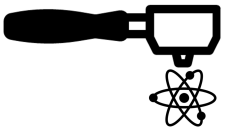
This chapter features a short overview over the implementations of DFT (see chapter 5). The software suites employed in this thesis are discussed in the sections 7.1 to 7.3, along with an example calculation that should enable an inexperienced user to recreate the calculations conducted within this thesis. Section 7.5 will showcase the automation code developed for this thesis and how it can be used to simplify the calculations presented in the preceding chapters. While the code developed in section 7.5 is original to this work, part III assumes the user perspective and thus only deals with technical aspects. It is not required to understand the results and their scientific implications.

7.1 Quantum Espresso and Critic2

The *Quantum opEn Source Package for Research in Electronic Structure, Simulation, and Optimization* (Quantum Espresso - QE) was and is being developed and maintained by Paolo Giannozzi et al., together with the QUANTUM ESPRESSO foundation. The first version was released in 2001 and called *pw.1.0.0*[165–168]. QE utilizes plane waves and pseudopotentials (see chapter 5.5.2) to conduct DFT calculations. Published as open-source software under the GNU General Public License, QE can be freely used and modified. Because of this open-source approach, the QE software suite is composed of multiple independent, but inter-operable codes[165]. This software design streamlines the process of adding new and maintaining existing components. Within the scope of this thesis, Quantum Espresso 6.4 is used.

Critic2 is a software tool published by Alberto Otera-de-la-Roza, Erin R. Johnson and Víctor Luaña in 2014[169], improving on the original release of Critic[170]. Its purpose is to analyze quantum-mechanical interactions in periodic solids, including amongst others the calculation of the localization and delocalization indices (see chapter 6) using the Wannier function approach (see chapter 6.3). While Critic2 utilizes the results of DFT calculations, it itself does not provide the framework to conduct DFT calculations and uses the outputs of other compatible DFT frameworks, one of which is Quantum Espresso. Like QE, Critic2 is publicly available under the GNU General Public License.

7.1.1 The SCF-Cycle in QE



In this section, an exemplary SCF-cycle calculation (see chapter 5.5) using QUANTUM ESPRESSO and the required inputs are presented. The inputs for a QE calculation are stored in an input file with an arbitrary name. For this example, an input file called *paw.scf.in* is used, setting up a calculation of gallium nitride (GaN):

```
&control
    calculation = 'scf'
    restart_mode = 'from_scratch'
    pseudo_dir='/path/to/PPs/PAW'
    outdir='tmp'
    prefix ='GaN'
    tstress= .true.
/
&system
    ibrav=0
    celldm(1)=1.0
    nat=4
    ntyp=2
    occupations='fixed'
    ecutwfc=100
/
&electrons
    mixing_beta=0.2
    conv_thr=1.0d-10
    electron_maxstep=400
/


ATOMIC_SPECIES
Ga 69.723 Ga.PAW.UPF
N 14.007 N.PAW.UPF
ATOMIC_POSITIONS {crystal}
Ga 0.6667 0.3333 0.5000
Ga 0.3333 0.6667 0.0000
N 0.6667 0.3333 0.8759
N 0.3333 0.6667 0.3759
CELL_PARAMETERS
6.0835 0.0 0.0
-3.0418 5.2685 0.0
```

```
0.0 0.0 9.9097
K_POINTS automatic
10 10 7 0 0 0
```

While a detailed description of each (or possible) parameter shown can be found on the QE website[165, 171], a brief overview is provided here:
The first set of inputs is listed under **& control**. The inputs within this bracket manage some general, non-material dependent settings which are required for QE to run properly.

- ◇ **calculation = 'scf'** : Defines the type of calculation to be carried out. In order to compute the SCF-cycle, "scf" is chosen accordingly.
- ◇ **restart_mode = 'from_scratch'**: The calculation starts from scratch. Alternatively, "restart" would allow to continue a previously interrupted run, given it was stopped properly.
- ◇ **pseudo_dir**: The directory where the pseudopotentials (see chapter 5.5.2) are stored that should be used for the calculation.
- ◇ **outdir='tmp'**: Name of the directory in which the calculation results (e.g. electron density) are stored.
- ◇ **prefix ='GaN'**: Prefix for the in- and output files generated by QE. While any string is possible, it is advisable to relate it to the material system to be calculated.
- ◇ **tstress= .true.**: Activates the calculation of the stress tensor. The stress tensor can later be used to assess whether the unit cell is sufficiently relaxed.

The second bracket of inputs is labeled **& system** and defines properties of the unit cell:

- ◇ **ibrav=0**: The Bravais lattice index for the crystal system used. "0" indicates a "free" structure, i.e. the lattice vectors are provided separately.
- ◇ **celldm(1)=1.0**: Lattice parameter for the unit cell. Can be understood as a scaling factor given "ibrav=0". Using "celldm(1)=1.0" implies that all lengths are given in units of Bohr radii (a_0).
- ◇ **nat=4**: Number of atoms in the unit cell (set to "4" for the example GaN unit cell).
- ◇ **ntyp=2**: Number of different elements in the unit cell (in this case "2", as Ga and N are present).
- ◇ **occupations='fixed'**: This parameter can, inter alia, be set to "smearing" to employ gaussian smearing for metallic compounds, or "fixed" for semiconductors with a non-vanishing bandgap.

◇ **ecutwfc=100**: Kinetic cutoff energy for wave functions, given in units of Rydberg (see 5.5.1). Higher values (tend to) improve precision, but increase the computational resources required.

The third bracket of inputs is labeled **& electrons** and contains settings regarding the convergence of the SCF-cycle:

◇ **mixing_beta=0.2**: Influences the mixing ratio of non-converged results (electron densities) between steps of the SCF-cycle (see chapter 5.5).

◇ **conv_thr=1.0d-10**: Threshold for the total energy error in units of Rydberg, below which convergence is assumed to be achieved.

◇ **electron_maxstep=400**: Maximum number of calculation steps allowed within the SCF-cycle before the computation is stopped.

At the end of the input file, the remaining parameters of the unit cell as well as the $\vec{k}$-point sampling are listed. **ATOMIC_SPECIES** defines each element, its weight in atomic units as well as the pseudopotential to be used. In this case "PAW"-potentials are used (see section 5.5.3). The section **ATOMIC_POSITIONS {crystal}** then defines the position of each element in relative coordinates of the unit cell vectors, which are consequently given under **CELL_PARAMETERS**. Finally, the $\vec{k}$-point grid is defined under **K_POINTS automatic**, where the first three numbers define the number of points in each unit cell vector direction, while the last three can be used to offset the $\vec{k}$-point grid[1].
The SCF calculation is the foundation and first step of each property to be calculated with DFT. Starting from the SCF-cycle calculation described in this section, the following sections will illustrate how ES and ET can consequently be calculated using Critic2 and follows the tutorial of Alberto Otera-de-la-Roza[172].

7.1.2 Calculation of ES and ET with QE and Critic2

Step 1

After converging the SCF-cycle, the all-electron density is extracted to the *rhoae.cube* file, using **pp.x**. As the all-electron density is required for the correct determination of the Bader basins (see chapter 6), PAW pseudopotentials have to be used for the SCF calculation. The input file for **pp.x**, *pp.input.plot.rhoae.in*, reads:

[1] To start a calculation, the command `pw.x < paw.scf.in > result.paw.scf.out` can be executed. This tells QE (**pw.x**) to read the inputs from *paw.scf.in* and to redirect all outputs that would have been printed to the console to the file *result.paw.scf.out*

```
&inputpp
    prefix='GaN'
    outdir='tmp'
    filplot='rhoae.dat'
    plot_num=21
/
&plot
 nfile=1
 filepp(1)='rhoae.dat'
 weight(1)=1.0
 iflag=3
 output_format=6
 fileout='rhoae.cube'
/
```

Despite *prefix='GaN'*, all settings in this file are system-independent and merely define the correct format for the consecutive calculations.

Step 2

Next, another SCF calculation with the input file *nc.scf.in* is conducted. It features almost identical parameters as the first one but uses norm-conserving pseudopotentials instead (see chapter 5.5.2). Hence:

```
ATOMIC_SPECIES
Ga 69.723 Ga.PAW.UPF
N 14.007 N.PAW.UPF
```

becomes

```
ATOMIC_SPECIES
Ga 69.723 Ga.NC.UPF
N 14.007 N.NC.UPF
```

From this second SCF run, the electron (pseudo) charge density is extracted using **pp.x** again with the input:

```
&inputpp
    prefix='GaN'
    outdir='tmp'
    filplot='rhops.dat'
```

```
    plot_num=0
/
&plot
 nfile=1
 filepp(1)='rhops.dat'
 weight(1)=1.0
 iflag=3
 output_format=6
 fileout='rhops.cube'
/
```

Step 3

Symmetry operations are used internally within QE to reduce the number of $\vec{k}$-points that have to be computed. To carry out the consecutive calculations, the $\vec{k}$-point grid has to be restored to the original shape. This is done by a so-called non-self-consistent-field (NSCF) cycle calculation. This kind of calculation uses the already converged charge density from the SCF run to reconstruct the KS Hamiltonian. This can e.g. be used to calculate the eigenvalues at different $\vec{k}$-points other than the ones used in the SCF calculation it is based on. In this case, however, its only purpose is to unpack the $\vec{k}$-point grid. For this matter, the $\vec{k}$-points have to be stated explicitly, and prefixed by the total number of points instead of generated automatically as is done for the SCF calculation. The parameters that change for the *nscf.in* input file compared to the *nc.scf.in* file are hence:

```
&control
    calculation = 'nscf'
[...]


K_POINTS crystal
700
0.00000000  0.00000000  0.00000000  1.428571e-03
0.00000000  0.00000000  0.14285714  1.428571e-03
0.00000000  0.00000000  0.28571429  1.428571e-03
[...]
0.90000000  0.90000000  0.85714286  1.428571e-03
```

The explicit $\vec{k}$-point grid can (and should) be generated using additional tools like **kmesh.pl**, which is included in the wannier90 package of QE[2]. The fourth entry is the weight of each $\vec{k}$-point, i.e. $(n_\mathrm{x} \times n_\mathrm{y} \times n_\mathrm{z})^{-1}$, where n_i are the numbers of $\vec{k}$-points in each dimension.

[2] In this case by running: `kmesh.pl 10 10 7`

Step 4

Critic2 employs maximally localized Wannier functions (MLWFs, see chapter 6.3) to calculate the delocalization index. In order to compute the MLWFs, the KS-state coefficients need to be extracted from the QE calculations. This is done by using the **pw2critic.x** tool, which comes with the Critic2 software package. The corresponding input file *pw2critic.in* contains:

```
&inputpp
 outdir='tmp/'
 prefix='GaN'
 seedname='wannier'
/
```

running `pw2critic.x < pw2critic.in` yields a *.pwc* file that contains the desired KS-coefficients.

Step 5

Besides the KS-state coefficients, the rotations of the KS-states are required as well to compute the MLWFs. This is done by the **wannier90.x** program that is included in the QE software package. To prepare the MLWF calculation, the information required for the run is prepared by creating a *wannier.win* input file:

```
num_wann=18
num_iter=10000
conv_window=3
conv_tol=1d-10

begin projections
random
end projections

mp_grid : 10 10 7
begin unit_cell_cart
bohr
6.0835 0.0 0.0
-3.0418 5.2685 0.0
0.0 0.0 9.9097
end unit_cell_cart
begin atoms_frac
Ga 0.6667 0.3333 0.5000
```

```
Ga 0.3333 0.6667 0.0000
N 0.6667 0.3333 0.8759
N 0.3333 0.6667 0.3759
end atoms_frac
begin kpoints
0.00000000  0.00000000  0.00000000  1.428571e-03
0.00000000  0.00000000  0.14285714  1.428571e-03
0.00000000  0.00000000  0.28571429  1.428571e-03
[...]
0.90000000  0.90000000  0.85714286  1.428571e-03
end kpoints
```

Most of the parameters are identical to the ones of previous input files. The parameter **num_wann=18** corresponds to the number of bands in the system, which can be obtained from the output of the norm-conserving SCF calculation result. The parameters **conv_window=3** and **conv_tol=1d-10** simply enforce a strict convergence criterion regarding the minimization of the Wannier functions. This input file is first executed via `wannier90.x -pp wannier.win`, where the `-pp` flag tells the program to only generate a list of required overlaps and then to terminate. Before the MLWF can finally be obtained, the required integral files must be obtained via `pw2wannier90.x < pw2wan.in`:

```
&inputpp
 outdir='tmp/'
 prefix='GaN'
 seedname='wannier'
 write_mmn=.true.
 write_amn=.true.
/
```

Ultimately, running `wannier90.x wannier.win` carries out the MLWF computation and generates a *.chk* file containing the orbital rotation matrices.

Step 6

With both the *.pwc* file containing the KS coefficients and the *.chk* file containing the rotations, Critic2 can be started using the input *critic2.cri* via `critic2 < critic2.cri > critic2.cro`:

```
crystal wannier.pwc
load rhoae.cube id rho
load wannier.pwc wannier.chk
```

```
reference rho
integrable 2
integrable 2 deloc
yt
```

The output file *critic2.cro* then contains the delocalization indices (DIs) and electron populations of each basin, as exemplified below for the Ga site of GaN:

```
[...]
# Attractor 3 (cp=3, ncp=2, name=Ga, Z=31) at: 0.6667000  0.3333000  0.4991000
# Id   cp   ncp   Name  Z    Latt. vec.  --- Cryst. coordinates ---    Distance        LI/DI
  Localization index..................................................... 10.07655504
  282  2    1     N     7    1   0   0   1.3333    0.6667    0.3759    3.7181087    0.61024119
  254  2    1     N     7    0  -1   0   0.3333   -0.3333    0.3759    3.7181386    0.61061230
  2    2    1     N     7    0   0   0   0.3333    0.6667    0.3759    3.7191312    0.60923291
  1    1    1     N     7    0   0   0   0.6667    0.3333    0.8759    3.7339750    0.59338402
  308  4    2     Ga    31   1   0  -1   1.3333    0.6667   -0.0009    6.0732475    0.01599691
  284  4    2     Ga    31   1   0   0   1.3333    0.6667    0.9991    6.0732475    0.01599690
  280  4    2     Ga    31   0  -1  -1   0.3333   -0.3333   -0.0009    6.0732658    0.01600964
  [...]
  2060 4    2     Ga    31   7   3  -4   7.3333    3.6667   -3.0009   49.3617060    0.00000000
  Total (atomic population)..................................................... 11.45910247
  [...]
```

The DIs are stated at the end of each line, multiplying by the factor of two then yields the value of electrons shared ES $= 2 \times$ DI between the attractor (in this case Ga), and the atomic site corresponding to the respective row. The total number of electrons transferred is obtained by subtracting the formal charge of the respective element from the entry in the last row. This yields the following values for the GaN example presented in this chapter: TET $\approx 11.46 - 13 = -1.54$[3] and ES ≈ 1.22 (between nearest neighbors).

7.1.3 Band Structure and Density of States Calculation using QE

In this chapter, the calculation of the band structure (BS) and density of states (DOS) using QE is exemplified on GaN again (see also chapter 5.6). As a first step, a SCF calculation has to be conducted to obtain the ground-state density, as already shown in section 7.1.1. Next, a so-called "bands" calculation has to be conducted, which is a non-self-consistent calculation, meaning the electronic ground state results from the SCF calculation are required. Within the input file *bands.in* the $\vec{k}$-points for which the band structure shall be calculated are stated, instead of the uniform $\vec{k}$-point grid. It is then executed with the command `bands.x < bands.in > bands.out`:

[3] As only valence electrons are considered, the nominal charge of Ga consists of ten 3d, two 4s and one 4p electrons. How many valence electrons are used can be checked within the PP file.

```
&control
 calculation = 'bands'
 prefix='GaN'
 tstress = .true.
 pseudo_dir='/path/to/PPs/NC'
 outdir='tmp'
/
&system
 ibrav =0,
 nat=4
 ntyp=2
 ecutwfc = 100
 occupations = 'fixed'
 nbnd = 34
/
&electrons
 mixing_beta = 0.2
 conv_thr = 1.0d-10
/
ATOMIC_SPECIES
N 14.007 N.upf
Ga 69.723 Ga.upf
ATOMIC_POSITIONS {crystal}
Ga  0.666667  0.333333  0.499136
Ga  0.333333  0.666667  0.999136
N  0.666667  0.333333  0.875864
N  0.333333  0.666667  0.375864
CELL_PARAMETERS
6.08354  0.0  0.0
-3.04177  5.268499  0.0
0.0  0.0  9.909726
K_POINTS {crystal_b}
12
0.0  0.0  0.0  20.0
0.5  -0.0  0.0  20.0
0.33333  0.33333  0.0  20.0
0.0  0.0  0.0  20.0
0.0  0.0  0.5  20.0
0.5  -0.0  0.5  20.0
0.33333  0.33333  0.5  20.0
0.0  0.0  0.5  20.0
0.5  -0.0  0.5  20.0
```

```
0.5   -0.0   0.0   20.0
0.33333   0.33333   0.0   20.0
0.33333   0.33333   0.5   20.0
```

The coordinates of the $\vec{k}$-points are listed under **K_POINTS {crystal_b}**, prefixed by the total number of points. The number after the three coordinates denotes the number of interpolation points between two given $\vec{k}$-points. Another important parameter is the number of bands, **nbnd = 34**. This variable is system-dependent and must be adapted to suit the system. Generally, the number of bands has to be at least half the number of electrons in the unit cell, as two electrons can occupy the same state. Generally, this number should be higher than that (especially for metallic compounds), as empty states are relevant for convergence. Furthermore, the conduction band would (generally) not be visible choosing only the minimum number of bands. The pseudopotential for Ga used here assumes 13 valence electrons (i.e. ten 3d, two 4s and one 4p electrons), while the one for N assumes 5 (i.e. two 2s and three 2p electrons). With two Ga and Na atoms each per unit cell, choosing **nbnd = 34** should hence be sufficient[4]. Plotting the data of the output files yields the band structure shown in figure 7.1: In order to calculate the density of states, it can be advantageous to use a denser

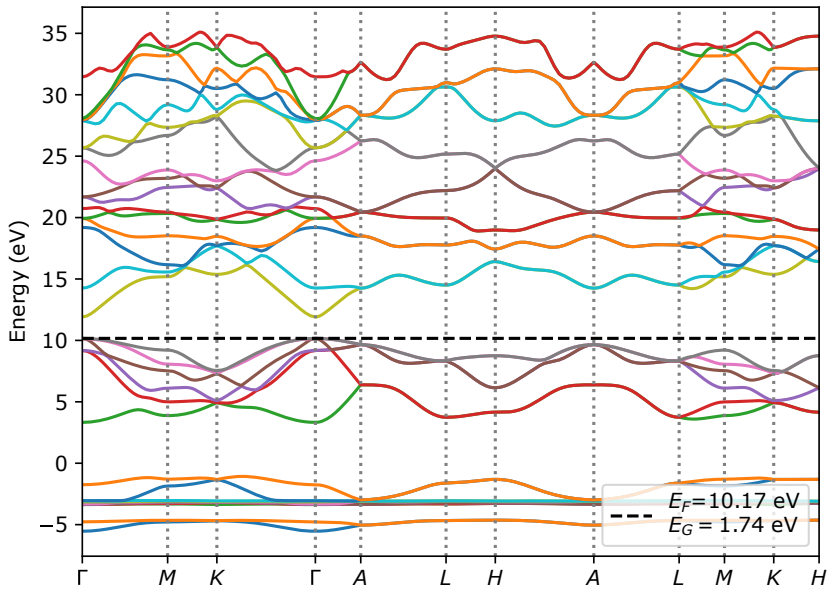


Figure 7.1: Band structure of GaN, as obtained by an exemplary Quantum Espresso calculation. The colors indicate different bands. The direct band gap of $E_\mathrm{G} = 1.74\,\mathrm{eV}$ underestimates the experimental value of approx. $E_\mathrm{G}^\mathrm{Ex.} = 3.4\,\mathrm{eV}$[173].

[4]The number of considered valence electrons is listed within the pseudopotential file. While it often corresponds to the number of electrons in the outermost shell, electrons from the second outermost shell are sometimes included as well, depending on the element and the settings chosen to create the pseudopotential.

$\vec{k}$-point grid, resulting in a smoother DOS. For that matter, another NSCF calculation can be used with the exact same inputs as the SCF calculation, only replacing **calculation = 'scf'** by **calculation = 'nscf'** and increasing the $\vec{k}$-point density. As an NSCF calculation does not update the ground-state density, choosing a denser $\vec{k}$ mesh in this step has computationally less impact than in the SCF calculation. Nevertheless, the $\vec{k}$-point density in the SCF must still be sufficient to ensure proper convergence. On the other hand, if the $\vec{k}$-point density was already reasonably high in the SCF calculation, this NSCF calculation can be omitted.
To compute the DOS, another input file, *DOS.in*, needs to be prepared and executed via `dos.x < DOS.in > DOS.out`. It merely contains in- and output file names, however:

```
&dos
   outdir='tmp'
   prefix='GaN'
   fildos='GaN.dos'
/
```

Plotting the output results in a file as shown in figure 7.2.

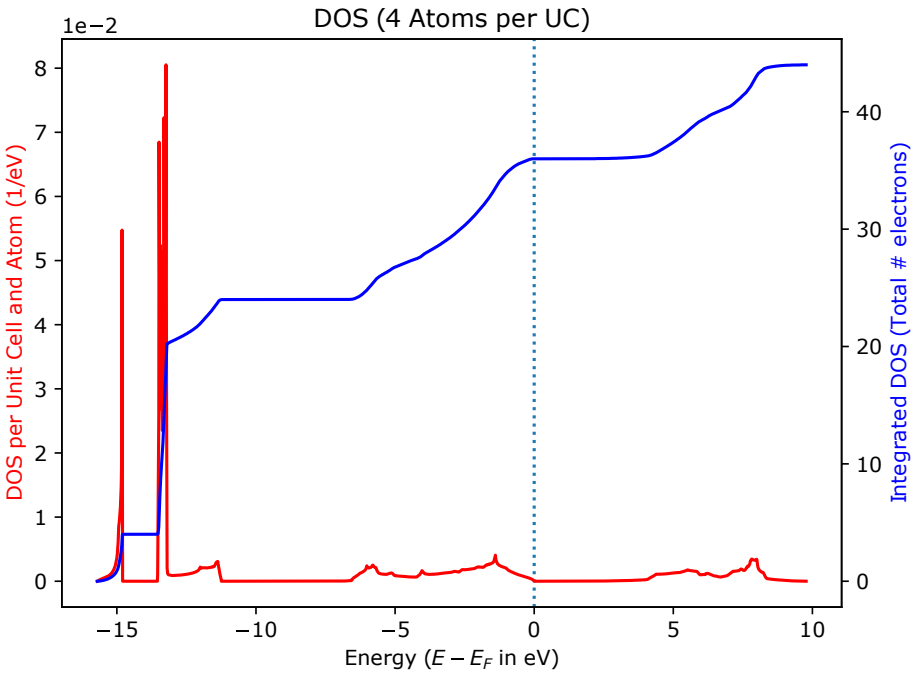


Figure 7.2: Density of states of GaN, as obtained by an exemplary Quantum Espresso calculation. The red line indicates the DOS, while the blue one denotes the integrated density of states (IDOS). The energy scale has been shifted so that $E_\mathrm{F} = 0\,\mathrm{eV}$.

Orbital-Resolved Band Structure and DOS

It can often be insightful to project the atomic orbitals onto each band of the band structure and to decompose the DOS to the orbital contributions (also see chapter 5.6). The process starts similar to the regular band structure and DOS calculation. After the bands calculation, however, a different input has to be prepared *kpdos.in*:

```
&projwfc
   outdir='tmp'
   prefix='GaN'
   ngauss=0
   degauss=0.015
   DeltaE=0.01
   kresolveddos=.true.
   filpdos='GaN.k'
/
```

The parameters **ngauss=0** and **degauss=0.015** introduce some broadening (gaussian broading with a value of 0.015 Ry, respectively). The variable **kresolveddos=.true.** indicates that the projection is to be calculated for each $\vec{k}$-point individually, and not to be summed. Executing `projwfc.x < kpdos.in > kpdos.out` and plotting the resulting files yields figure 7.3: This variant of visualization is called "FAT Bands", as the size of the markers

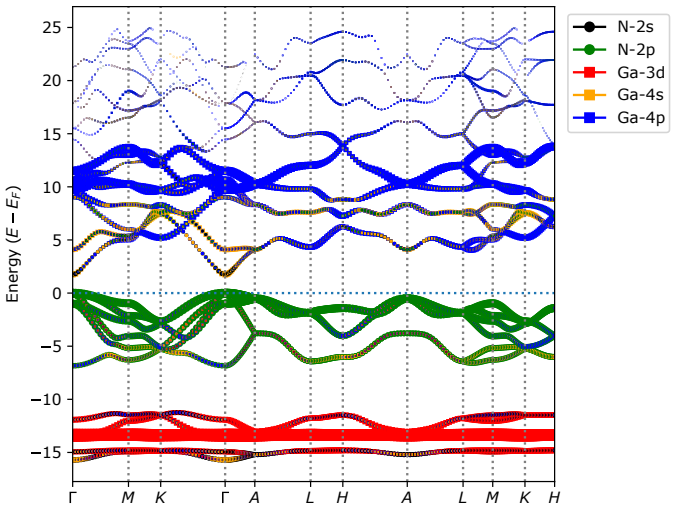


Figure 7.3: FAT bands of GaN. The color and shape of each marker represent a respective orbital, while the size corresponds to the relative contribution to a specific band at a given $\vec{k}$-point. This type of visualization can be helpful in assessing the general orbital character of a band, but lacks in precision. Especially overlapping symbols can be misleading.

indicates the relative contribution of an orbital to a band at a given $\vec{k}$-point, resulting in

"fat" bands. It should further be noted that the output of the *kpdos.in* calculation is not yet formatted in a way that facilitates this kind of plot. Furthermore, the output is more detailed than shown in this graph. The orbitals are resolved by the quantum number m (i.e. contributions are calculated p_x, p_y, p_z, d_{xy}, d_{xz} etc.) and the contributions are resolved by atomic site, not atomic type (i.e. values for Ga_1, Ga_2, N_1 and N_2 exist). Reasonable simplifications have to be employed to keep the band structure legible. Figure 7.3 exemplifies another effect, namely the apparent "fading" of the bands for higher energies. This occurs as the pseudopotentials used do not incorporate the higher orbitals that would be occupied at these energies (e.g. Ga 4d), hence only the remaining orbital contributions of the energetically lower orbitals remain.
The orbital-resolved DOS can be generated with almost the same input *DOS_kpdos.in*:

```
&projwfc
   outdir='tmp'
   prefix='GaN'
   ngauss=0
   degauss=0.015
   DeltaE=0.01
   kresolveddos=.false.
   filpdos='GaN.k'
/
```

The difference being **kresolveddos=.false.** instead of **kresolveddos=.true.**. Plotting the respective output produces figure 7.4:

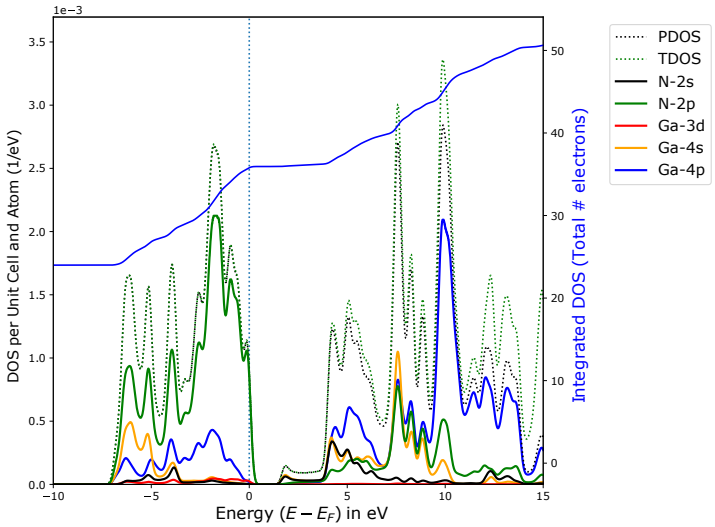


Figure 7.4: Orbital-Resolved DOS of GaN. The colors indicate different orbital contributions. PDOS refers to the sum of all projected DOS contributions, while TDOS is the total DOS. At around 5 eV PDOS and TDOS start to deviate as orbitals not included in the pseudopotentials start to become relevant.

7.1.4 Structural Relaxation using QE

So far, the unit cell and the position of the atomic cites has been considered and consequently treated to be static (see section 5.2). However, unit cells are not necessarily in their equilibrium state. Especially unit cells obtained by experiment might originate from materials under stress and/or strain, in addition to thermal expansion yielding different unit cell sizes at room temperature, while (standard) DFT is limited to 0 K. It is hence possible to use DFT to relax a unit cell to its equilibrium configuration by adapting lattice vector lengths, angles, and atomic positions. This is done by converging the SCF-cycle for the initial unit cell and computing the stress tensor and forces on atoms and adapting the unit cell parameters accordingly. Then another SCF-cycle is computed, using the updated unit cell. This relaxation cycle is then repeated until the forces and the stress vanish (and consequently the total energy of the system reaches a minimum). Fortunately, this scheme is readily implemented in QE. The following input file is an example for HfC in the $Fm\bar{3}m$ structure, where the first component of the first lattice vector has been increased ($6.203\,a_0 \rightarrow 6.403\,a_0$):

```
&control
    calculation = 'vc-relax'
    restart_mode = 'from_scratch'
    pseudo_dir='/path/to/PPs/NC'
    outdir='tmp_relax'
    prefix ='relax'
    tstress= .true.
    forc_conv_thr=1.0D-5
    etot_conv_thr=1.0D-10
/
&system
    ibrav=  0
    celldm(1)=1.0
    nat=2
    ntyp=2
    occupations='fixed'
    ecutwfc=80

/
&electrons
    mixing_beta=0.2
    conv_thr=1.0d-10
    electron_maxstep=400
/

&ions
```

```
    ion_dynamics = 'bfgs'

/
&cell
  cell_dynamics='bfgs'
  press=0.0
  press_conv_thr=0.1
/


ATOMIC_SPECIES
C 12.011 C.upf
Hf 178.49 Hf.upf
ATOMIC_POSITIONS {crystal}
C 0.55 0.5 0.5
Hf 0.0 0.0 0.0
CELL_PARAMETERS
6.402614102 0.0 0.0
3.101307051 5.371621383 0.0
3.101307051 1.790540461 5.10076543
K_POINTS automatic
6 6 6 0 0 0
```

Most of the parameters are identical to the SCF-cycle and have already been mentioned in section 7.1.1. The new (or changed) parameters are:

- ◇ **calculation = 'vc-relax'**: A relaxation is carried out with a variable cell (vc), meaning the lattice vectors are allowed to be adapted.
- ◇ **forc_conv_thr=1.0D-5**: Minimum value (a.u.), below which the ionic forces are considered to be sufficiently small to be considered converged.
- ◇ **ion_dynamics = 'bfgs'**: Name of optimization algorithm used to adapt the atomic positions.
- ◇ **press=0.0**: Target pressure for the cell to reach. "0" is chosen here, as the cell is supposed to relax to its equilibrium.
- ◇ **press_conv_thr=0.1**: Threshold value for the pressure below which convergence is assumed be achieved.

Within the output file, the stress tensor after the first SCF-cycle (before updating any cell parameters) reads:

```
(kbar)     P=  -69.81
       -127.52     25.56     17.88
         25.56    -27.96      4.09
         17.88      4.09    -53.93
```

This stress tensor indicates that the cell is indeed not at equilibrium. After the relaxation is complete, the lattice vectors have been optimized to:

```
     crystal axes: (cart. coord. in units of alat)
               a(1) = (   6.232640   0.039543   0.027886 )
               a(2) = (   3.052269   5.399599   0.019716 )
               a(3) = (   3.052337   1.818527   5.084263 )
```

The first component of the first lattice vector has been restored sufficiently close to the original value, while the remaining components are almost identical. And indeed, the final stress tensor reads:

```
(kbar)     P=   -0.23
         -0.37      0.05      0.04
          0.05     -0.12      0.02
          0.04      0.02     -0.21
```

which can be considered sufficiently small pressures.
As there are many free parameters to optimize, it can be advisable to e.g. optimize the atomic positions first, while locking the lattice vectors, followed by optimizing the cell parameters and fixing the atomic positions and only leaving all parameters open in the final computation. This process is usually more robust, even for unit cells that are far off their equilibrium state.

7.2 Abinit and DGrid

Abinit is another DFT implementation, similar to Quantum Espresso, but also features density-functional perturbation theory, many-body perturbation theory and other first-principle approaches. It is a project initiated by Xavier Gonze (Universite Catholique de Louvain) and Douglas C. Allen (Corning Inc.) and further developed by many other institutions and individuals, as it has been published under the GNU General Public License since 1999[174, 175]. Within the scope of this thesis, version 7.10.5 was used.
DGrid was developed by Miroslav Kohout. It is a tool that operates in real-space to calculate QTAIM properties such as Bader basins, localization and delocalization indices, domain averaged Fermi holes, electron localization function (ELF) and many others (see chapter 6). Within this thesis, version 4.7 was used[176].

7.2.1 Calculation of ES and ET with Abinit and DGrid

In this chapter, the calculation of ES and ET will be outlined using the example of aluminum (Al).

Step 1 - SCF Calculation

To begin the calculation, the ground electronic state density has to be determined using an SCF-cycle. This process is conceptually identical to the calculation done for GaN with Quantum Espresso, but the input files have to adapted to suit Abinit. Abinit requires two input files to conduct an SCF-cycle. The first is called *Al.files*, as it merely contains in- and output file names, as well as the names of the pseudopotentials:

```
Al.in
Al.out
Al_i
Al_o_DS1
Al_m
/path/to/PPs/Pseudos_GGA/Al.paw.abinit
```

Al.in and **Al.out** are the names of the formatted in- and output files, while **Al_i**, **Al_o_DS1** and **Al_m** are (arbitrary) names for the generic input, output and temporary files. After these five file names, the name (and path to) the required pseudopotentials are appended, in this case only the PP for Al is required. The second required file, as already requested in *Al.files*, is the *Al.in* file, which contains the parameters for the DFT calculation:

```
istwfk 10000*1

# Electronic Ground state calculation

kptopt 3
tolvrs 1.0d-10
iscf 17

#Definition of the unit cell

acell 3*1.0
rprim
4.709479 0.032459 2.634708
1.600415 4.430055 2.634731
-0.03029 -0.021389 5.321882

#Definition of the atom types
```

```
ntypat 1
znucl 13
#Definition of the atoms
natom 1
typat 1*1
xred

  0.0  0.0  0.0

#electron section
nband 6
occopt 3
tsmear 0.005

#Exchange-correlation functional
ixc 11

#Definition of the planewave basis set
ecut 25.0
pawecutdg 38.0

#Definition of the k-point grid
ngkpt 10 10 10

shiftk 0 0 0

#Definition of the SCF procedure
nstep 50
diemac 1000000
```

A short description of the input parameters is given below, while a more detailed description can be found in the ABINIT documentation, e.g. [177].

- ◇ **istwfk 10000*1**: Defines the save format of the wave function to ensure compatibility with DGRID.
- ◇ **kptopt 3**: Disables the reduction of the $\vec{k}$-point grid using symmetry operations. Required to ensure compatability with DGRID.
- ◇ **tolvrs 1.0d-10**: Tolerance on the potential residual. This parameter defines the convergence criterion for the SCF-cycle.
- ◇ **iscf 17**: Choice of SCF-cycle algorithm. 17 corresponds to so-called Pulay mixing.
- ◇ **acell 3*1.0**: Scaling factor of each of the lattice vectors. As the factor is chosen to be 1.0 for all vectors, they must be stated in units of Bohr radii (a_0).
- ◇ **rprim**: Stated below this keyword are the lattice vectors.
- ◇ **ntypat 1**: Number of different element types present in the unit cell.

- ◇ **znucl 13**: Atomic number of the elements present in the unit cell, separated by spaces (hence for GaN the correct input would be: **znucl 31 7**. The order of the elements must correspond to the order of PPs given in the *Al.files* file.
- ◇ **natom 1**: Total number of atoms in the unit cell.
- ◇ **typat 1*1**: Stating the type of atoms for the consequent listing of atomic positions. In this case this means: One atom of type 1. (In the case of GaN, it would read **typat 2*1 2*2**, as there are two instances of Ga, and two instances of N present.)
- ◇ **xred**: Stated below this keyword are the (relative) atomic position of the atoms in the unit cell. The order must match the order stated in **typat**.
- ◇ **nband 6**: Number of bands to be used for the calculation. As aluminum features only 3 valence electrons and one atom per unit cell, 6 is sufficient.
- ◇ **occopt 3**: Defines the type of smearing employed, "3" corresponds to Fermi-Dirac type smearing.
- ◇ **tsmear 0.005**: Broadening of smearing given in units of Hartree.
- ◇ **ixc 11**: Choice of exchange-correlation functional. "11" corresponds to GGA-PBE (see chapter 5.5.4).
- ◇ **ecut 25.0**: Energy cutoff given in Hartree (see chapter 5.5.1).
- ◇ **pawecutdg 38.0**: Energy cutoff specific to calculations using PAW PPs (see chapter 5.5.3). Related to the transition between the inside and outside region of the augmentation spheres (see [177] for more information). **pawecutdg** must be larger than **ecut**. A factor of 1.5 is usually a reasonable estimate.
- ◇ **ngkpt 10 10 10**: Definition of the $\vec{k}$-point grid.
- ◇ **shiftk 0 0 0**: Shift of the $\vec{k}$-point grid.
- ◇ **diemac 1000000**: Initial guess of the dielectric constant for the system, helping the SCF algorithm to converge faster. For metals, high values are chosen, for insulators "10" is a reasonable guess.

With both files set up, ABINIT can be started via: `abinit < Al.files > log`, generating inter alia the output file *Al_o_DS1_WFK* containing the wave function data required for DGRID.

Step 2 - Basis File

DGRID is in general compatible with various DFT implementations. Hence the output of the ABINIT calculations as well as the charge information (within and outside the augmentation spheres of the PAW PPs) have to be combined and reformatted. This is conveniently done by simply running `dgrid-4.7 Al.files Al.abi` on the same input file used for the ABINIT calculation, creating the files *Al.abi* and *Al.abi.coeffs*.

Step 3 - Charge Density

Next, the charge density grid is calculated in real space, using the input file *dg.inp*:

```
:TITLE
:-------------------------------------------------------------------|
:: Al
:-------------------------------------------------------------------|

:KEYWORDS
:--------------------------------------------------------------------
 basis=Al.abi

:CHOOSE THE DESIRED PROPERTIES
:-----------------------------------------

compute=rho

:-----------------------------------------

vectors
origin: 0. 0. 0.
i-vector:    4.7095    0.0325    2.6347  54
j-vector:    1.6004    4.4301    2.6347  54
k-vector:    -0.0303    -0.0214    5.3219  54
END
```

In this file, the basis file must be stated as **basis=Al.abi**, as well as the lattice vectors in Bohr radii after the keyword **vectors**. The fourth entry of each vector denotes the grid spacing. Generally a resolution of 0.1 a_0 is sufficient. The **origin** keyword can be used to shift the grid, which can be useful to center specific atomic sites of interest. Executing `dgrid-4.7 dg.inp` then creates a *Al.abi.rho_r* file, containing the charge density grid.

Step 4 - Translation/Super Cell

After generating the charge density grid, the grid has to be translated to create a supercell. This step is technically optional, but results for atomic sites at the border of the grid are not guaranteed to be correct. Hence especially for small unit cells such a translation is strongly recommended. To conduct the translation, execute `dgrid-4.7 Al.abi.rho_r op1`. The `op1` flag signals DGrid to expect further (manual) inputs. As Al only features a single atom, a double translation in each direction is advisable. For this, the following inputs are required:

1. translation in
2. translation in
3. translation jn
4. translation jn
5. translation kn
6. translation kn
7. save

For each call of **translation**, DGrid doubles the dimension of the unit cell, hence a $4 \times 4 \times 4$ cell has been created. For most structures, a single call of **translation** per direction is more than sufficient, as the computational cost and memory requirements increase exponentially. A shifted charge density file *Al.abi.rho_r_.trans_in.trans_in.trans_jn.trans_jn.trans_kn.trans_kn* is created, the suffix indicating the translations that have been carried out.

Step 5 - Grid Refinement

The accuracy of the integration performed by DGrid to evaluate delocalization indices, etc., strongly depends on the density of the charge density grid as defined in step 3. Increasing the grid density would benefit the accuracy, however, while also strongly increasing the computational load and memory requirements. DGrid features the option to refine the grid in areas of highly non-linear behavior only, while using the original (less dense) grid everywhere else. The refinement can be conducted via `dgrid-4.7 Al.abi.rho_r_.trans_in.[...].trans_kn refine 1`, where the "1" denotes the precision of the refinement. Lower values (e.g. "0.5") can be chosen to increase precision; "1" however is usually sufficient. A new file is generated called *Al.abi.rho_r_.trans_in.trans_in.trans_jn.trans_jn.trans_kn.trans_kn.rfn*.

Step 6 - Generation of Bader Basins

The refined grid can consequently be used to calculate the Bader basins (see chapter 6). This requires another input file, called *comp_basins*:

```
::total_density_3D
property=Al.abi.rho_r.trans_in.trans_in.trans_jn.trans_jn.trans_kn.trans_kn
integrate=Al.abi.rho_r.trans_in.trans_in.trans_jn.trans_jn.trans_kn.trans_kn.rfn
symmetry=translation i j k
top=0.5
output = .
end
```

The **property** keyword denotes which file contains the grid values for the determination of the Bader basins. **integrate** denotes the file containing the property to be integrated over, i.e. the refined charge density is integrated over each Bader basin. The keyword **symmetry=translation i j k** enables DGRID to utilize translation symmetries at the edge of the grid (which might not be desired for molecular calculations, for example). Finally, the **top=0.5** parameter is a threshold value for the basin calculation. If the number of basins does not match the number of atomic sites, lowering it merges smaller basins together (in the event "empty" basins appear in the calculation, i.e. basins with no atom inside), while increasing it generally creates more basins (e.g. useful if the border between two basins could not be resolved and a combined basin with two or more atomic sites inside is obtained).
Executing `dgrid-4.7 comp_basins` then creates two new files *Al.[...].bas* and *Al.[...].bsn*, containing the information regarding the calculated basins[5]

Step 7 - Computation of DI, LI and Domain Populations

Finally, the QTAIM properties can be calculated, using an input file similar to the one used for the basin calculation, *overlap.inp*:

```
:TITLE
:-----------------------------------------------------------------|
::overlap
:-----------------------------------------------------------------|

:KEYWORDS
```

[5] A small handguide on how to use PARAVIEW to plot the basins: Use PARAVIEW 5.2.0: 1: Load the plugins "elf_mk" and "SMMDGridPlugin" (inlcuded in *Project DIO*, see chapter 7.5) via *Tools → Manage Plugins...* in the toolbar. 2: Drag the *Al.[...].bsn* file into PARAVIEW. 3: Click on *Structure* in the *Pipeline Browser* and apply the *Plot Structure* filter via *Filters* in the toolbar to show atoms and the unit cell. 4: Select the *Property* object of the *Al.[...].bsn* file in the *Pipeline Browser* and apply the *Plot Basin* filter.

```
:-----------------------------------------------------------------------
property=Al.abi.rho_r.trans_in.trans_in.trans_jn.trans_jn.trans_kn.trans_kn.bsn
integrate=Al.abi.rho_r.trans_in.trans_in.trans_jn.trans_jn.trans_kn.trans_kn.rfn

overlap
output= .

end
```

Again, this input file is executed via `dgrid-4.7 overlap.inp`, yielding an *Al.[...].sij*, containing the overlap matrices for each basin combination (see chapter 6.1) and an *Al.[...].ovl* file, containing human-readable outputs, i.e. inter alia the delocalization indices and domain populations. For the localization indices and the domain population, the following table is produced:

```
+----------------------+
| Localization indices |
+----------------------+
```

Basin	Descriptor	Q	Daa	Dbb	Dab	sigma2	LI	LIaa	LIbb
1	Al_1	3.000	0.883	0.883	0.000	2.031	0.969	0.485	0.485
2	Al_1	3.000	0.883	0.883	0.000	2.031	0.969	0.485	0.485
3	Al_1	3.000	0.883	0.883	0.000	2.031	0.969	0.485	0.485
4	Al_1	3.000	0.883	0.883	0.000	2.031	0.969	0.485	0.485
5	Al_1	3.000	0.883	0.883	0.000	2.031	0.969	0.485	0.485

```
[...]
```

"Q" denotes the domain population, which is equal to the number of valence electrons for Al, resulting in a TET = 0, as expected from a monoatomic compound. "LI" corresponds to the localization index, and "sigma2" is half the bond order, which is the total sum of all DI to all other basins:

$$\text{Bond Order}(\Omega_i) = \sum_{j \neq i} \text{DI}(\Omega_i, \Omega_j) = 2\sigma_i^2 \tag{7.1}$$

The delocalization indices are given in a grid-like table, stating the basin numbers:

```
+------------------------+        +-------+
| Delocalization indices |        | Total |
+------------------------+        +-------+
```

Basin	1	2	3	4	5	6	7	8	9	10
1	x	0.0028	0.0030	0.1356	0.0007	0.0002	0.0002	0.0046	0.0016	0.0007
2	0.0028	x	0.0025	0.0046	0.1356	0.0030	0.0001	0.1356	0.0007	0.0002
3	0.0030	0.0025	x	0.1356	0.0046	0.0028	0.0002	0.0006	0.0002	0.0001

```
4   0.1356   0.0046   0.1356      x       0.0028   0.0007   0.0002   0.0025   0.0005   0.0002
5   0.0007   0.1356   0.0046   0.0028       x       0.1356   0.0001   0.0030   0.0002   0.0001
[...]
```

As basin 1 and 5 are nearest neighbors, the calculated $DI_{DGrid} = 0.1356$. It is important to consider that DGRID uses a different convention regarding DI (also see equations 6.12 and 6.13 in chapter 6):

$$ES = 2 \times DI_{CRITIC2} = 4 \times DI_{DGrid} = 0.5424 \tag{7.2}$$

In general, all formulas or DI values stated in this thesis correspond to the CRITIC2 definition of DI, hence DI = ES/2, if not declared otherwise. Raw output files of DGRID calculations, however, will use the DGRID definition.
DGRID and CRITIC2 also differ in compute time. For small unit cells (two atoms), DGRID finishes the computation about two times faster than CRITIC2 (identical $\vec{k}$-point grid and CPU/RAM ressources, while remaining default settings differ). However, CRITIC2 becomes more efficient for larger cells (six atoms), for which then CRITIC2 computes two times faster than DGRID. CRITIC2 should hence be used for larger cells, if possible.

7.2.2 Generation of Pseudopotentials using AtomPaw

In order to obtain pseudopotentials compatible with ABINIT, the ATOMPAW tool can be utilized[178]. Within the scope of this thesis, ATOMPAW version 4.1.0.6 was used. The easiest approach it to utilize the PAW pseudopotentials provided by *pseudo-dojo.com* in the *.xml*, as they contain the required parameters for ATOMPAW at the end of the file. The required lines from the *.xml* file, exemplified for Ag, are:

```
Ag 47
XC_GGA_X_PBE+XC_GGA_C_PBE scalarrelativistic loggrid 712 200.0 2.5
[...]
XMLOUT
default
END
```

The shown lines must be adapted in a new file *Ag.in* as follows to ensure compatability with ABINIT:

```
Ag 47
GGA-PBE scalarrelativistic loggrid 712 200.0 2.5
[...]
ABINITOUT
prtcorewf
END
```

Running this input via `atompaw < Ag.in` creates two PP files: *Ag.GGA-PBE-paw.abinit* and *Ag.GGA-PBE-paw.abinit.corewf*. The first one is used for the ABINIT calculation, while she second one is required by DGRID to determine the correct amount of core electrons. For that matter, it must be renamed from *Ag.GGA-PBE-paw.abinit.corewf* → *Ag.GGA-PBE-corewf.abinit* and placed in the same folder as the *Ag.GGA-PBE-paw.abinit* file[6].

7.2.3 Calculation of the Domain Averaged Fermi Hole using DGrid

For further analysis, the domain averaged fermi hole (DAFH, see chapter 6.2) of a compound can be calculated using the *[...].sij* file of the previous DGRID calculation.

Orbital Occupation

The process is rather simple, requiring only one DGRID call via: `dgrid-4.7 Al[...].sij fermi 55 > &fermi_log_55_Al` (without any additional input files). The keyword `fermi` signals DGRID to conduct a DAFH calculation, while the consecutive number (`55` in this example) specifies the respective basin. This calculation can be rather time- and memory-intensive, depending on the size and complexity of the used supercell (i.e. the total number of basins). The results are stored in files *Al.abi.FSO-X*, where *X* is the basin number, and a log file, as stated in the DGRID call (*fermi_log_55_Al* in this example). The latter contains the orbital occupations as well as overlaps of the orbitals with other basins:

```
[...]
1705    0.00448400832     -         0.00448400832     -         0.00896801664     -
1706    0.00188781909     -         0.00188781909     -         0.00377563818     -
1707    0.03592576526     -         0.03592576526     -         0.07185153052    0.1
1708    0.03370886251     -         0.03370886251     -         0.06741772503    0.1
1709    0.03266715635     -         0.03266715635     -         0.06533431271    0.1
1710    0.03530788180     -         0.03530788180     -         0.07061576360    0.1
1711    0.03619886444     -         0.03619886444     -         0.07239772888    0.1
1712    0.23315210195    0.2        0.23315210195    0.2        0.46630420390    0.5
1713    0.23594632497    0.2        0.23594632497    0.2        0.47189264994    0.5
1714    0.23649698322    0.2        0.23649698322    0.2        0.47299396644    0.5
1715    0.55687541299    0.6        0.55687541299    0.6        1.11375082598    1.1
--------------------   -----       -------------   -----       -------------   -----
       1.50006612934    1.5        1.50006612934    1.5        3.00013225869    3.0

                +-----------------------------------------------+
                | FERMI ORBITAL INTEGRALS OVER SELECTED BASINS |
                +-----------------------------------------------+
```

[6] The PPs for most elements are already provided in *Project DIO* (see chapter 7.5).

```
-------------------------------------------------------------------------------------------
  FO  /  Basins:       1             [...]            54             55             56
-------------------------------------------------------------------------------------------
 1712             0.00455604021      [...]       0.01393737681  0.24878852029  0.01393880897
 1713             0.00064331688      [...]       0.01457363288  0.25163921114  0.01457230866
 1714             0.00067797926      [...]       0.13372348751  0.25099326930  0.13375628871
 1715             0.00073854835      [...]       0.03350877687  0.57219493261  0.03350526999
-------------------------------------------------------------------------------------------
```

The numbers at the start of each line denote a unique orbital. As the DAFH orbitals are not (necessarily) identical to atomic orbitals, they must be matched to the (corresponding) atomic orbital by considering symmetries, occupations and visualization (if necessary). The top part of the output denotes the orbital occupation (top right entry). Orbital 1715 is filled with about 1.1 electrons, orbitals 1714-1712 with 0.5 each, respectively. As aluminum features s and p orbitals in its valence shell, this splitting strongly indicates orbital 1715 corresponds to the 3s orbital, while 1714-1712 are 3p orbitals. The bottom part displays the overlap of each (significantly) filled orbital with the remaining basins. As the output is the DAFH for basin 55, the overlap with itself is naturally the largest: The 3s orbital (1715) is about 57% localized in the native basin, while the 3p orbitals (1714-1712) are only localized for about 25%. Basins 54 and 56 are basins neighboring basin 55. The 3s orbital only overlaps with about 3% with these basins, while 3p orbital 1714 overlaps about 13%. The remaining 3p orbitals, however, only overlap about 1%. This is expected, as s orbitals are isotropic, while p orbitals are directed. Hence the two lobes of orbital 1714 enter basins 54 and 46, while the lobes of the orbitals 1713 and 1712 face another direction and overlap with different basins (which are omitted in the table above). These values can also be used to calculate the DI contribution of an orbital to a basin pair:

$$\mathrm{DI}_{Orb,\Omega'} = 2 \times q_{Orb} \times \Xi_{Orb,\Omega'}, \tag{7.3}$$

where $\mathrm{DI}_{Orb,\Omega'}$ is the DI (contribution) of orbital "Orb" with basin Ω', q_{Orb} the occupation of orbital "Orb" and $\Xi_{Orb,\Omega'}$ the overlap of orbital "Orb" with basin Ω'. The DI value corresponds to the CRITIC2 definition, hence DI = ES/2. The DI contribution of orbital 1714 (3p-like) towards basin 56 hence amounts to about: $\mathrm{DI}_{1714,\Omega(56)} = 2 \times 0.47 \times 0.134 = 0.126$

Orbital Visualization

In order to visualize the calculated orbitals, another calculation is required, using the following input file *dg_fermi_FSO-55.inp*:

```
:TITLE
:-------------------------------------------------------------------|
:: Al
:-------------------------------------------------------------------|
```

```
:KEYWORDS
:----------------------------------------------------------------------
 basis=Al.abi.FSO-55

:CHOOSE THE DESIRED PROPERTIES
:----------------------------------------

compute=phi 1715 alpha real
compute=phi 1715 alpha imag

:----------------------------------------

vectors
origin: 0. 0. 0.
i-vector:  18.83800000  0.13000000  10.53880000     217
j-vector:   6.40160000  17.72040000 10.53880000    217
k-vector: -0.12120000  -0.08560000  21.28760000     217
END
```

This file is very similar to the *dg.inp* file used in step 3. The main difference lies in the different property to be computed, i.e. **compute=phi 1715 alpha real** and **compute=phi 1715 alpha imag**. These commands tell DGrid to compute the orbital amplitude (**phi**) of orbital 1715 for one spin component (**alpha**) as well as to compute the real and imaginary part of the orbital (which contains the phase information), respectively. As basis, the *FSO* file generated in the previous step has to be stated. Also note that the real-space grid at the end of the file must encompass the whole supercell, and not only the size of the initial unit cell. As DGrid does not compute any values outside the supercell, it is advisable to use an atom located at the center of the unit cell. Executing `dgrid-4.7 dg_fermi_FSO-55.inp` yields two output files *[...]r_r_a* (real part) and *[...]i_r_a* (imaginary part). Figure 7.5 shows the visualization of the real part[7].

[7] A small handguide on how to use ParaView to plot these orbitals: Use ParaView 5.2.0: 1: Load the plugins "elf_mk" and "SMMDGridPlugin" (included in *Project DIO*, see chapter 7.5) via *Tools* → *Manage Plugins...* in the toolbar. 2: Separately drag the *[...]r_r_a* and *[...]i_r_a* files into ParaView and select *DGrid Property File* in the prompt that opens respectively. 3: Click on *Structure* (any will do) in the *Pipeline Browser* and apply the *Plot Structure* filter via *Filters* in the toolbar to show atoms and the unit cell. 4: Select the *Property* object of the *[...]r_r_a* and *[...]i_r_a* files in the *Pipeline Browser* at the same time and apply the *Append Attributes* filter. 5: Select the newly created *AppendAttributes X* object and apply the *Contour Filter*. 6: Select the *Contour X* object and set *Contour By* to the *[...]r_r_a* file and *Coloring* to the *[...]i_r_a* file and adjust the isosurface value to display the orbital with the phase information encoded in color.

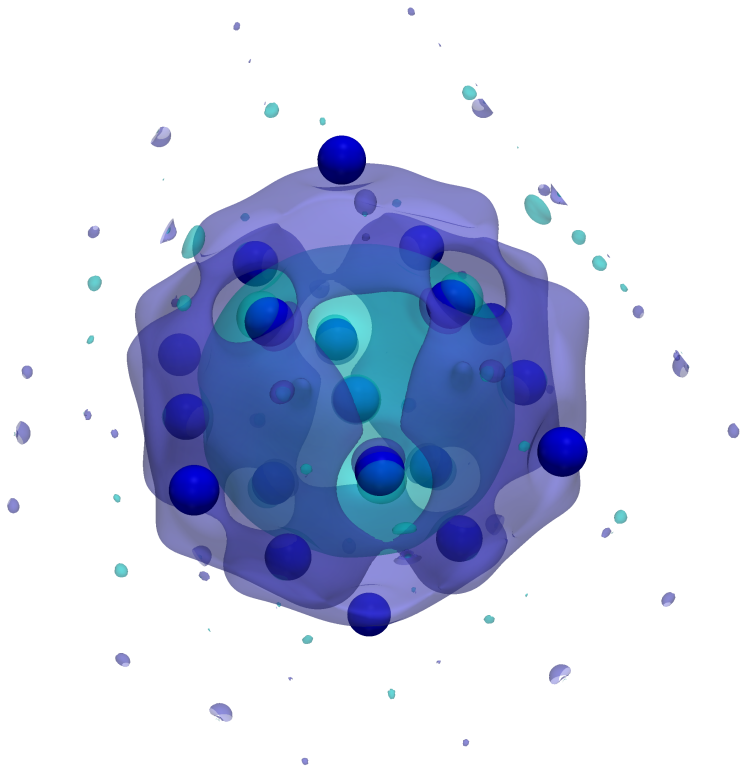


Figure 7.5: Visualization of the Al 1715 DAFH orbital. The blue spheres are Al atoms, an isosurface value of 0.005 is used for the dark blue part of the orbital, while a value of -0.005 is used for the turquoise one. This visualization underlines the 3s orbital character of the DAFH orbital, showing its isotrope character. Created using PARAVIEW version 5.2[179] and the additional plugins "elf_mk" and "SMMDGridPlugin", obtained from Pavlo Golub (these plugins do not seem to be compatible with newer versions of PARAVIEW).

7.2.4 Structural Relaxation using Abinit

Similar to the relaxation performed with QUANTUM ESPRESSO described in chapter 7.1.4, ABINIT also features the option of relaxing the structure of a given unit cell. This section will briefly exemplify the calculation using $KCuCl_3$ in the $Pm\bar{3}m$ structure. The *[...].files* is comparable to the one used for the ES/ET calculation:

```
KCuCl3.in
KCuCl3.out
KCuCl3_xi
KCuCl3_xo
```

```
KCuCl3_x
/path/to/PPs/Pseudos_Paw_DGRID/K.xml
/path/to/PPs/Pseudos_Paw_DGRID/Cu.xml
/path/to/PPs/Pseudos_Paw_DGRID/Cl.xml
```

The *[...].in* file is also almost identical:

```
chksymbreak=0
chkprim=0
#Optimization of the lattice parameters
optcell 2
ionmov  2
ntime  100
ecutsm  0.3
#electron section
nband 29
occopt 3
tsmear 0.005

#Definition of the k-point grid
kptopt 1
nshiftk 1
shiftk  0.0 0.0 0.0
ngkpt 6 6 6
#Definition of the unit cell
acell 3*1.00000
rprim
9.185767 0.0 0.0
0.0 9.185767 0.0
0.0 0.0 9.185767

#Definition of the atom types
ntypat 3
znucl 19, 29, 17
#Definition of the atoms
natom 5
typat 1*1 1*2 3*3
xred

  0.00000000  0.00000000  0.00000000
  0.50000000  0.55000000  0.50000000
  0.00000000  0.50000000  0.50000000
```

```
  0.50000000  0.00000000  0.50000000
  0.50000000  0.50000000  0.00000000

#Definition of the planewave basis set
ecut 30.0
pawecutdg 60.0
#Definition of the SCF procedure
nstep 100
diemac 1000000
```

The new or changed settings are:

- ◇ **chksymbreak=0**: Do not abort calculation when symmetry breaking occurs. This can happen when the symmetry of the cell changes during relaxation.
- ◇ **chkprim=0**: Allow non-primitive cells to be used.
- ◇ **optcell 2**: Enable optimization of lattice vector lengths and angles (i.e. all cell parameters).
- ◇ **ionmov 2**: Optimize positions of ions using the BFGS algorithm (see [177] for more information).
- ◇ **ntime 100**: Maximum number of optimization steps before stopping.
- ◇ **ecutsm 0.3**: Internal parameter for the relaxation process, can usually remain at 0.3 (see [177] for more information).

This time, the y-position of the center Cu atom has been shifted out of equilibrium from 0.5 → 0.55. Running ABINIT with these input files yields the following relative atomic positions:

```
0.0     0.00979313374     0.0 (K)
0.5     0.51016589278     0.5 (Cu)
0.0     0.50997996672     0.5 (Cl 1)
0.5     0.01008104004     0.5 (Cl 2)
0.5     0.50997996672     0.0 (Cl 3)
```

The y-position of the artificially moved Cu has been reduced from 0.55 → 0.51, while for each other site the y-position has increased by 0.01. All y-positions are now 0.01 above the original state. However, due to the translational symmetry of unit cells, a global shift of each atomic site within a unit cell preserves the structure. The relaxation has hence restored the original (equilibrium) positions. The lattice vector lengths (and angles) have remained virtually identical during relaxation.

7.3 Orbital-Resolved Dielectric Function

The Vienna Ab initio Software Package (VASP) is yet another software implementation of DFT (although also featuring the possibility to utilize the Hartree-Fock approximation to solve the Roothaan equations). VASP was (almost) exclusively used to calculate dielectric functions within this thesis, using a slightly modified version 5.4.4 (see following paragraphs), hence the description will be relatively brief. VASP is a commercial software - purchase of a respective license is required[180–183].
LOBSTER (Local-Orbital Basis Suite Towards Electronic-Structure Reconstruction) is a tool that enables the calculation of COHP (Crystal Orbital Hamilton Populations), COOP (Crystal Orbital Overlap Populations) and the projected DOS using plane-wave output from QE, ABINIT or VASP[39, 184]. Within the scope of this thesis, version 4.0.0 was used. Within this chapter, a brief overview on how the (orbital-resolved) dielectric function is calculated using VASP and LOBSTER is illustrated using the example of GeTe in the $R3m$ phase.

Step 1 - Generation of Input Files

VASP requires four input files. The first is called *POSCAR* and contains structural information for the unit cell:

```
Ge1 Te1
1.0
        4.398         0.016         0.011
        2.356         3.714         0.011
        2.356         1.301         3.479
   Ge   Te
    1    1
Direct
     0.994      0.994     0.994
     0.527     0.527      0.527
```

The top entries contain the lattice vectors (horizontally), while the bottom part contain the relative atomic coordinates in the order they are listed after the lattice vectors (in this case Ge first, then Te). The second required input file is named *INCAR*, containing the parameters of the calculation.

```
$system    =  GeTe
PREC       =  Accurate
```

```
ADDGRID   = .TRUE.
LOPTICS=.TRUE.
ISTART=0

ISMEAR = 0
SIGMA=0.005
CSHIFT=0.0


ENCUT     =  550
EDIFF     =  1E-09
IBRION    =  1
ISIF      =  3

LWAVE=.TRUE.
LCHARG=.TRUE.

EMIN=-20.0
EMAX=20.0
NEDOS = 20000
LORBIT=11
LMAXMIX=4

NBANDS = 19
```

A brief description of the (most important) input parameters are given below. A more detailed discussion can be found in the VASP manual[185].

- ◇ **ADDGRID = .TRUE.**: Increases the precision of the augmentation charge calculation by utilizing an additional, high resolution support grid.
- ◇ **LOPTICS=.TRUE.**: The (frequency dependent) dielectric matrix is calculated after the SCF-cycle is converged.
- ◇ **ISTART=0**: Calculation is started from scratch.
- ◇ **ISMEAR = 0**: Use Gaussian smearing.
- ◇ **SIGMA=0.005**: Width of smearing in units of eV.
- ◇ **CSHIFT=0.0**: Sets complex shift η in the Kramers-Kronig transformation to zero (see equation 5.68).
- ◇ **ENCUT = 550**: Cutoff energy in eV (see chapter 5.5.1).

- ◇ **EDIFF = 1E-09**: Break condition for the SCF-cycle in eV. If the energy change between two cycles is smaller than this value, convergence is assumed to be reached.
- ◇ **IBRION = 1**: Chooses a specific algorithm for ionic relaxation (see VASP wiki[185] for more details).
- ◇ **ISIF = 3**: Calculate stress tensor.
- ◇ **LWAVE=.TRUE.**: Save wave function to *WAVECAR* file.
- ◇ **LCHARG=.TRUE.**: Save charge densities to the files *CHGCAR* and *CHG*.
- ◇ **EMIN=-20.0**: Lowest eigenenergy to be calculated.
- ◇ **EMAX=20.0**: Highest eigenenergy to be calculated.
- ◇ **NEDOS = 20000**: Number of grid points for DOS calculation.
- ◇ **LORBIT=11**: Sets correct format for output.
- ◇ **LMAXMIX=4**: l-quantum number up to which the one-center PAW charge densities are passed through the charge density mixer see VASP wiki[185] for more details).
- ◇ **NBANDS = 19**: Number of bands to be considered.

The third file is called *KPOINTS* and simply contains the $\vec{k}$-point grid resolution:

```
Automatic mesh
0
Gamma
24 24 24
0.0 0.0 0.0
```

The first "0" indicates the automatic determination of the number of $\vec{k}$-points, while "Gamma" centers the grid around the Γ-point. The numbers below are the numbers of $\vec{k}$-points in each direction, while the last row can be used to shift the grid. The last file is called *POTCAR* and contains the concatenated pseudopotential data of all elements in the unit cell[8].
Executing VASP in the same directory as these input files creates an output file called *OUTCAR*, containing the regular dielectric function. Calculating the orbital-resolved variant requires some additional steps.

[8] The VASP website offers a variety of compatible PPs[183].

Step 2 - Generation of Explicit $\vec{k}$-Point Grid

In order to avoid discrepancies between LOBSTER and VASP, an explicit list of $\vec{k}$-points must be used (to avoid any potential mismatch due to VASP reducing the effective $\vec{k}$-point grid via symmetry operations). Generating the $\vec{k}$-point list in the correct format is achieved easiest using a small trick. Upon executing VASP, it generates an output file named *IBZKPT*, containing the required list of $\vec{k}$-points, before starting the SCF-cycle. Therefore, while not especially elegant, starting VASP and stopping it again after a couple of seconds is probably the simplest way of obtaining said $\vec{k}$-point list. The contents of the *KPOINTS* file are then replaced by the $\vec{k}$-points from the *IBZKPT* file. Additionally, the list of $\vec{k}$-points must be appended a second time into the same file, with all weights set to zero. The total number of $\vec{k}$-points in the header of the file has to be doubled accordingly. This step is required to accommodate input requirements for the subsequent LOBSTER calculation:

```
Automatically generated mesh
    2626
Reciprocal lattice
    0.00000000000000    0.00000000000000    0.00000000000000           1
    0.04166666666667   -0.00000000000000    0.00000000000000           6
    0.08333333333333    0.00000000000000    0.00000000000000           6
    [...]
   -0.45833333333333    0.50000000000000    0.45833333333333           6
    0.50000000000000    0.50000000000000    0.50000000000000           1
    0.00000000000000    0.00000000000000    0.00000000000000           0
    0.04166666666667   -0.00000000000000    0.00000000000000           0
    0.08333333333333    0.00000000000000    0.00000000000000           0
    [...]
   -0.45833333333333    0.50000000000000    0.45833333333333           0
    0.50000000000000    0.50000000000000    0.50000000000000           0
```

VASP is then executed again to calculate the SCF-cycle and (standard) dielectric function (or rather for the first time, as the first (incomplete) run merely served to obtain the $\vec{k}$-point list). At this point, the slight modification to VASP mentioned earlier, comes into play. In order to orbitally resolve the dielectric function, the transition matrix elements have to be saved to an output file. While VASP internally computes these, there is no option to save them in the standard version of VASP. This feature was implemented by Jean-Yves Raty, by modifying the *linear_optics.F* file of the VASP source code to save the matrix elements into a file called *Matrix_elements001*. This modification is not publicly available, but can be requested by contacting Jean-Yves Raty from Liège université (it is also included in *Project DIO*, see chapter 7.5). The content of *Matrix_elements001* looks like this:

```
 ikpt, band1, band2, E1, E2, ETRANS, MATRIX, WEIGHT
    1     1     2  -0.788D+01   0.185D-01   0.789D+01   0.264D-02   0.000D+00   0.000D+00
    1     1     3  -0.787D+01   0.327D+01   0.111D+02   0.461D-02   0.000D+00   0.000D+00
```

```
1    1    4  -0.787D+01   0.339D+01   0.112D+02   0.153D-01   0.000D+00   0.000D+00
1    1    5  -0.787D+01   0.339D+01   0.112D+02   0.249D-01   0.000D+00   0.000D+00
[...]
```

The first column, **ikpt**, states the $\vec{k}$-point index, **band1** and **band2** the indices of the bands between the transition occurs, **E1** and **E2** the energies of the respective bands at that specific $\vec{k}$-point, **ETRANS**=**E1**-**E2** the energy difference and **MATRIX** the transition matrix element and **WEIGHT** a weighting factor of the respective transition:

$$\textbf{WEIGHT} = (N_{\text{Initial}} - N_{\text{Final}}) \times \frac{W_{\vec{k}}}{\#K}, \tag{7.4}$$

where N_{Initial} and N_{Initial} are the electronic occupations of the initial and final state (band) of the respective transition, $W_{\vec{k}}$ the weight of the $\vec{k}$-point and $\#K$ the total number of considered $\vec{k}$-points. The VASP calculation also creates a **PROCAR** file, containing orbital projections of each $\vec{k}$-point and band and the occupation of each band. However, as LOBSTER generally produces more precise orbital projections, only the occupations from the **PROCAR** file are used.

Step 3 - Generation of Lobster Input File

The input file for the LOBSTER calculation is relatively simple, as LOBSTER extracts most of the information from the VASP calculation. The file *lobsterin* contains for the example of GeTe:

```
COHPstartEnergy -30
COHPendEnergy 30
usebasisset pbeVaspFit2015
createFatband Ge 4s 4p_x 4p_y 4p_z
createFatband Te 5s 5p_x 5p_y 5p_z
doNotIgnoreExcessiveBands
basisfunctions Ge 4s 4p
basisfunctions Te 5s 5p
cohpbetween atom 1 and atom 2
```

The only non-static parameters are **createFatband** and **basisfunctions**, after which the element and the to be calculated atomic orbitals need to be stated. LOBSTER can then be executed via `lobster < lobsterin`, creating *FATBAND_[...]* files for each atomic orbital.

Step 4 - Combining the Outputs

Finally, another tool created by Jean-Yves Raty, called *decomp_epsilon_lobsterfiles*, is used to combine the various output files and create the orbital-resolved dielectric function. This tool reads data from the **PROCAR** file. However, the number of stated $\vec{k}$-points in the header of this file must be halved, as this number still includes the doubled amount of $\vec{k}$-points from step 2. Furthermore, an input file named *in.decomp_lobster* is required:

```
18
FATBAND_Te2_5s.lobster
FATBAND_Te2_5p_x.lobster
FATBAND_Te2_5p_y.lobster
FATBAND_Te2_5p_z.lobster
VOID
VOID
VOID
VOID
VOID
FATBAND_Ge1_4s.lobster
FATBAND_Ge1_4p_x.lobster
FATBAND_Ge1_4p_y.lobster
FATBAND_Ge1_4p_z.lobster
VOID
VOID
VOID
VOID
VOID
```

The first line contains the total number of projections, that is #Sites × 9, as one s, three p and five d orbitals are considered for each atomic site. The name of each orbital-resolved LOBSTER band file is then stated afterwards, using "VOID", if no projection exists for the specific orbital (for GeTe, no d orbitals exist)[9]. Executing `decomp_epsilon_lobsterfiles > DECOMP_ABS` (and stating "9" for the number of projections) yields the *DECOMP_ABS* files with the orbital-resolved dielectric function (imaginary part).

```
nr of projections ?
nr proj =           9
       1313           24           2
 0.000D+00  0.000D+00  0.000D+00  0.000D+00  0.000D+00  0.000D+00  0.000D+00  0.000D+00
 0.502D-01  0.000D+00  0.000D+00  0.000D+00  0.000D+00  0.000D+00  0.000D+00  0.000D+00
 0.100D+00  0.000D+00  0.000D+00  0.000D+00  0.000D+00  0.000D+00  0.000D+00  0.000D+00
 [...]
```

[9] f electrons and higher orbitals have not been implemented so far.

Ignoring the header, the first column denotes the energy in eV, while the last column is "dead", as the values it contains are incorrect. Columns 2 to 7 contain the relative contribution of the $(s \Leftrightarrow s)$, $(s \Leftrightarrow p)$, $(s \Leftrightarrow d)$, $(p \Leftrightarrow p)$, $(p \Leftrightarrow d)$ and $(d \Leftrightarrow d)$ transitions, respectively. Plotting the dielectric function yields figures 7.6 and 7.7.

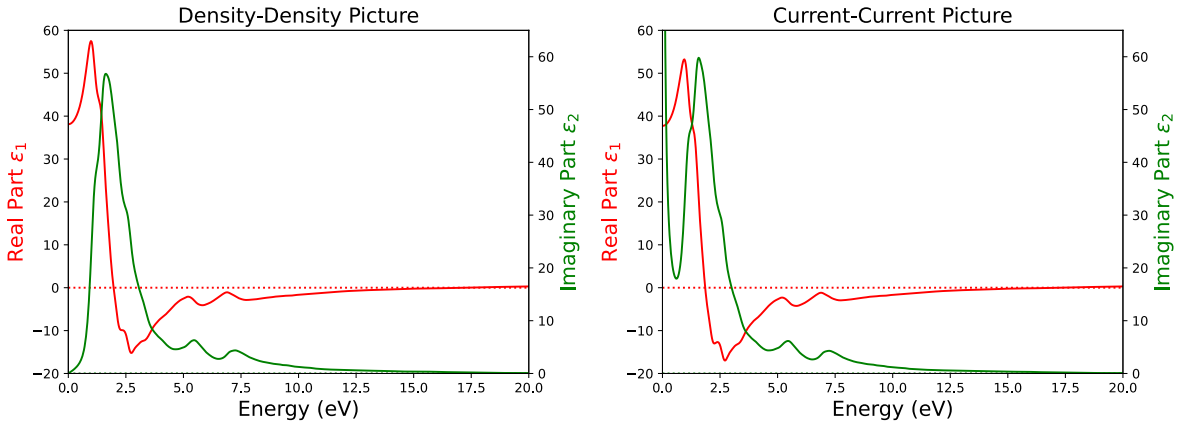


Figure 7.6: Real (red) and imaginary (green) part of the dielectric function of GeTe $R3m$ in the density-density picture (left) and the current-current picture (right).

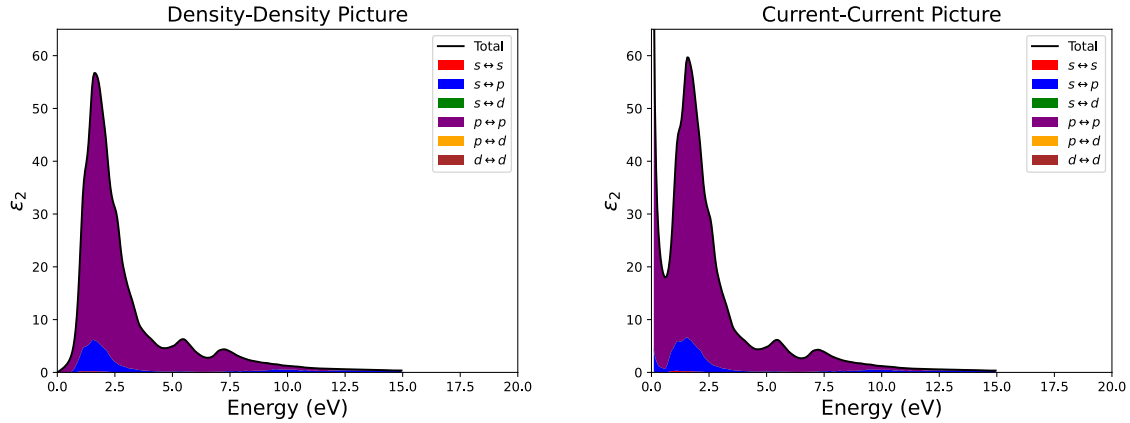


Figure 7.7: The orbital-resolved imaginary part of the dielectric function of GeTe $R3m$ in the density-density picture (left) and the current-current picture (right). The labels indicate transitions in both direction, e.g. $(s \Leftrightarrow p)$ includes transitions from s states to p states, as well as transitions from p states to s states.

Besides the so-called density-density picture, VASP also calculates the current-current picture. Both formalisms yield almost identical results, despite a spike in the imaginary part of the dielectric function for the current-current picture for small energies close to 0 eV. Figures 7.6 (right) and 7.7 (right) show the results for GeTe using the current-current picture.

While the dielectric function (tensor) within the density-density and current-current picture

are formally equivalent, they are computed in a different fashion:

$$\epsilon\left[\chi_{\text{dens.}}(\vec{q},\omega)\right] = 1 - \frac{4\pi}{|\vec{q}\,|^2}\chi_{\text{jj}}(\vec{q},\omega), \quad \chi_{\text{dens.}}(\vec{q},\omega) = \frac{\partial\rho(\vec{q},\omega)}{\partial\phi(\vec{q},\omega)}, \tag{7.5}$$

$$\epsilon\left[\chi_{\text{curr.}}(\vec{q},\omega)\right] = 1 - \frac{4\pi}{\omega^2}\chi_{\text{jj}}(\vec{q},\omega), \quad \chi_{\text{curr.}}(\vec{q},\omega) = c\frac{\partial j(\vec{q},\omega)}{\partial A(\vec{q},\omega)}, \tag{7.6}$$

assuming a small perturbation with small momentum transfer $\vec{q}$. Furthermore is ρ the charge density, j the current density (considering only the component parallel to $\vec{q}$). The χ are the response functions to the macroscopic, external perturbations. These perturbations are $\partial\phi$ (scalar potential) and ∂A (strictly a vector potential, restricted to be parallel to $\vec{q}$ as well). Both response functions are linked by the continuity equation $\vec{q}\cdot\vec{j} = \omega\rho$:

$$|\vec{q}\,|^2\chi_{\text{curr.}}(\vec{q},\omega) = \omega^2\chi_{\text{dens.}}(\vec{q},\omega) \tag{7.7}$$

In the limit of $\vec{q} \to 0$, the dielectric functions obtained by equations 7.5 and 7.6 should yield identical results. However, at the non-analytic point $(\vec{q},\omega) = (\vec{0},0)$, the limit has to be evaluated, with the results depending on the chosen direction[186, 187].

While formally equivalent, the numerical equivalence is not guaranteed in general. Usually, the density-density picture is examined in the literature and other publications, as it tends to be more stable to compute numerically. Therefore, it is also used in this thesis. However, the density-density picture suppresses the Drude term of the imaginary part of the dielectric function for metallic systems. The current-current picture (or rather its implementations in the respective DFT software packages) can reproduce it to some extent. A more detailed perspective on the differences and similarities between the density-density and current-current picture can be found in [186, 187].

7.4 The Effective Coordination Number and ECoN Weighted ES

The effective coordination number (ECoN) denotes a distance-weighted average of the number of neighbors of a given atomic site and is a suitable gauge to quantify trends in the atomic arrangement of locally distorted systems. The ECoN value for each atomic site can be calculated according to:

$$\mathrm{ECoN} = \sum_m \exp\left(1 - \left(\frac{d_m}{d_\mathrm{r}}\right)^6\right), \tag{7.8}$$

where d_m is the distance to the m-th neighbor, and d_r is an effective distance defining the first coordination shell (see figure 7.8):

$$d_\mathrm{r} = \frac{\sum_m d_m \exp\left(1 - \left(\frac{d_m}{d_1}\right)^6\right)}{\sum_m \exp\left(1 - \left(\frac{d_m}{d_1}\right)^6\right)} \tag{7.9}$$

The ECoN averaged ES can be calculated in an almost identical fashion, where each ES value is weighted by the ECoN contribution corresponding to the respective bonding partner and its distance:

$$\mathrm{ES}_\mathrm{ECoN} = \frac{\sum_m \mathrm{ES}_m \exp\left(1 - \left(\frac{d_m}{d_\mathrm{r}}\right)^6\right)}{\mathrm{ECoN}} \tag{7.10}$$

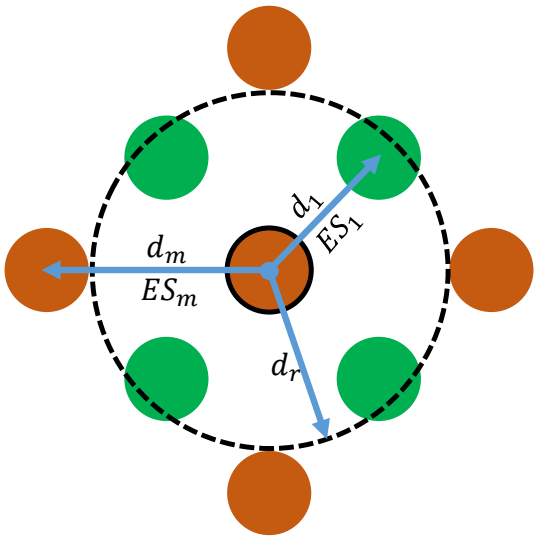


Figure 7.8: Sketch of the relevant distances and ES values to calculate the effective coordination number (ECoN) and the ECoN averaged ES.

7.5 Project DIO

The process of setting up the calculations described in the previous sections is in principle relatively easy, but can become tedious and repetitive for large numbers of calculations. Each compound should generally be relaxed before the other properties can be calculated. This means for each calculation the input files have to be generated and the unit cell parameters inserted (and potentially extracted first from the relaxation calculation). If done manually, this procedure is prone to typos and human error, which will in the best case cause the calculation to crash, in the worst-case yield deceptively skewed results.
To alleviate this complexity, the *DFT Integration Optimization (DIO) project* has been developed within the scope of this thesis. *Project DIO* is a software package that enables the user to generate the input files for a multitude of DFT and QTAIM calculations, requiring only a structure file in the CIF (Crystallographic Information File) or VASP/POSCAR format as input. It also produces shell scripts to execute the respective calculation with minimum effort. *Project DIO* consists of two automation layers.

Layer 1

The first layer includes the functionality to generate the input files and start individual calculations, requiring only the unit cell data. The options are:

1. **Unit Cell Relaxation using QE**: As described in section 7.1.4, a relaxation of the unit cell provided is carried out using QUANTUM ESPRESSO. This is done by six consecutive calculations, each optimizing only a subset of parameters, until all are optimized at once in the last. This is done to improve quality of relaxation and to reduce the chance of crashes. The output is then converted to a new unit cell structure file in the POSCAR/.vasp format, enabling easy visualization and ensuring compatability with other calculation options provided by *Project DIO*.

2. **Unit Cell Relaxation using Abinit**: As described in section 7.2.4, a relaxation of the provided unit cell is carried out using ABINIT. In contrast to the relaxation using QE however, only a single calculation is conducted, optimizing all parameters at once. After the calculation is done, the results are again converted into a structure file in the POSCAR/.vasp format.

3. **ES/ET Calculation using QE and Critic2**: Calculates ES and ET using QUANTUM ESPRESSO and CRITIC2, as described in section 7.1.2. Optionally, the Born effective charge Z^* and the dielectric constant ϵ_∞ can be calculated on top, which can be activated in the script starting the calculation, if so desired[10]. After the calculation is finished,

[10]The script to generate the inputs is based on code initially provided by Jean-Yves Raty.

the most relevant outputs are compiled into a report file[11], which states the TET values, localization indices, bond orders, the ECoN and the ECoN averaged ES values for each site, as well as the stress tensor and convergence behavior as a final check of the validity of the calculation. The Z^* and ϵ_∞ tensors are included as well, if calculation was enabled.

4. **ES/ET Calculation using Abinit and DGrid**: Calculates ES and ET using ABINIT and DGRID, as described in section 7.2.1. After calculation, the DGRID output is reformatted into a more legible file, similar to the output produced by CRITIC2. This includes the mapping of the basin numbers to atomic sites, as well as ordering all bond partners by distance. Additionally, a report file similar to the one created after the QE/CRITIC2 calculation is created.

5. **Band Structure and DOS Calculation using QE**: Calculates the band structure and density of states using QUANTUM ESPRESSO (see section 7.1.3). The required high-symmetry points and Brillouin zone paths can also be generated automatically, featuring either the use of the SeeK-path[188] or Pymatgen library[189]. In case a band gap is detected, the hole and electron effective masses are calculated by fitting a parabola to the maximum and minimum of the valence and conduction band, respectively. Furthermore, the DOS effective mass m^*_{DOS} is evaluated by fitting the DOS. In addition to a file containing the band structure and DOS data, band structure and DOS are also directly plotted for an easy overview of the results. In the event the effective mass tensor is required, an experimental feature can be enabled that uses a finite differences method to compute the effective mass tensor for each high-symmetry point (HSP). This is done by resampling the $\vec{k}$-point grid around and close to each HSP to obtain the change in energy in each $\vec{k}$-direction. This feature uses a modified version of software written by Alexandr Fonari and Christopher Sutton[190] and a parser written by Alireza Faghaninia. The results of the tensor calculations should be taken with caution, as they do not seem to be consistent with the literature all the time. Further testing is required.

6. **FAT Bands and Orbital-Resolved DOS Calculation using QE**: In principle similar to the non-orbital-resolved option and the calculation described in section 7.1.3, the FAT bands and orbital-resolved DOS are computed, as well as the orbital-resolved integrated DOS (IDOS). All results are saved to a file for later use, as well as plotted automatically.

7. **Calculation of the JDOS**: Calculates and plots the joint density of states using QUANTUM ESPRESSO.

8. **Calculation of Z^* and ϵ_∞ using QE**: Calculates the Born-Effective Charge Z^* and the dielectric constant ϵ_∞ using QUANTUM ESPRESSO, similar to the optional calculation of these properties available for the ES/ET calculation using QE (see above). The output is a report file containing only the relevant tensors.

[11] The report file is named *#Final_report_[...].txt*, where *[...]* is the name of the structure file used to generate the input.

9. **Calculation of the Total Energy / SCF-Cycle**: Only calculates the SCF-cycle and consequently the total energy of a system (see section 7.1.1). This can be used to compare the total energies of variations of similar unit cells (e.g. different degrees of distortion) or in the event another calculation not implemented in *Project DIO* is supposed to be carried out next.

10. **(Orbital-Resolved) Dielectric Function using Vasp and Lobster**: Calculates the orbital-resolved and regular dielectric function $\epsilon(\omega)$ using VASP and LOBSTER, as exemplified in section 7.3. The results are saved to a file for later use, as well as plotted automatically.

11. **Effective Mass Tensor utilizing the KP-Method and Vasp**: An alternative way of calculating the effective mass tensors, by utilizing the KP-method and an implementation written by Oleg Rubel et al.[191]. This option is experimental and requires more testing, as the results do not seem to be consistent with the literature in all cases[12]. A report file is generated with tensors for each band and HSP.

Besides these main functionalities, *Project DIO* also includes some "little helper" tools, that can be handy however:

1. **LEA Algorithm**: The algorithm described in chapter 2.3.1. This tool calculates the effective coordination number of each site in a given structure, as well as the elemental contributions (i.e. how much a given element in a unit cell contributes to the total ECoN value). It also features the option of searching for cutting planes that leave the ECoN and/or coordination number (the number of nearest neighbors) intact. This can be helpful in identifying layered/2D compounds.

2. **DAFH Input/Output Collection**: A collection of tools is included that simplifies the calculation of the DAFH orbitals. This incorporates the generation of the input files, a tool that creates input files with a reduced number of basins to further reduce RAM/compute time requirements (so-called *Fake Basins*) and a tool that reformats the DAFH orbital output into a MICROSOFT EXCEL compatible *.csv* file. Additionally, a more sophisticated tool exists that prepares DGRID input files with a shifted grid so that each basin for which a DAFH orbital is calculated is perfectly centered in the supercell. This last tool was developed by Jakob Lötfering.

An attempt is made to guess the system-specific inputs automatically, but these can be modified manually before the calculation is started. In order to decide whether a compound is metallic or insulating, *Project DIO* connects to the Materials Project (MP) database[80] and searches for compounds with similar stoichiometry, assuming they will show similar behavior in general (which is not always correct of course). If no compound is found, or *Project DIO* cannot connect to the Materials Project[13], a metal is assumed for an uneven number

[12] There could be a problem with the ordering of degenerate bands, i.e. band A becomes band B and vice versa after crossing the degenerate maximum or minimum, mixing up the elements of the two resulting tensors.
[13] Due to maintenance, etc.

of valence electrons, otherwise an insulator. This information is then used to set properties like *occupations* and *smearing* in QE to reasonable values. Furthermore, the $\vec{k}$-point grid density is calculated as follows:

$$\#\text{Sites} \times n_\text{x} \times n_\text{y} \times n_\text{z} \begin{cases} \geq 100 & E_\text{G} > 0\,\text{eV (Insulators)} \\ \geq 1000 & E_\text{G} = 0\,\text{eV (Metals)} \end{cases} \tag{7.11}$$

where #Sites is the number of sites in the unit cell and $n_{\text{x,y,z}}$ are the number of $\vec{k}$-points for each direction. The sampling density $n_{\text{x,y,z}}$ in each direction is furthermore inversely proportional to the magnitude of the corresponding real-space lattice vector. This allows the shape of the Brillouin zone to be taken into account, i.e. shorter dimensions are sampled by fewer $\vec{k}$-points, while larger dimensions are attributed more $\vec{k}$-points. For some (although not all) calculations included in *Project DIO*, there is an option to (manually) activate the Grimme-D3 van der waals correction[79].

Layer 2 - JoJo

While layer 1 streamlines the DFT and QTAIM calculations, layer 2 aims to (almost) fully automate it. Layer 2 consists of one additional tool called *Juggler of Joint Operations* (*JoJo*). *JoJo* can be instructed to carry out any of the calculations of layer 1 for all compound structure files in a directory. Within the settings of *JoJo*, the type of relaxation (QE/Abinit/no relaxation) as well the consecutively executed main calculation can be set (see figure 7.9). Limits regarding maximum storage use and file number on the drive as well as the maximum number of calculations to be run at the same time can also be configured. When started, *JoJo* begins the execution of calculations for a number of compounds up to the maximum number specified in the settings. It will then periodically check on the calculation to determine whether it has finished, or crashed, and label the calculation accordingly. Afterwards (if activated), *JoJo* deletes the memory-intensive temporary files (like the *.sij* files for the ES/ET calculations) to keep disk storage available for other calculations. Then a new calculation is started, repeating this cycle until all compounds in the input directory have been processed.
Using *JoJo*, ES/ET values for about 4000 and FAT Bands (+orbital-resolved DOS) for about 2000 compounds were calculated without any manual input.
The code of *Project DIO* was written and tested for the RWTH Compute Cluster, but should work with other compute clusters or regular workstations running Linux[14].

[14] Some compiled programs like Vasp and Abinit etc. would probably need to be recompiled though, or the paths changed to the local installations, if already present. Also, for use on a workstation, SLURM-specific functions of *JoJo* need to be adapted.

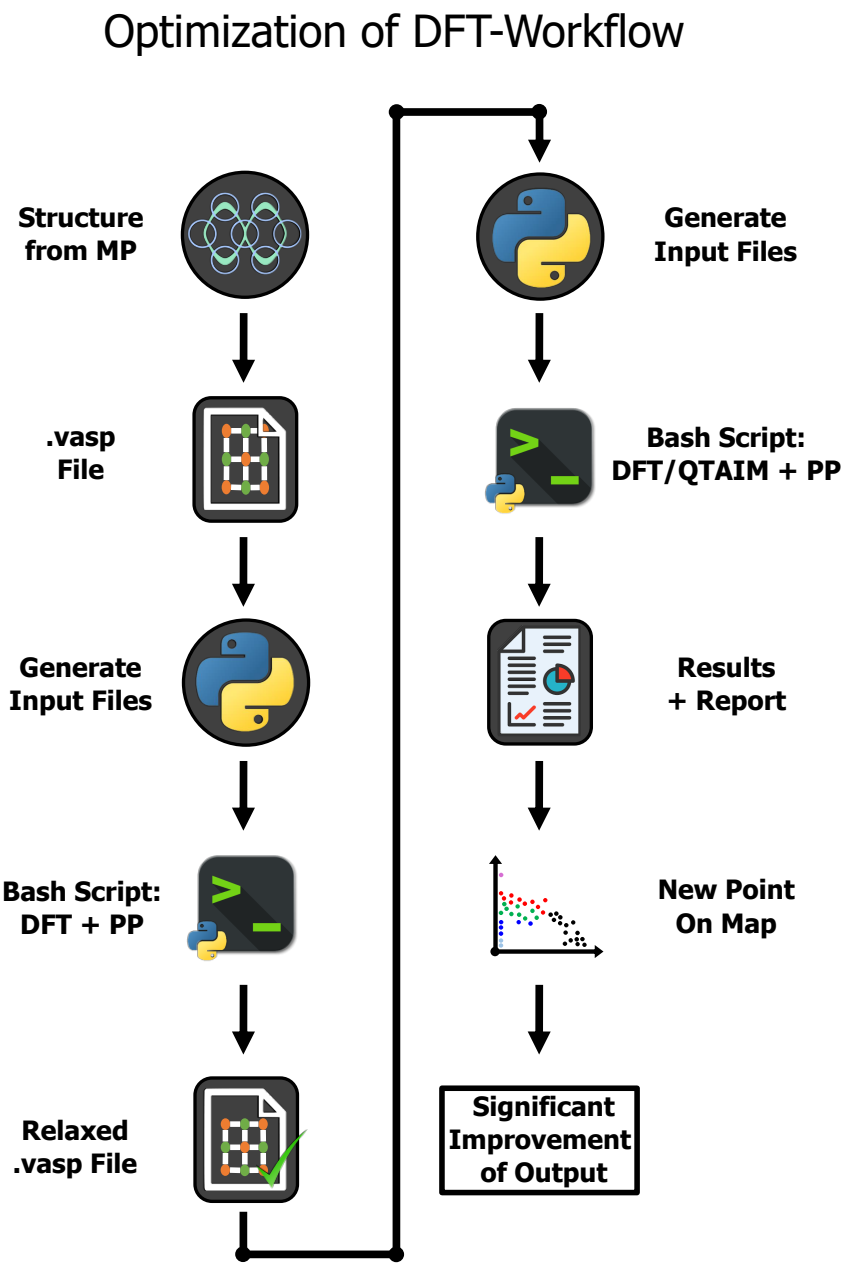


Figure 7.9: DFT Workflow of *JoJo*, exemplified for an ES/ET calculation. The sequence from the 2nd to the 8th icon corresponds to one layer 1 calculation (ES/ET), while layer 2 (*JoJo*) schedules and monitors a set of such calculations for multiple compounds at the same time.

Appendix

Chapter 8

Additional Figures

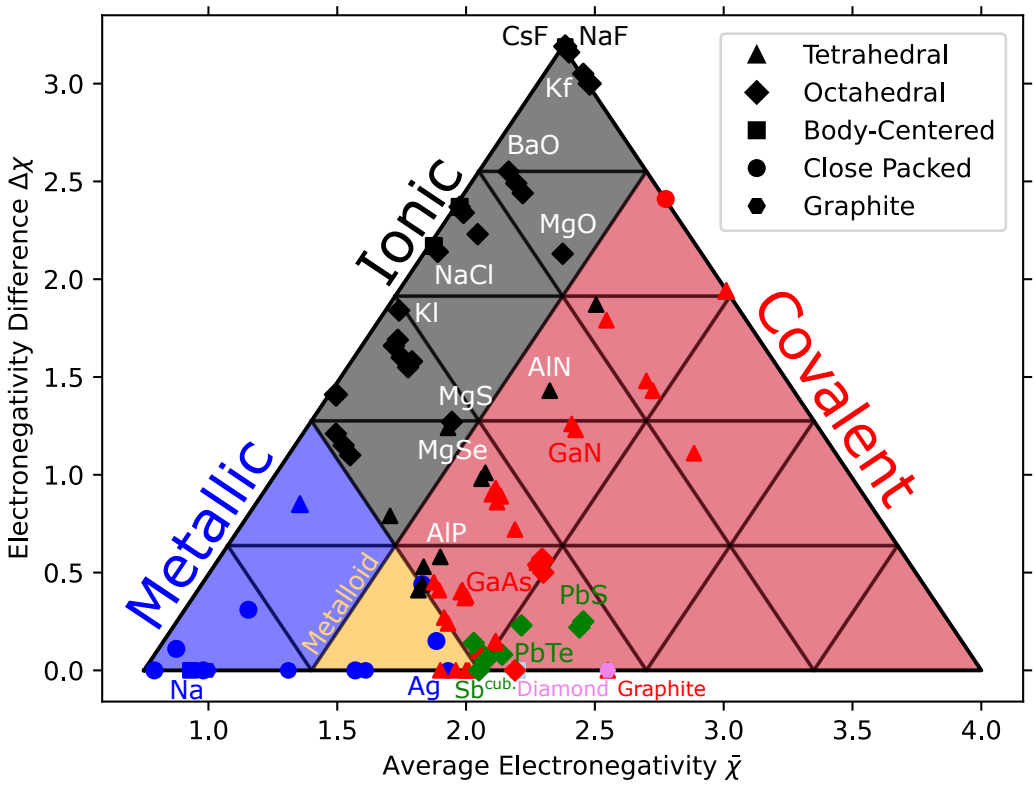


Figure 8.1: The van Arkel-Ketelaar triangle uses the average electronegativity $\bar{\chi} = \frac{\chi_A + \chi_B}{2}$ and the electronegativity difference $\Delta\chi = |\chi_A - \chi_B|$ to classify bonding types[192, 193]. The colors of the datapoints reflect the bonding type: Covalent (red), metavalent (green), ionic (black), metallic (blue), resonant (violet). As structural information is not taken into account, allotropes such as diamond and graphite are located at the same position. No distinct region for metavalent compounds emerges, while the general separability of the classical bonding types is also relatively weak. Also published in [10].

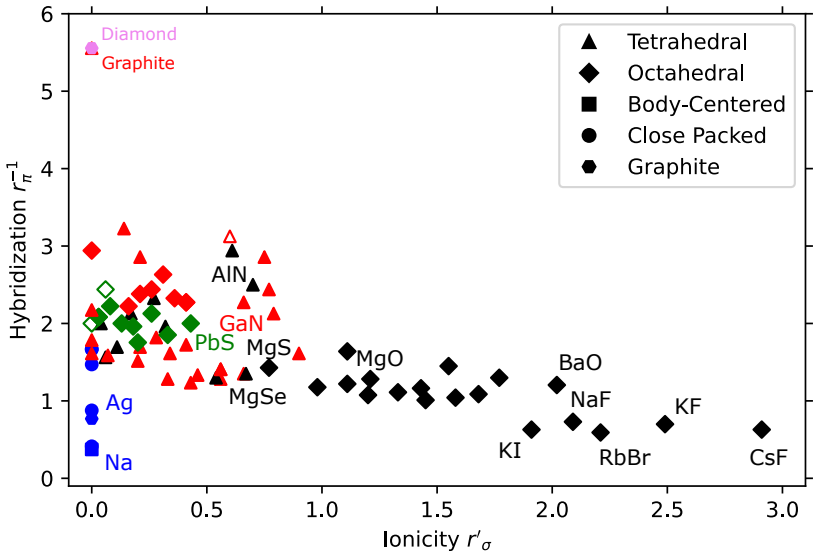


Figure 8.2: The Littlewood map uses the ionicity $r_\sigma = |r_\mathrm{p}^\mathrm{A} - r_\mathrm{p}^\mathrm{B}|$ and hybridization $r_\pi^{-1} = \left[\left(r_\mathrm{p}^\mathrm{A} - r_\mathrm{s}^\mathrm{A}\right) + \left(r_\mathrm{p}^\mathrm{B} - r_\mathrm{s}^\mathrm{B}\right)\right]^{-1}$, to classify bonding types, where r_s and r_p are the valence radii of the s and p orbitals respectively[194]. The colors of the datapoints reflect the bonding type: Covalent (red), metavalent (green), ionic (black), metallic (blue), resonant (violet). As structural information is not taken into account, allotropes such as diamond and graphite are located at the same position. Also published in [10].

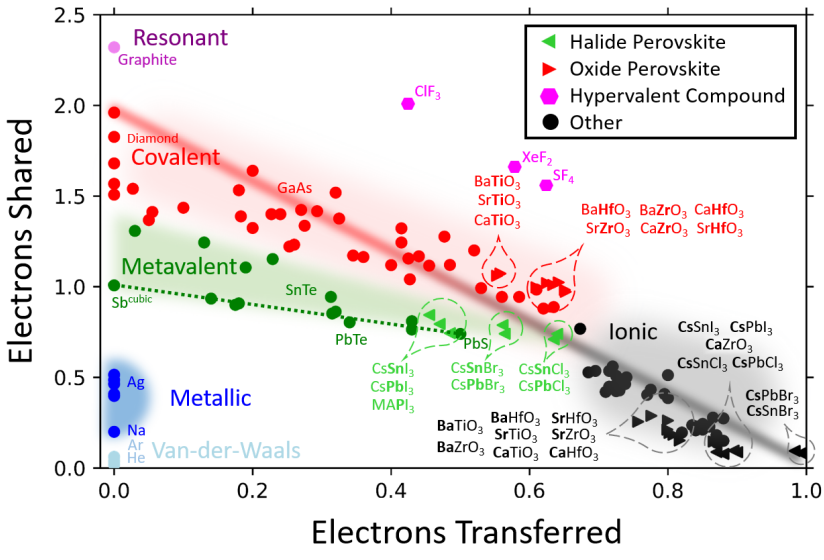


Figure 8.3: ES/ET map featuring perovskites and hypervalent compounds. Hypervalent compounds computed by Jean-Yves Raty. Modified from [3].

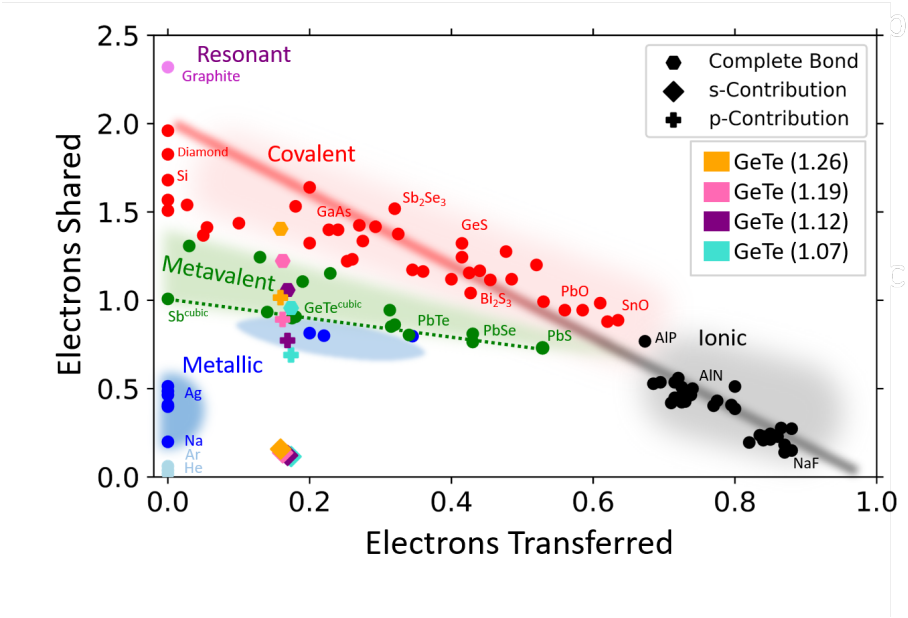


Figure 8.4: ES/ET map with orbital contributions for GeTe with different degrees of distortion. The s-contributions are denoted by diamonds, the p-contributions by crosses and the complete bond by hexagons. The number in brackets in the legend denotes the short to long bond ratio: $\mathrm{PD} = \frac{d_{\mathrm{long}}}{d_{\mathrm{short}}}$. As already shown for the PbX systems and the different GeTe phases in figure 2.6, the s-contribution is minor and remains mostly constant, while the changes in ES originate almost exclusively from increased p-contributions. The ET/ES values of the respective full bonds are: $\mathrm{PD} = 1.07 \rightarrow (0.17/0.96)$, $\mathrm{PD} = 1.12 \rightarrow (0.17/1.06)$, $\mathrm{PD} = 1.19 \rightarrow (0.16/1.22)$, $\mathrm{PD} = 1.26 \rightarrow (0.16/1.40)$. Tables 11.3 to 11.6 show the respective DAFH values.

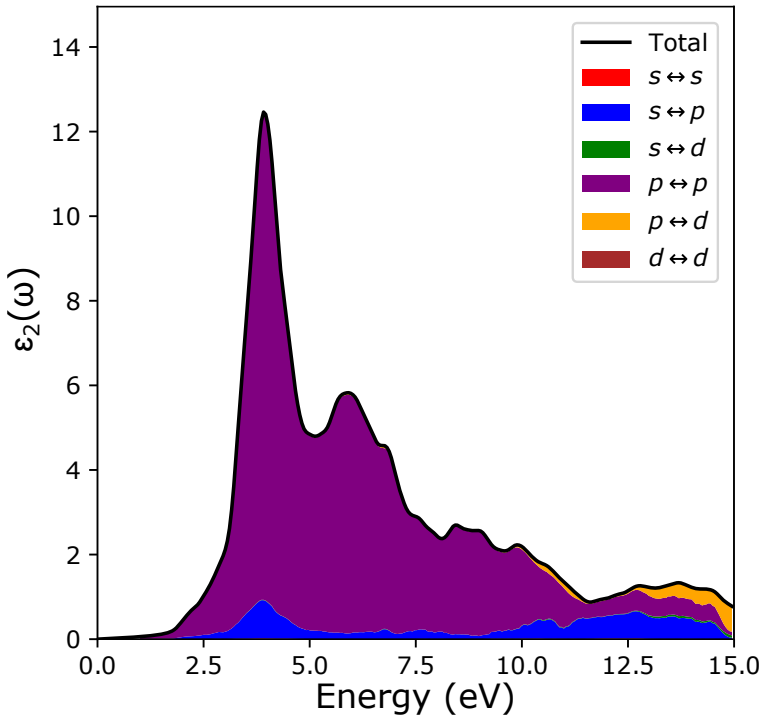


Figure 8.5: Orbital-resolved imaginary part of the dielectric function of PbO.

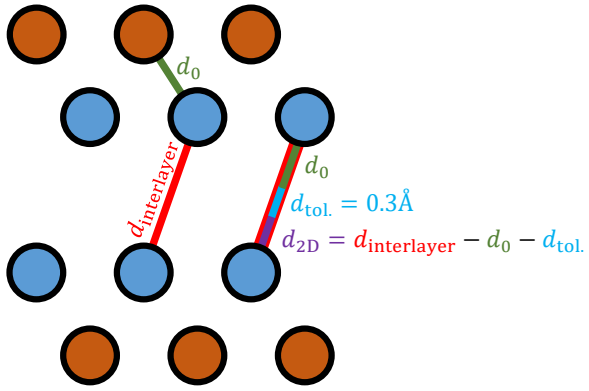


Figure 8.6: Illustration of the distances relevant for the 2D criterion by Stevanović et al.[76].

Chapter 9

Journal Covers

Modern art = I could do that
+ Yeah, but you didn't.

Craig Damrauer

Within the scope of this thesis, several cover images for scientific journals have been designed to accompany corresponding publication, as well as an alternative cover for this thesis. They are presented here, along with a photograph from the lecture "Physik für Maschinenbauer", for no particular reason.

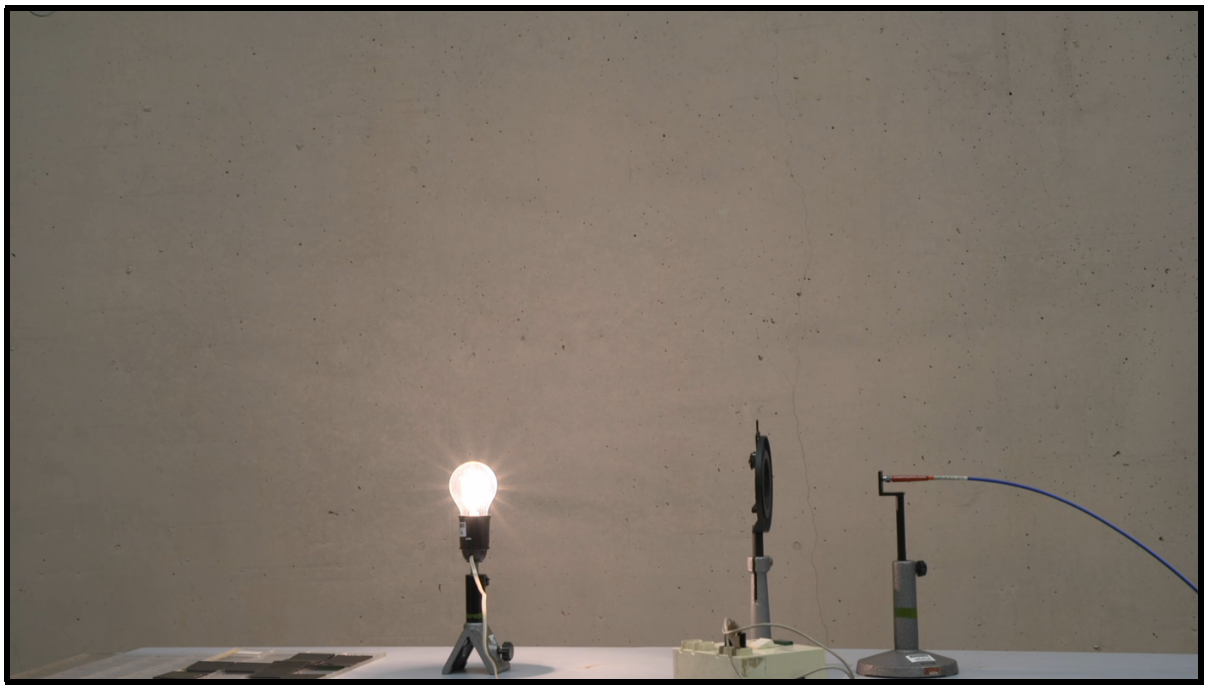


Figure 9.1: Solitude of the Astigmatism - Impression from the "Physik für Maschinenbauer" (Physics for Mechanical Engineers) lecture in the winter semester 2020. Experiment to measure thermal radiation: Light bulb, aperture and fiber-optic cable connected to a spectrometer.

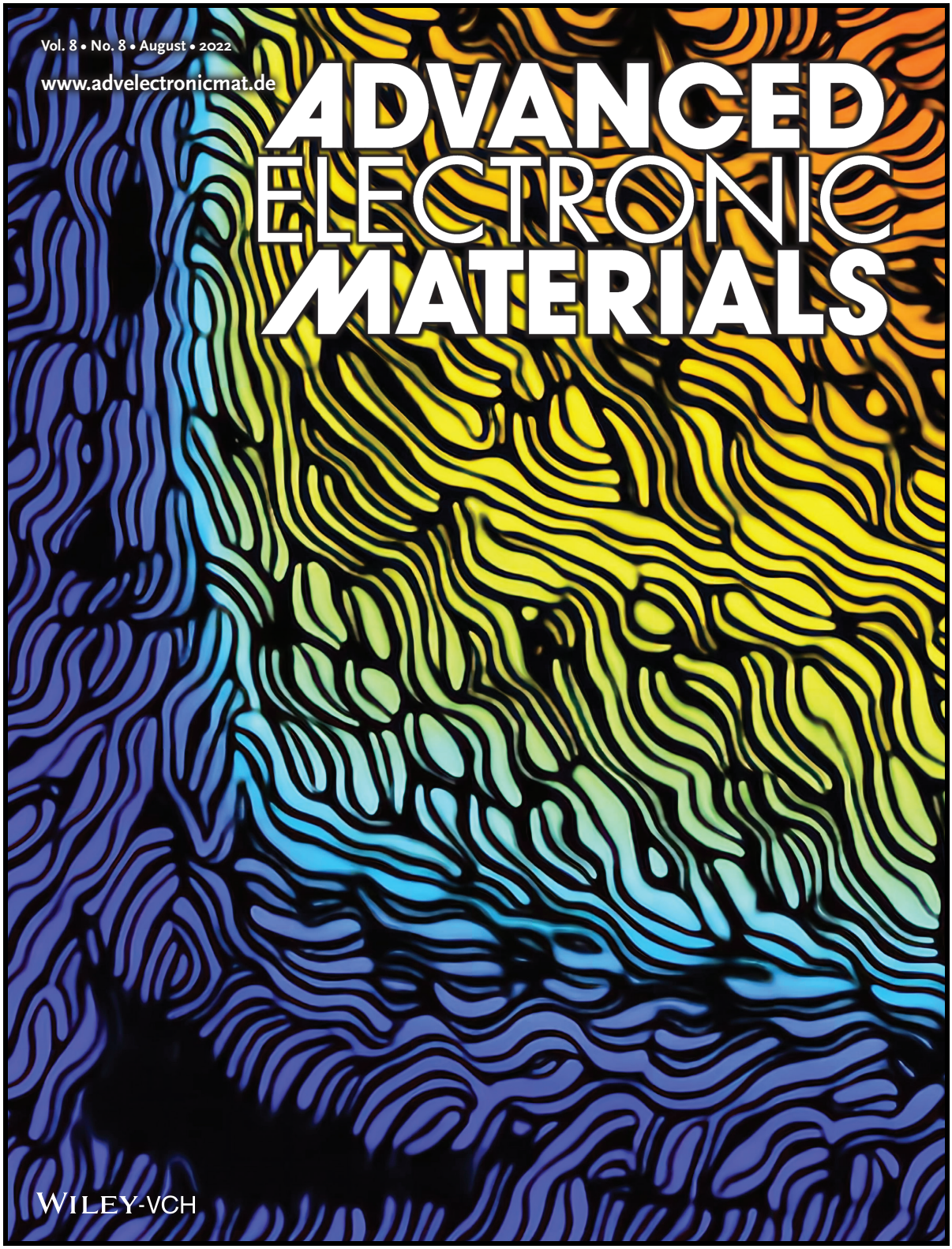


Figure 9.2: Phase Ocean - Advanced Electronic Materials. This cover has been designed for the publication *Tailoring Crystallization Kinetics of Chalcogenides for Photonic Applications* by Maximillian Müller et al.[195] and was published as the front inside cover. It was inspired by figure 3 of this publication and uses a method of transfer learning to achieve this visual effect. Name reference to [196].

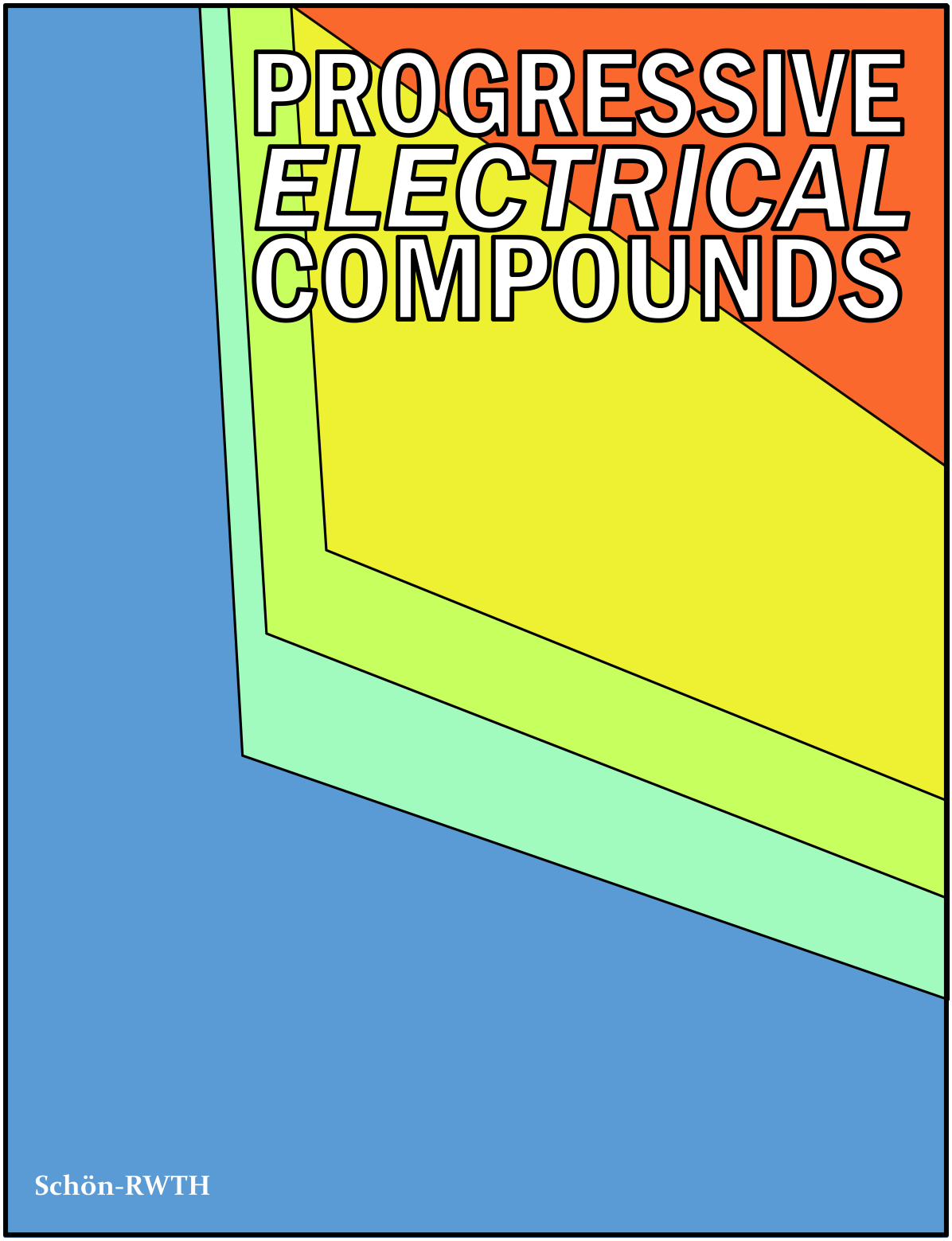


Figure 9.3: Colorful Experience - Unpublished. This cover has been designed for the publication *Tailoring Crystallization Kinetics of Chalcogenides for Photonic Applications* by Maximillian Müller et al.[195]. It was inspired by figure 3 of this publication. Name reference to [197].

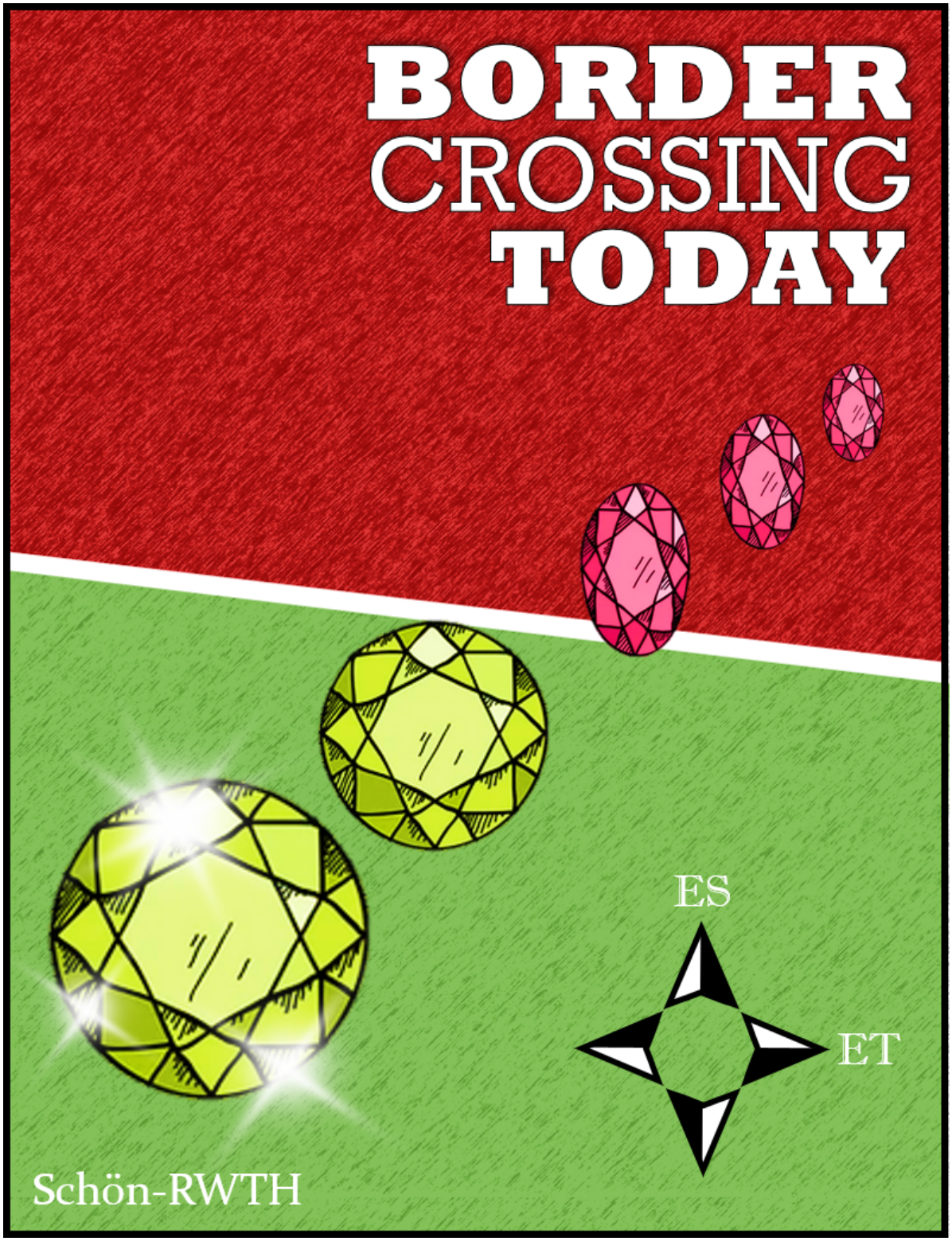


Figure 9.4: Gemstones Are Unbreakable - Unpublished. This cover has been designed for the publication *Tailoring Crystallization Kinetics of Chalcogenides for Photonic Applications* by Maximillian Müller et al.[195]. It was inspired by figure 5 of this publication. Name reference to [198].

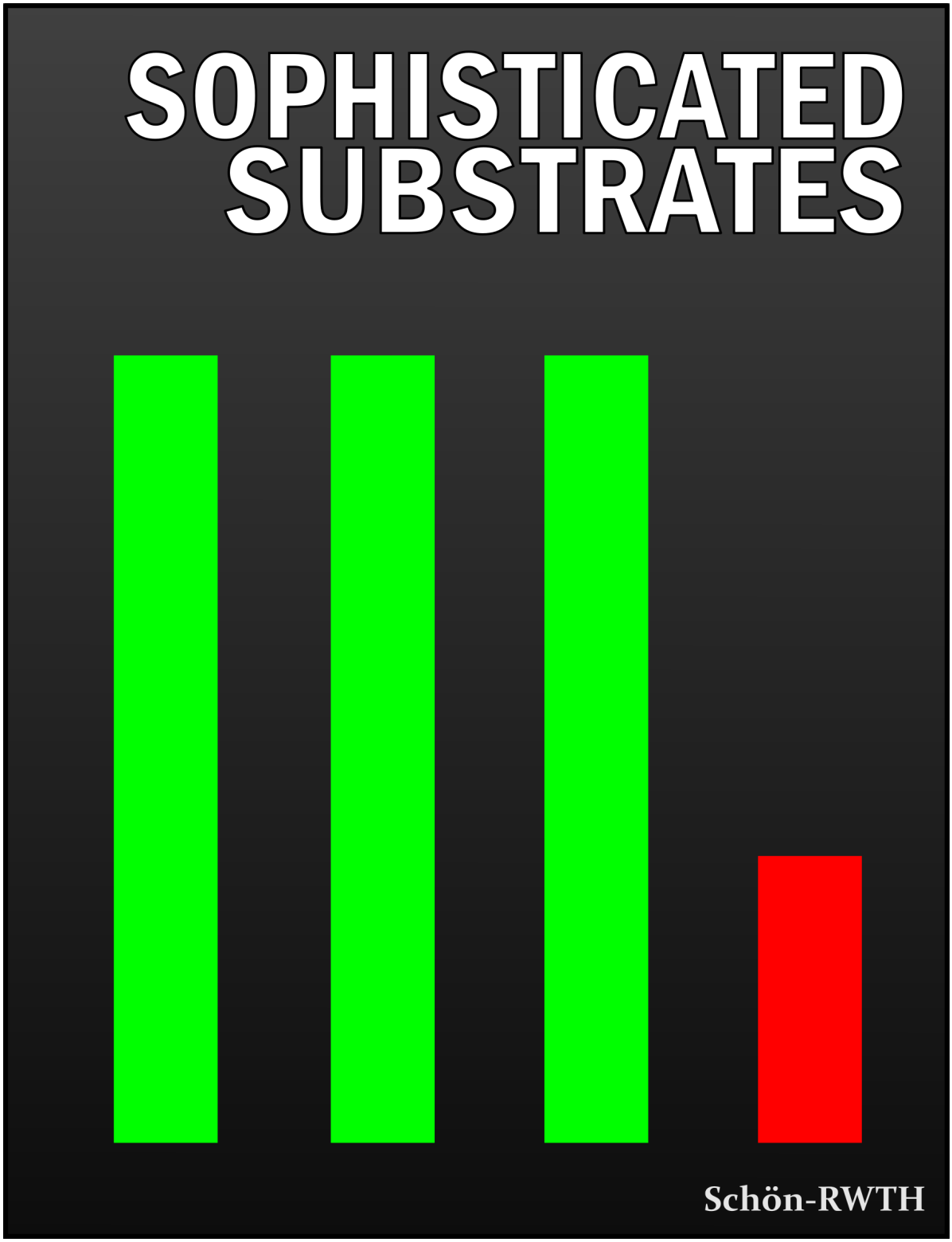


Figure 9.5: Pillar Tendency - Unpublished. This cover has been designed for the publication *Discovering Electron-Transfer-Driven Changes in Chemical Bonding in Lead Chalcogenides* by Stefan Maier et al.[24]. It was inspired by figure 2 of this publication. Name reference to [199].

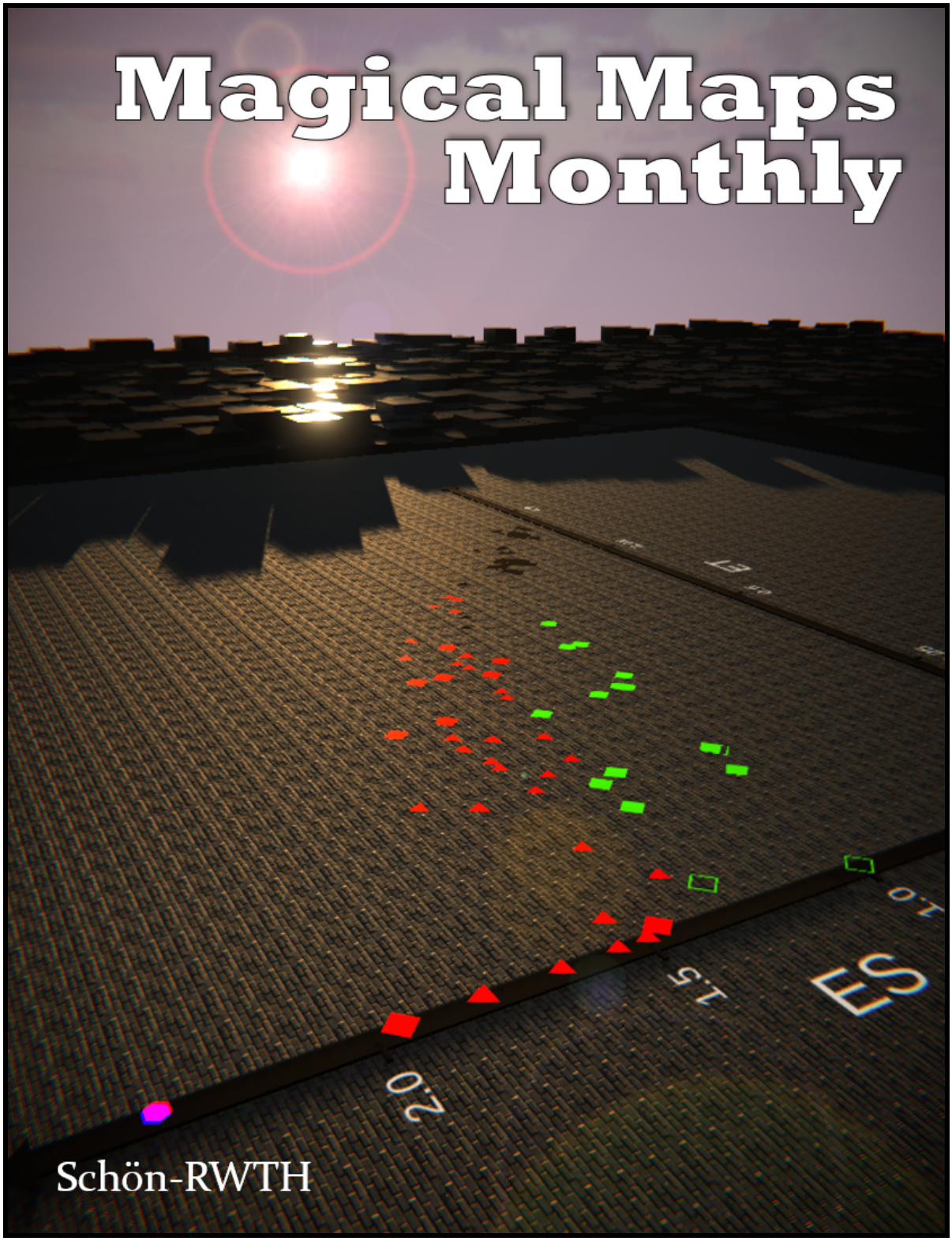


Figure 9.6: Phantom Map - Unpublished. This cover has been designed for the publication *Discovering Electron-Transfer-Driven Changes in Chemical Bonding in Lead Chalcogenides* by Stefan Maier et al.[24]. It was inspired by figure 5 of this publication. Name reference to [200].

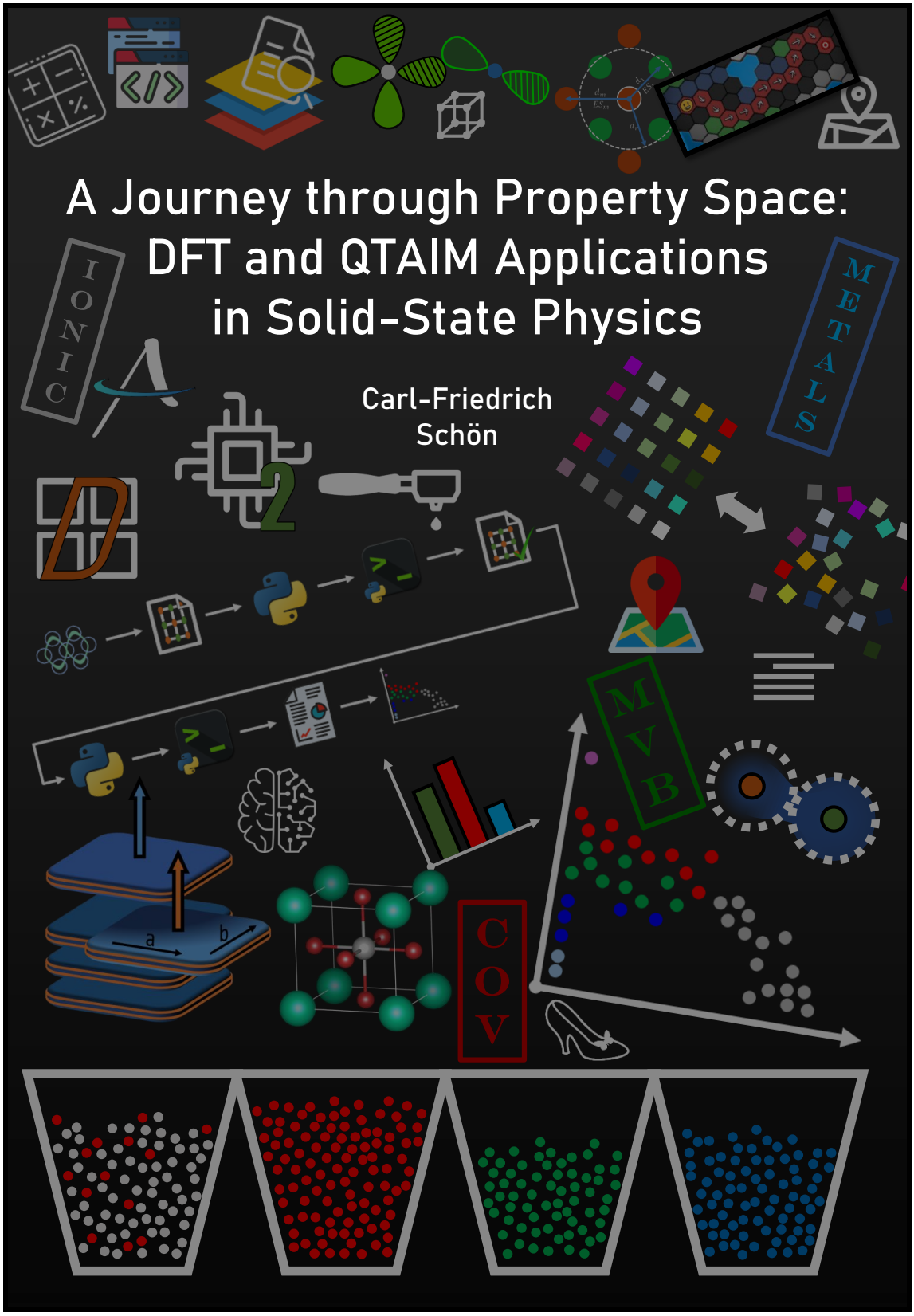


Figure 9.7: Icon Crusaders - Unpublished. Alternative cover for this thesis. Name reference to [201].

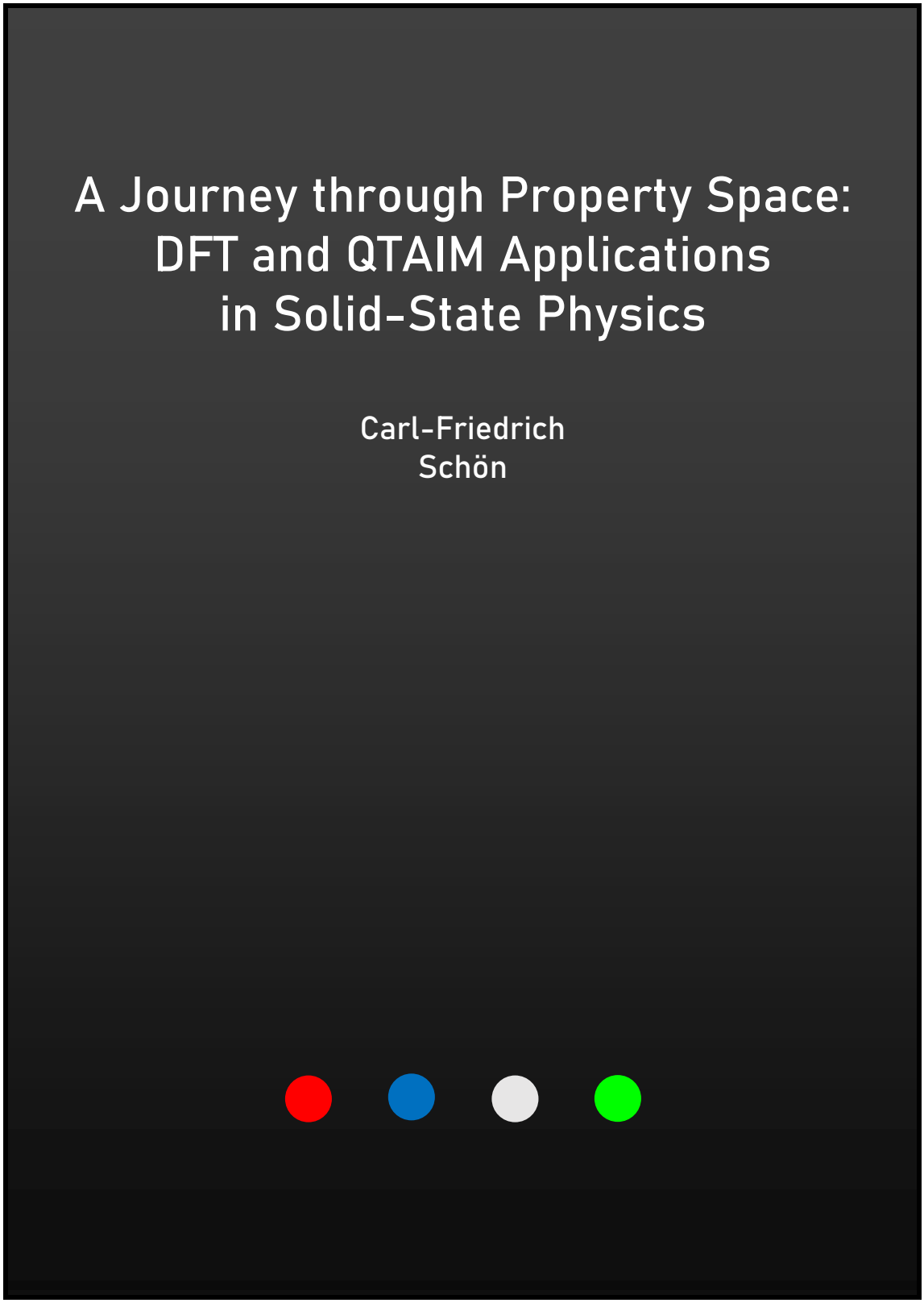


Figure 9.8: All Balls Run - Unpublished. Alternative cover for this thesis. Name reference to [202].

Chapter 10

Interactive Map

The interactive version of the ES/ET map was developed using UNITY3D[203]. It can also be used to view the van Arkel-Ketelaar triangle (see figure 8.1), as well as the Littlewood map (see figure 8.2). At the time of publication, the interactive map is hosted at:

materials-map.rwth-aachen.de

The dataset of compounds can be updated by adapting a *.csv file* and consequently recompiling the software. Potential target platforms are inter alia *Windows*, *MacOS* and *Linux*, as well as *WebGL* for use on a website. The following screenshots depict some of the features of the interactive map.

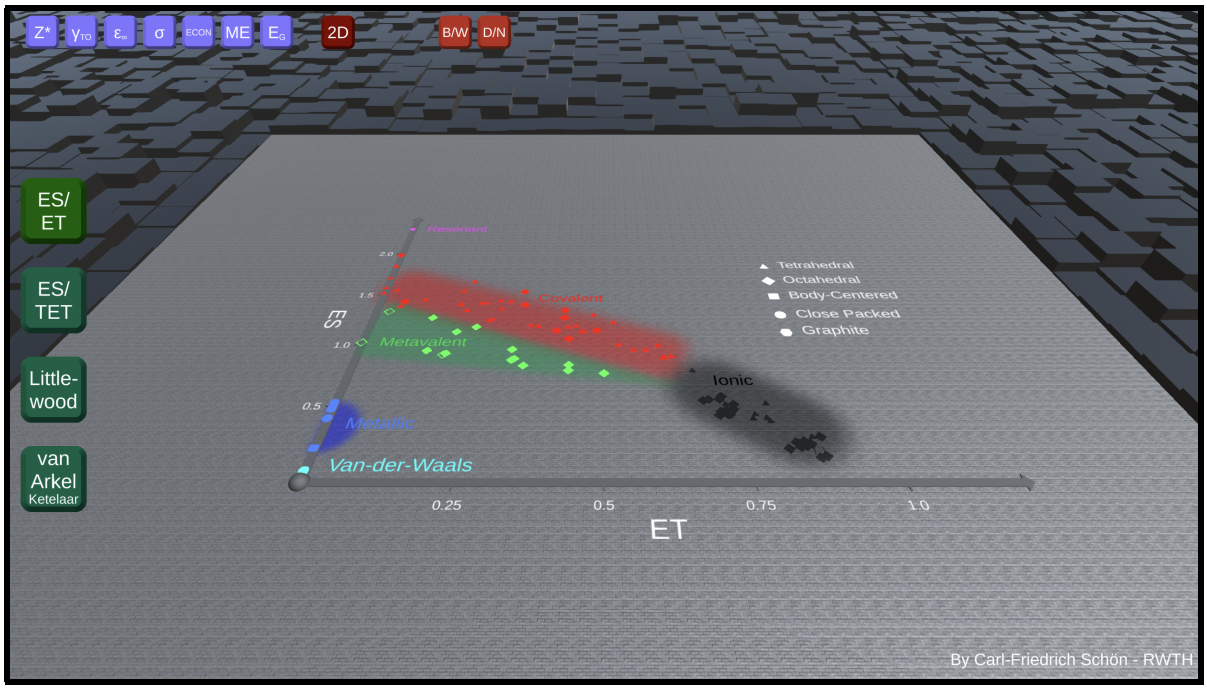


Figure 10.1: Standard view of the interactive map, showing ES and ET.

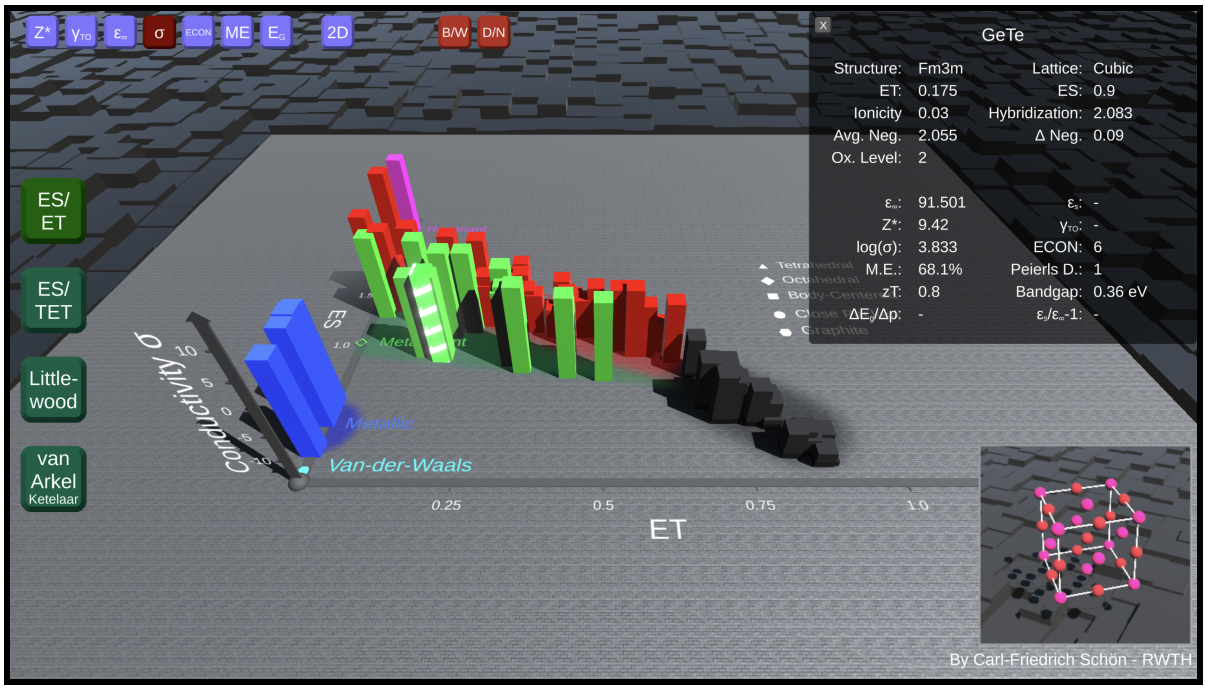


Figure 10.2: By selecting σ from the interface, the conductivity $\log(\sigma)$ is depicted on the z-axis. Additional information regarding a specific compound can be displayed by clicking on it.

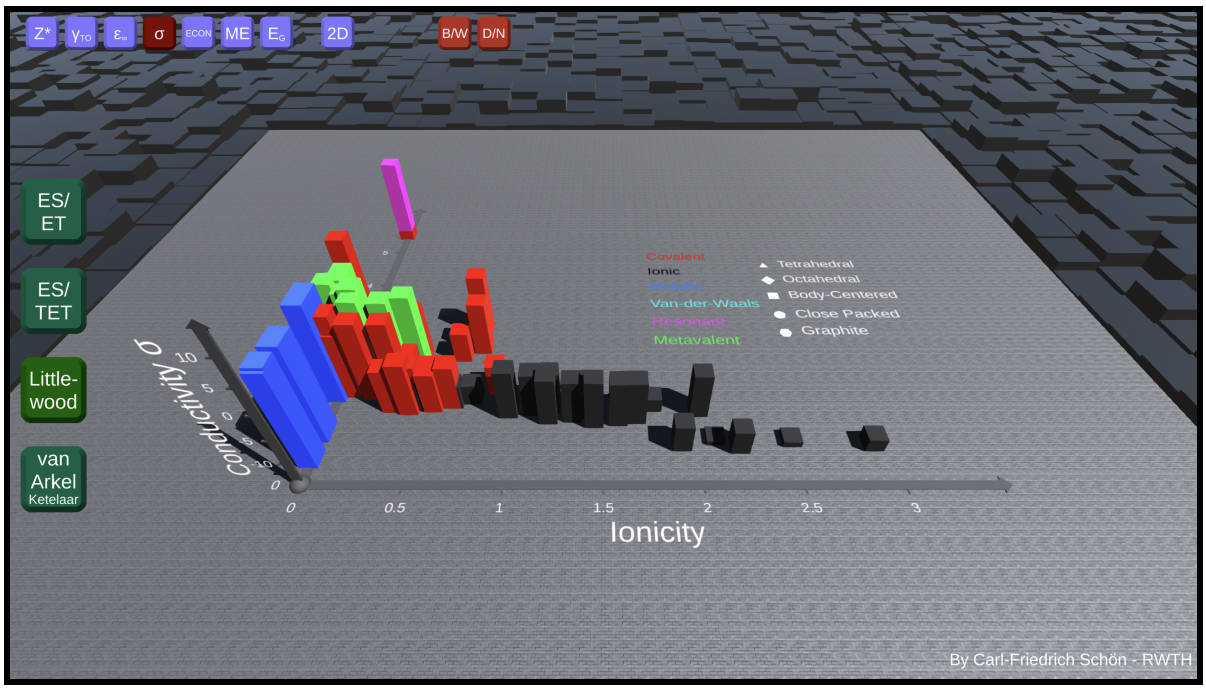


Figure 10.3: By selecting "Littlewood" from the interface, ionicity and hybridization are used as x- and y-axis respectively (see figure 8.2).

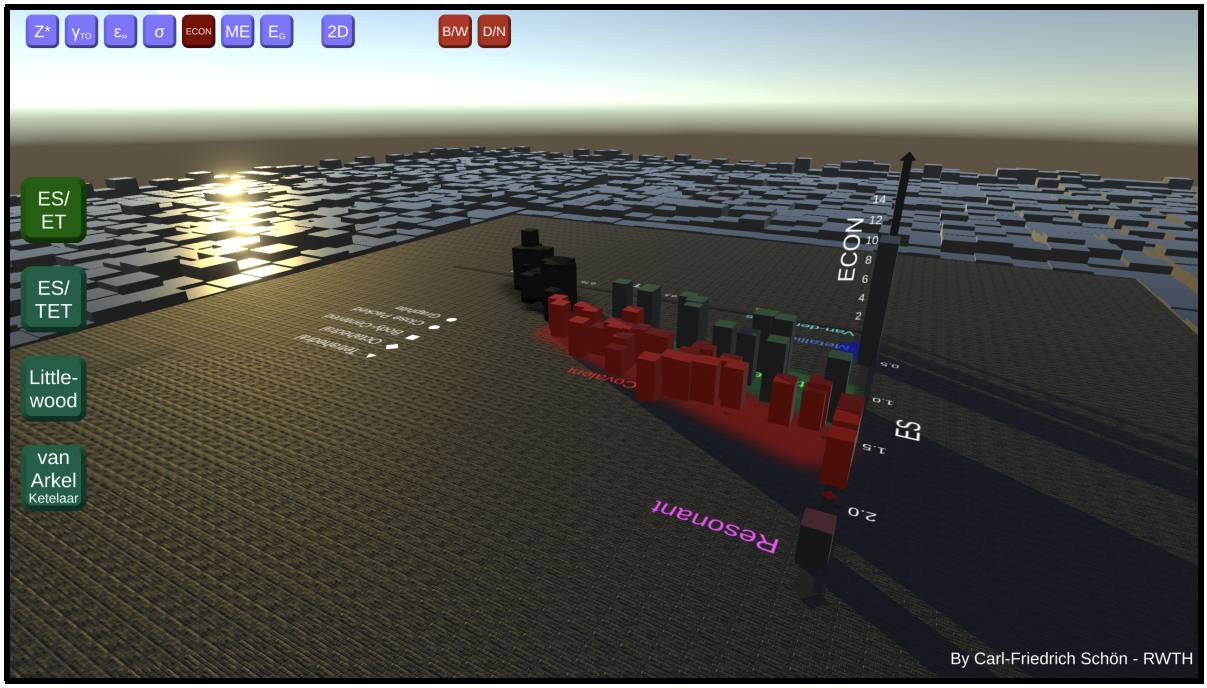


Figure 10.4: By selecting "D/N" (Day/Night), the scene transitions into night-mode and vice versa. An intermediate state is depicted here.

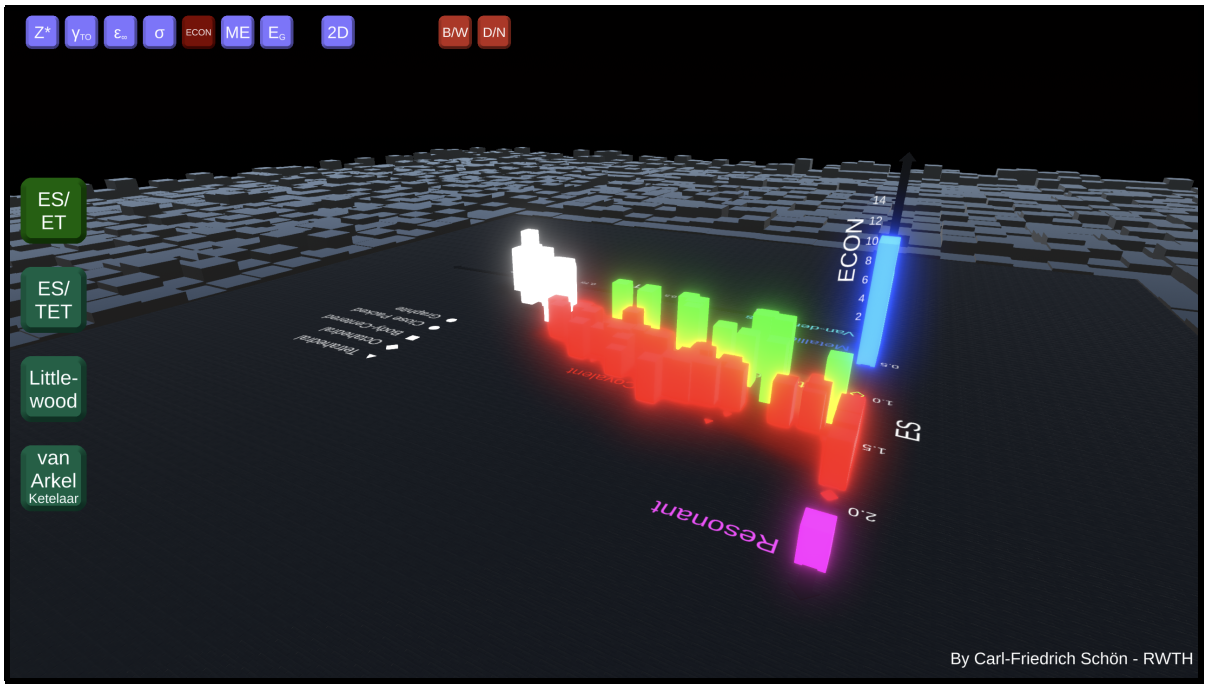


Figure 10.5: By selecting "D/N" (Day/Night), the scene transitions into night-mode and vice versa.

Chapter 11

Additional Tables

Table 11.1: Domain averaged Fermi hole (DAFH) table for cubic PbSe. The delocalization index (DI) denotes the number of shared electron pairs: $\mathrm{ES} = 2 \times \mathrm{DI}$. Adapted from [3, 24].

PbSe cubic	Occupation (e^-)	Localization in Native Basin	Overlap to Partner Basin	DI Contribution
		Lead		
6p (3x)	0.32	17.7%	40.8%	0.263 (Pb-Se)
6s	1.81	90.5%	1.4%	0.050 (Pb-Se)
5d (5x)	2.0	99.4%	-	-
		Selenium		
4p (3x)	1.4	75.6%	8.2%	0.246 (Se-Pb)
4s	1.9	97.0%	0.4%	0.018 (Se-Pb)

Table 11.2: Domain averaged Fermi hole (DAFH) table for cubic PbS. The delocalization index (DI) denotes the number of shared electron pairs: $\mathrm{ES} = 2 \times \mathrm{DI}$. Adapted from [3, 24].

PbS cubic	Occupation (e^-)	Localization in Native Basin	Overlap to Partner Basin	DI Contribution
		Lead		
6p (3x)	0.31	15.8%	20.9%	0.209 (Pb-S)
6s	1.79	89.4%	1.6%	0.056 (Pb-S)
5d (5x)	2.0	99.2%	-	-
		Sulfur		
3p (3x)	1.57	78%	7.3%	0.227 (S-Pb)
3s	1.93	96.1%	0.5%	0.018 (S-Pb)

Table 11.3: Domain averaged Fermi hole (DAFH) table for GeTe with a short to long bond ratio: PD $= \frac{d_{\text{long}}}{d_{\text{short}}} = 1.07$. The delocalization index (DI) denotes the number of shared electron pairs: ES $= 2\times$DI.

GeTe (1.07)	Occupation (e^-)	Localization in Native Basin	Overlap to Partner Basin	DI Contribution
		Germanium		
4p (3x)	0.52	26.6%	39.2% 14.8%	0.41 (Ge-Te) 0.15 (Ge-Te)
4s	1.73	86.3%	2.0% (3x) 1.8% (3x)	0.07 (Ge-Te) 0.06 (Ge-Te)
3d (5x)	2.00	>99.9%	-	-
		Tellurium		
5p (3x)	1.30	65.1%	15.0% 8.1%	0.39 (Te-Ge) 0.21 (Te-Ge)
5s	1.88	93.8%	0.9% (3x) 0.8% (3x)	0.03 (Te-Ge) 0.03 (Te-Ge)

Table 11.4: Domain averaged Fermi hole (DAFH) table for GeTe with a short to long bond ratio: PD $= \frac{d_{\text{long}}}{d_{\text{short}}} = 1.12$. The delocalization index (DI) denotes the number of shared electron pairs: ES $= 2\times$DI.

GeTe (1.12)	Occupation (e^-)	Localization in Native Basin	Overlap to Partner Basin	DI Contribution
		Germanium		
4p (3x)	0.53	27.1%	43.8% 10.6%	0.46 (Ge-Te) 0.11 (Ge-Te)
4s	1.73	86.5%	2.1% (3x) 1.6% (3x)	0.07 (Ge-Te) 0.06 (Ge-Te)
3d (5x)	2.00	>99.9%	-	-
		Tellurium		
5p (3x)	1.30	65.2%	17.1% 6.1%	0.45 (Te-Ge) 0.16 (Te-Ge)
5s	1.88	93.9%	0.9% (3x) 0.8% (3x)	0.03 (Te-Ge) 0.03 (Te-Ge)

Table 11.5: Domain averaged Fermi hole (DAFH) table for GeTe with a short to long bond ratio: $\mathrm{PD} = \frac{d_{\mathrm{long}}}{d_{\mathrm{short}}} = 1.19$. The delocalization index (DI) denotes the number of shared electron pairs: $\mathrm{ES} = 2{\times}\mathrm{DI}$.

GeTe (1.19)	Occupation (e^-)	Localization in Native Basin	Overlap to Partner Basin	DI Contribution
		Germanium		
4p (3x)	0.54	27.6%	46.9% 7.7%	0.51 (Ge-Te) 0.08 (Ge-Te)
4s	1.74	86.8%	2.2% (3x) 1.4% (3x)	0.08 (Ge-Te) 0.05 (Ge-Te)
3d (5x)	2.00	>99.9%	-	-
		Tellurium		
5p (3x)	1.30	65.2%	18.8% 4.6%	0.49 (Te-Ge) 0.12 (Te-Ge)
5s	1.88	94.1%	0.9% (3x) 0.7% (3x)	0.03 (Te-Ge) 0.03 (Te-Ge)

Table 11.6: Domain averaged Fermi hole (DAFH) table for GeTe with a short to long bond ratio: $\mathrm{PD} = \frac{d_{\mathrm{long}}}{d_{\mathrm{short}}} = 1.26$. The delocalization index (DI) denotes the number of shared electron pairs: $\mathrm{ES} = 2{\times}\mathrm{DI}$.

GeTe (1.26)	Occupation (e^-)	Localization in Native Basin	Overlap to Partner Basin	DI Contribution
		Germanium		
4p (3x)	0.55	27.9%	48.8% 5.8%	0.53 (Ge-Te) 0.06 (Ge-Te)
4s	1.74	87.3%	2.4% (3x) 1.2% (3x)	0.08 (Ge-Te) 0.04 (Ge-Te)
3d (5x)	2.00	>99.9%	-	-
		Tellurium		
5p (3x)	1.31	65.6%	19.7% 1.7%	0.51 (Te-Ge) 0.04 (Te-Ge)
5s	1.89	94.4%	1.0% (3x) 0.6% (3x)	0.04 (Te-Ge) 0.02 (Te-Ge)

Table 11.7: Domain averaged Fermi hole (DAFH) table for Al. The delocalization index (DI) denotes the number of shared electron pairs: $\mathrm{ES} = 2 \times \mathrm{DI}$. Adapted from [3].

Al	Occupation (e^-)	Localization in Native Basin	Overlap to Partner Basin	DI Contribution
3p (3x)	0.47	25.1%	5.6%-7.5% (4x) 1.4%-2.2% (4x) 11.3%-13.4% (2x)	0.05-0.07 (Al-Al) 0.01-0.02 (Al-Al) 0.11-0.13 (Al-Al)
3s	1.11	57.2%	2.6%-3.4% (12x)	0.06-0.07 (Al-Al)

Table 11.8: Domain averaged Fermi hole (DAFH) table for diamond. The delocalization index (DI) denotes the number of shared electron pairs: $\mathrm{ES} = 2 \times \mathrm{DI}$. Adapted from [3].

C Diamond	Occupation (e^-)	Localization in Native Basin	Overlap to Partner Basin	DI Contribution
$3sp^3$ (4x)	0.95	48.7%	42.2%	0.80 (C-C)

Table 11.9: Domain averaged Fermi hole (DAFH) table for NaCl. The delocalization index (DI) denotes the number of shared electron pairs: $\mathrm{ES} = 2 \times \mathrm{DI}$. Adapted from [3].

NaCl	Occupation (e^-)	Localization in Native Basin	Overlap to Partner Basin	DI Contribution
		Sodium		
3p (3x)	1.99	99.4%	-	-
3s	2.0	99.9%	-	-
		Chlorine		
3p (3x)	1.90	95.2%	1.1% (2x)	0.04 (Cl-Na)
3s	1.99	99.3%	-	-

Table 11.10: Domain averaged Fermi hole (DAFH) table for $CsPbI_3$. The delocalization index (DI) denotes the number of shared electron pairs: $\mathrm{ES} = 2 \times \mathrm{DI}$. Adapted from [47].

$CsPbI_3$	**Occupation (e^-)**	**Localization in Native Basin**	**Overlap to Partner Basin**	**DI Contribution**
		Lead		
6p (3x)	0.35	17.3%	33.7%	0.232 (Pb-I)
6s	1.81	90.4%	1.4%	0.052 (Pb-I)
		Iodine		
5p	1.60	79.8%	7.9%	0.253 (I-Pb)
5p (2x) Lone Pair	1.88	93.9%	0.8%	0.031 (I-Pb)

Table 11.11: Domain averaged Fermi hole (DAFH) table for $CsPbBr_3$. The delocalization index (DI) denotes the number of shared electron pairs: $\mathrm{ES} = 2 \times \mathrm{DI}$. Adapted from [47].

$CsPbBr_3$	**Occupation (e^-)**	**Localization in Native Basin**	**Overlap to Partner Basin**	**DI Contribution**
		Lead		
6p (3x)	0.30	14.8%	35.1%	0.208 (Pb-Br)
6s	1.80	89.9%	1.5%	0.054 (Pb-Br)
		Bromine		
4p	1.66	82.9%	6.9%	0.229 (Br-Pb)
4p (2x) Lone Pair	1.89	94.6%	0.8%	0.032 (Br-Pb)

Table 11.12: Domain averaged Fermi hole (DAFH) table for $CsPbCl_3$. The delocalization index (DI) denotes the number of shared electron pairs: $ES = 2 \times DI$. Adapted from [47].

$CsPbCl_3$	Occupation (e^-)	Localization in Native Basin	Overlap to Partner Basin	DI Contribution
		Lead		
6p (3x)	0.31	11.8%	32.1%	0.201 (Pb-Cl)
6s	1.80	90.1%	1.5%	0.054 (Pb-Cl)
		Chlorine		
3p	1.70	84.9%	6.3%	0.213 (Cl-Pb)
3p (2x) Lone Pair	1.91	95.2%	0.8%	0.032 (Cl-Pb)

Table 11.13: Domain averaged Fermi hole (DAFH) table for $BaTiO_3$. The delocalization index (DI) denotes the number of shared electron pairs: $ES = 2 \times DI$. Adapted from [47].

$BaTiO_3$	Occupation (e^-)	Localization in Native Basin	Overlap to Partner Basin	DI Contribution
		Titanium		
3d (3x)	0.24	11.9%	19.0%	0.091 (Ti-O)
Mixed Type (6x)	0.18	14.1%	66.8%	0.240 (Ti-O)
		Oxygen		
2p	1.62	81.2%	7.6%	0.246 (O-Ti)
2p (2x) Lone Pair	1.70	85.2%	2.9%	0.099 (O-Ti)

Table 11.14: Compounds used for the analysis of layeredness (Part I). Compiled by Alexander Kiehn[68].

Name	ID	Space Group	d_{2D}	Interlayer ES	ΔE_{Shift}	LEA Class.	Bonding Type
C	sd-1925964	194	1.63	0.04	28.26	True vdW	Resonant
P	sd-1251849	64	0.98	0.15		True vdW	Covalent
BN	sd-0457583	194	1.59	0.01	35.37	True vdW	Ionic
As_2Te_3	sd-0455575	166	0.28	0.46	730.64	Pseudo-vdW	Metavalent
Sb_2Te_3	sd-0250935	166	0.29	0.37	696.35	Pseudo-vdW	Metavalent
Bi_2Se_3	sd-0305047	166	0.25	0.25	486.43	Pseudo-vdW	Metavalent
Bi_2Te_3	sd-0310425	166	0.28	0.35	722.07	Pseudo-vdW	Metavalent
Ga_2S_3 (HP)	sd-1831448	166	0.15	0.47	902.86	Pseudo-vdW	Not assigned
In_2Te_3 (HP)	sd-0534031	166	0.19		1456.82	Pseudo-vdW	Not assigned
SnS_2 (1T)	sd-1503981	164	0.75	0.13	134.07	True vdW	Covalent
$SnSe_2$ (1T)	sd-0379304	164	0.79	0.15	149.750	True vdW	Covalent
PbS_2 (1T)	sd-1903552	164	0.72	0.13	115.34	Pseudo-vdW	Not assigned
GaS	sd-0310970	194	1.15	0.10	73.90	True vdW	Covalent
GaSe	sd-0454324	194	1.18	0.12	96.143	True vdW	Covalent
GaTe	sd-0457825	194	1.27	0.16	134.84	True vdW	Covalent
InS	sd-1301541	58	-0.04	1.14		Not Lay.	Covalent
InSe	sd-0379302	194	1.26	0.09	56.49	True vdW	Covalent
TiS_2 (1T)	sd-0313908	164	0.73	0.16	97.64	True vdW	Not assigned
$TiSe_2$ (1T)	sd-0381309	164	0.73	0.18	107.73	True vdW	Not assigned
$TiTe_2$ (1T)	sd-0455326	164	0.90	0.22	127.64	True vdW	Not assigned
VS_2 (1T)	sd-0380499	164	0.77	0.15	83.66	True vdW	Not assigned
VSe_2 (1T)	sd-0452492	164	0.85	0.16	87.44	True vdW	Not assigned
VTe_2 (1T')	sd-1501087	12	0.35			Pseudo-vdW	Not assigned
VTe_2 (1T, HT)	sd-1501085	164	0.92	0.21	128.60	True vdW	Not assigned
$CrSe_2$ (1T)	sd-0456707	164	0.79	0.17	121.20	True vdW	Not assigned
$NiTe_2$ (1T)	sd-1404525	164	0.56	0.33	382.25	Pseudo-vdW	Not assigned
$CuTe_2$ (1T)	sd-1404528	164	0.61	0.30	69.93	Pseudo-vdW	Not assigned
$ZnTe_2$ (1T)	sd-1404529	164	0.54	0.37	161.55	Pseudo-vdW	Not assigned
ZrS_2 (1T)	sd-0260479	164	0.80	0.10	58.77	True vdW	Not assigned
$ZrSe_2$ (1T)	sd-0260478	164	0.73	0.18	152.66	Pseudo-vdW	Not assigned
$ZrTe_2$ (1T, HT)	sd-0455325	164	0.90	0.19	131.276	True vdW	Not assigned
NbS_2 (2H, HT)	sd-0313921	194	0.82	0.13	48.13	True vdW	Not assigned
NbS_2 (3R)	sd-0525227	160	0.73		60.72	True vdW	Not assigned
$NbSe_2$ (2H)	sd-0310422	194	0.93	0.13	37.01	True vdW	Not assigned
$NbSe_2$ (3R)	sd-0528384	160	0.62		94.29	Pseudo-vdW	Not assigned
$NbTe_2$ (1T)	sd-0531421	164	0.97	0.16	106.10	True vdW	Not assigned
$NbTe_2$ (1T')	sd-1947005	12	0.37			Pseudo-vdW	Not assigned
MoS_2 (2H)	sd-0309036	194	0.77	0.13	94.10	True vdW	Not assigned
MoS_2 (3R)	sd-0526530	160	0.78		105.02	True vdW	Not assigned
$MoSe_2$ (2H)	sd-0310431	194	0.95	0.12	78.87	True vdW	Not assigned
$MoSe_2$ (3R, HP)	sd-0457993	160	0.96		85.31	True vdW	Not assigned
$MoTe_2$ (1T')	sd-1044911	12	0.75	0.22		True vdW	Not assigned
$MoTe_2$ (2H)	sd-0453840	194	1.06	0.15	110.18	True vdW	Not assigned
$RhTe_2$ (1T)	sd-1404523	164	0.59	0.20	416.05	Pseudo-vdW	Not assigned
$PdTe_2$ (1T)	sd-0260989	164	0.51	0.28	558.57	Pseudo-vdW	Not assigned
HfS_2 (1T)	sd-0544648	164	0.75	0.13	90.19	True vdW	Not assigned
$HfSe_2$ (1T)	sd-0453169	164	0.81	0.14	103.22	True vdW	Not assigned
$HfTe_2$ (1T)	sd-0453168	164	0.91	0.18	127.19	True vdW	Not assigned
TaS_2 (1T)	sd-0382778	164	0.79	0.13	99.66	True vdW	Not assigned
TaS_2 (2H)	sd-0457067	194	0.84	0.12	70.42	True vdW	Not assigned
TaS_2 (3R)	sd-1937509	160	0.76		76.79	True vdW	Not assigned

Table 11.15: Compounds used for the analysis of layeredness (Part II). Compiled by Alexander Kiehn[68].

Name	ID	Space Group	d_{2D}	Interlayer ES	ΔE_{Shift}	LEA Class.	Bonding Type
$TaSe_2$ (1T)	sd-0455259	164	0.86	0.13	83.65	True vdW	Not assigned
$TaSe_2$ (2H)	sd-0525233	194	0.90	0.13	60.75	True vdW	Not assigned
$TaSe_2$ (3R)	sd-0453300	160	0.80		89.93	True vdW	Not assigned
$TaTe_2$ (1T')	sd-0454049	12	0.38			Pseudo-vdW	Not assigned
WS_2 (2H)	sd-0455614	194	0.82	0.12	86.99	True vdW	Not assigned
WS_2 (3R)	sd-0455616	160	0.85		90.77	True vdW	Not assigned
WSe_2 (2H)	sd-0310430	194	0.96	0.11	66.16	True vdW	Not assigned
WTe_2 (1T')	sd-0453184	31	0.81			True vdW	Not assigned
WTe_2 (1T, Rotated)	sd-0531417	186	1.11		72.40	True vdW	Not assigned
$IrTe_2$ (1T)	sd-1404524	164	0.59	0.15	508.25	Pseudo-vdW	Not assigned
PtS_2 (1T)	sd-0455705	164	0.54	0.12	405.12	Pseudo-vdW	Not assigned
$PtSe_2$ (1T)	sd-0455703	164	0.53	0.15	490.83	Pseudo-vdW	Not assigned
$PtTe_2$ (1T)	sd-0455704	164	0.53	0.20	667.34	Pseudo-vdW	Not assigned
$AuTe_2$ (1T')	sd-1243537	12	-0.06			Pseudo-vdW	Not assigned
$AuTe_2$ (1T, HP)	sd-1215131	164	0.08	0.77		Pseudo-vdW	Not assigned
GeS	sd-0300094	62	0.53	0.24	271.50	Pseudo-vdW	Covalent
GeSe	sd-0458320	62	0.55	0.25	182.30	Pseudo-vdW	Covalent
GeTe	sd-0551187	62	0.65	0.26	18.75	Pseudo-vdW	Covalent
$GeTe^{cub}$	mp-2612	225	-0.3	0.9	139.66	Not Lay.	Metavalent
SnS	sd-0309394	62	0.58	0.25	397.40	Pseudo-vdW	Covalent
SnSe	sd-0380080	62	0.58	0.26	290.20	Pseudo-vdW	Covalent
SnTe	sd-0457338	62	0.59	0.30	176.58	Pseudo-vdW	Not assigned
PbSe	sd-0457339	62	0.63	0.25		Pseudo-vdW	Not assigned
PbTe. hp	sd-0457080	62	0.60	0.29		Pseudo-vdW	Not assigned
$Ge_2Sb_2Te_5$ - GST225	sd-1102022	164	0.32	0.42	329.60	Pseudo-vdW	Metavalent
MgI_2 (1T)	icsd-281551	164	1.05	0.13		True vdW	Not assigned
MoS_2 (1T)	icsd-254956	164	0.85	0.13		True vdW	Not assigned
NbS_2 (1T)	icsd-250594	164	0.80	0.12		True vdW	Not assigned
$ZrTe_2$ (1T)	icsd-7956	164	0.79	0.21	143.43	Pseudo-vdW	Not assigned
In_2Se_3	icsd-602266	164	0.69	0.21	47.79	True vdW	Covalent
$Fe3GeTe_2$	sd-1420956	194	0.81		323.23	True vdW	Not assigned
$Pt3Te_4$	sd-1321040	166	0.52		758.33	Pseudo-vdW	Not assigned
Ag	icsd-181730	225	-0.30	0.49	310.46	Not Lay.	Metallic
Al	icsd-18839	225	-0.30	0.5	279.84	Not Lay.	Metallic
Ca	sd-0251474	225	-0.30	0.41	169.57	Not Lay.	Metallic
Si	icsd-51688	227	-0.30	1.68	10496.22	Not Lay.	Covalent
Ge	icsd-121532	227	-0.30	1.57	8978.99	Not Lay.	Covalent
GaAs	icsd-107946	216	-0.30	1.40	10938.07	Not Lay.	Covalent
NaCl	icsd-29929	225	-0.30	0.14	411.05	Not Lay.	Ionic
MgO	icsd-52026	225	-0.30	0.24	1587.47	Not Lay.	Ionic
BeO	icsd-391224	186	-0.20	0.28	1141.17	Not Lay.	Ionic
CaS	sd-0300100	225	-0.30	0.43	995.54	Not Lay.	Ionic
KF	sd-0556890	225	-0.30	0.23	364.14	Not Lay.	Ionic
Mg	sd-0260118	225	-0.30	0.40	406.53	Not Lay.	Metallic
PbS	sd-0309096	225	-0.30	0.74	383.02	Not Lay.	Metavalent
PbSe	sd-0310439	225	-0.30	0.76	290.50	Not Lay.	Metavalent
PbTe	sd-0310485	225	-0.30	0.80	195.20	Not Lay.	Metavalent
SnTe	sd-0260901	225	-0.30	0.80	136.85	Not Lay.	Metavalent
Ti	icsd-253841	225	-0.30		542.69	Not Lay.	Metallic
Co	sd-0261742	225	-0.30		318.39	Not Lay.	Metallic
Sr	sd-0251473	225	-0.30		164.10	Not Lay.	Metallic
K	sd-1214736	225	-0.30		148.04	Not Lay.	Metallic

Table 11.16: Chalcogenide minerals, as compiled by Sebastian Gruner, Christian Stenz and Alexander Kiehn, RWTH Aachen University[68].

Name	Chem. Comp.	LEA Class.	Exp. Class.	Name	Chem. Comp.	LEA Class.	Exp. Class.
Covellin	CuS	Not Lay.	Layered	Volynskit	$AgBiTe_2$	Not Lay.	Not Lay.
Ikunolith	$Bi_3(S,Se)_3$	Not Lay.	Layered	Wassonit	TiS	Not Lay.	Not Lay.
Klockmannit	CuSe	Not Lay.	Layered	Wurtzit (2H)	(Zn,Fe)S	Not Lay.	Not Lay.
Kochkarit	$PbBi_3Te_7$	Not Lay.	Layered	Zlatogorit	$CuNiSb_2$	Not Lay.	Not Lay.
Pilsenit	Bi_3Te_3	Not Lay.	Layered	Pararealgar	As_3S_3	Not Lay.	Not Lay.
Achavalit	(Fe,Cu)Se	Not Lay.	Not Lay.	Petricekit	$CuSe_2$	Not Lay.	Not Lay.
$AgSnTe_2$	$AgSnTe_2$	Not Lay.	Not Lay.	Roterbärit	$PdBiCuSe_3$	Not Lay.	Not Lay.
Aikinit	$PbCuBiSe_3$	Not Lay.	Not Lay.	Vaesit	NiS_2	Not Lay.	Not Lay.
Alabandin	MnS	Not Lay.	Not Lay.	Anorpiment	As_2S_3	True vdW	Layered
Altait	PbTe	Not Lay.	Not Lay.	Antimon	Sb	Pseudo-vdW	Layered
Berzelianit	Cu_2Se	Not Lay.	Not Lay.	Arsen	As	Pseudo-vdW	Layered
Bohdanowiczit	$AgBiSe_2$	Not Lay.	Not Lay.	Arsenolamprit	As	True vdW	Layered
Bowieit	$(Rh,Ir,Pt)_2S_3$	Not Lay.	Not Lay.	As_2Te_3	As_2Te_3	Pseudo-vdW	Layered
Breithauptit	NiSb	Not Lay.	Not Lay.	Orpiment	As_2S_3	True vdW	Layered
Browneit	MnS	Not Lay.	Not Lay.	Berndtit (2T)	SnS_2	True vdW	Layered
Cadmoselit	CdSe	Not Lay.	Not Lay.	Bi_3Se_3 (synth.)	Bi_3Se_3	Pseudo-vdW	Layered
Caswellsilverit	$NaCrS_2$	Not Lay.	Not Lay.	Calaverit	$AuTe_2$	Pseudo-vdW	Layered
Cattierit	CoS_2	Not Lay.	Not Lay.	Chalkostibit	$CuSbS_2$	Pseudo-vdW	Layered
Cinnabarit	HgS	Not Lay.	Not Lay.	Demicheleite-(Cl)	BiSCl	Pseudo-vdW	Layered
Clausthalit	PbS	Not Lay.	Not Lay.	Emplektit	$CuBiS_2$	Pseudo-vdW	Layered
Cooperit	PtS	Not Lay.	Not Lay.	Graphen (synth.)	C	True vdW	Layered
Cuboargyrit	$AgSbS_2$	Not Lay.	Not Lay.	Herzenbergit	SnS	Pseudo-vdW	Layered
Empressit	AgTe	Not Lay.	Not Lay.	In_2Se_3	In_2Se_3	True vdW	Layered
Erlichmannit	OsS_2	Not Lay.	Not Lay.	Jacutingait	Pt_2HgSe_3	Pseudo-vdW	Layered
Ernigglit	$Tl_2SnAs_2S_6$	Not Lay.	Not Lay.	Kawazulite	$Bi_2(Te,Se,S)_3$	True vdW	Layered
Ferroselit	$FeSe_2$	Not Lay.	Not Lay.	Laphamit	$As_2(Se,S)_3$	True vdW	Layered
Freboldit	CoSe	Not Lay.	Not Lay.	Mackinawit	$(Fe,Ni)S_{0.9}$	True vdW	Layered
Frohbergit	$FeTe_2$	Not Lay.	Not Lay.	Melonit	$NiTe_2$	Pseudo-vdW	Layered
Galenit	PbS	Not Lay.	Not Lay.	Merenskyit	$(Pd,Pt)(Te,Bi)_2$	Pseudo-vdW	Layered
Greenockit	CdS	Not Lay.	Not Lay.	Mitrofanovite	Pt_3Te_3	Pseudo-vdW	Layered
Hauerit	MnS_2	Not Lay.	Not Lay.	Molybdänit (2H)	MoS_2	True vdW	Layered
Hawleyit	CdS	Not Lay.	Not Lay.	Molybdänit (2H, 1)	MoS_2	True vdW	Layered
In_2Se_3	In_2Se_3	Not Lay.	Not Lay.	Molybdänit (2H, 2)	MoS_2	True vdW	Layered
Jaipurit	CoS	Not Lay.	Not Lay.	Molybdänit (2H, 3)	MoS_2	True vdW	Layered
Kotulskit	Pd(Te,Bi)	Not Lay.	Not Lay.	Molybdänit (3R)	MoS_2	True vdW	Layered
Krutait	$CuSe_2$	Not Lay.	Not Lay.	MoS_2 (3R, synth. 2)	MoS_2	True vdW	Layered
Laurit	RuS_2	Not Lay.	Not Lay.	Drysdallit (synth. 1)	$MoSe_2$	True vdW	Layered
Markasit	FeS_2	Not Lay.	Not Lay.	Drysdallit (synth. 2)	$MoSe_2$	True vdW	Layered
Matildit	$AgBiS_2$	Not Lay.	Not Lay.	Paraguanajuatit	$Bi_2(Se,S)_3$	Pseudo-vdW	Layered
Mattagamit	$CoTe_2$	Not Lay.	Not Lay.	Shuangfengit	$IrTe_2$	Pseudo-vdW	Layered
Metacinnabarit	HgS	Not Lay.	Not Lay.	Skippenit (1963)	$Bi_2Se_2(Te,S)$	True vdW	Layered
Millerit	NiS	Not Lay.	Not Lay.	Skippenit (2003)	$Bi_2Se_2(Te,S)$	True vdW	Layered
Nickelin	NiAs	Not Lay.	Not Lay.	Sudovikovit	$PtSe_2$	True vdW	Layered
Pyrit	FeS_2	Not Lay.	Not Lay.	Tellurantimon	Sb_2Te_3	True vdW	Layered
Rambergit	MnS	Not Lay.	Not Lay.	Tellurobismuthit I	Bi_2Te_3	True vdW	Layered
Roquesit	$CuInS_2$	Not Lay.	Not Lay.	Tellurobismuthit II	Bi_2Te_3	True vdW	Layered
Schapbachit	$AgBiS_2$	Not Lay.	Not Lay.	Tetradymit	Bi_2Te_2S	True vdW	Layered
Sederholmit	NiSe	Not Lay.	Not Lay.	Tsumoit	BiTe	True vdW	Layered
Sobolevskit	PdBi	Not Lay.	Not Lay.	Tungstenit (2H)	WS_2	True vdW	Layered
Sphalerit	(Zn,Fe)S	Not Lay.	Not Lay.	Tungstenit (3R)	WS_2	True vdW	Layered
Stibarsen	AsSb	Not Lay.	Not Lay.	Vavrinit	Ni_2SbTe_2	True vdW	Layered
Stilleit	ZnSe	Not Lay.	Not Lay.	Vulcanit	CuTe	True vdW	Layered
Stistait	SnSb	Not Lay.	Not Lay.	Bi	Bi	Pseudo-vdW	Layered
Stumpflit	Pt(Sb,Bi)	Not Lay.	Not Lay.	Ingodit	Bi_2TeS	True vdW	Layered
Sudburyit	(Pd,Ni)Sb	Not Lay.	Not Lay.	Nevskit	Bi(Se,S)	True vdW	Layered
Trogtalit	$CoSe_2$	Not Lay.	Not Lay.	Telluronevskit	Bi_3TeSe_2	True vdW	Layered
Troilit	FeS	Not Lay.	Not Lay.				

Table 11.17: Compounds and respective properties used for classification (Part 1). Melting temperature T_{m} in (K), conductivity σ in (S/cm), atomic density ρ_{A} in (10^{-2}/Å^3), density ρ in (g/cm^3), band gap E_{G} in (eV). Properties mostly obtained from the literature[204].

Compound	Structure	Exp.	EMA	T_{m}	σ	Z^*	Z^*_+	ECoN	γ_{TO}	ρ_{A}	ρ	E_{G}	ϵ_∞	t-SNE X	t-SNE Y	TET	ET	ES
Ag	$Fm3m$	MB	MB	1235	6.80E+05	0.00	0.00	12.02	-	0.0587	9.94	0.00	-	19.93	5.51	0.00	0.00	0.49
$AgSbS_2$	$C2/c$	CB	IB	794	9.00E-10	2.08	1.04	3.10	-	0.0404	4.93	2.06	9.60	-5.08	-3.10	0.74	0.37	1.26
$AgSbSe_2$	$R\bar{3}m$	MVB	MVB	910	2.50E+01	3.99	2.00	6.03	-	0.0413	6.64	0.70	28.03	-0.99	-21.84	0.63	0.32	0.85
$AgBiSe_2$	$R3m$	MVB	MVB	1038	1.80E+02	3.97	1.99	6.03	6.63	0.0395	7.79	0.52	24.25	-1.07	-22.21	0.64	0.32	0.86
$AgBiTe_2$	$R3m$	MVB	MVB	828	1.30E+03	4.25	2.12	6.01	6.77	0.0333	7.92	0.16	57.55	-1.94	-21.05	0.36	0.18	0.91
$AgSbTe_2$	$R3m$	MVB	MVB	847	3.25E+02	4.30	2.15	6.02	3.03	0.0347	6.98	0.35	39.21	-2.18	-21.96	0.28	0.14	0.93
$AgSnSe_2$	$R3m$	MB	MB	900	1.00E+04	0.00	0.00	6.03	-	0.0423	6.75	0.00	-	8.77	-3.36	0.69	0.35	0.80
$AgSnTe_2$	$P4/mmm$	MB	MB	818	1.90E+03	0.00	0.00	6.02	-	0.0349	6.98	0.00	-	8.72	-3.39	0.14	0.22	0.80
Al	$Fm3m$	MB	MB	933	4.10E+05	0.00	0.00	12.02	-	0.0620	2.70	0.00	-	17.20	1.60	0.00	0.00	0.51
AlBi	$F43m$	CB	CB	923	7.57E-02	1.87	0.62	4.00	1.53	0.0298	5.84	0.02	22.20	6.24	-3.78	1.59	0.53	0.99
AlN	$P63mc$	IB	IB	2690	1.00E-11	2.55	0.85	4.00	1.13	0.0941	3.26	6.13	4.43	-6.29	11.42	2.40	0.80	0.51
AlP	$F43m$	IB	IB	2827	1.01E-06	2.25	0.75	4.00	1.20	0.0480	2.36	2.45	8.19	-4.54	9.59	2.02	0.67	0.77
AlSb	$F43m$	CB	CB	1333	7.50E-01	1.89	0.63	4.00	1.11	0.0330	4.08	2.21	12.67	2.27	-1.08	1.68	0.56	0.94
As_2Te_3	$R\bar{3}m$	MVB	MVB	894	6.50E+02	7.41	2.47	4.76	-	0.0342	6.05	0.24	58.03	-5.83	-22.44	0.09	0.03	1.31
Au	$Fm3m$	MB	MB	1337	4.50E+05	0.00	0.00	12.02	-	0.0551	18.03	0.00	-	23.09	5.98	0.00	0.00	-
BaO	$Fm3m$	IB	IB	2196	5.21E-06	2.79	1.40	6.04	-	0.0452	5.75	2.09	4.13	-2.46	9.00	1.44	0.72	0.56
BaS	$Fm3m$	IB	IB	2458	3.96E-06	2.60	1.30	6.04	2.59	0.0300	4.18	2.15	4.77	-3.47	8.90	1.43	0.72	0.54
BaSe	$Fm3m$	IB	IB	2053	1.00E-05	2.59	1.29	6.04	2.47	0.0268	4.81	1.95	5.25	-2.57	7.82	1.39	0.70	0.54
BaTe	$Fm3m$	IB	IB	1738	5.11E-05	2.59	1.30	6.04	2.36	0.0224	4.94	1.59	5.94	-2.10	7.15	1.37	0.69	0.53
BeO	$P63mc$	IB	IB	2823	6.02E-23	1.80	0.90	4.00	1.65	0.1428	1.88	10.60	3.15	-7.14	12.18	1.73	0.87	0.28
BeS	$F43m$	IB	IB	853	8.54E-13	1.61	0.80	4.00	1.56	0.0690	2.36	5.50	5.53	-12.79	15.93	1.60	0.80	0.39
BeSe	$F43m$	IB	IB	1203	3.67E-07	1.56	0.78	4.00	0.33	0.0574	4.19	2.67	6.26	-7.30	0.03	1.54	0.77	0.40
BeTe	$F43m$	IB	IB	1077	2.02E-07	1.34	0.67	4.00	1.76	0.0440	4.99	2.80	6.80	-7.15	-0.13	1.45	0.73	0.51
Bi_2Se3	$R\bar{3}m$	MVB	MVB	983	1.00E+03	4.98	1.66	4.76	-	0.0329	7.16	0.25	17.16	-3.12	-18.45	1.01	0.34	1.14
Bi_2Te_3	$R\bar{3}m$	MVB	MVB	853	6.60E+02	6.91	2.30	5.10	1.49	0.0281	7.48	0.13	60.00	-3.61	-21.68	0.69	0.23	1.15
BN	$F43m$	IB	IB	3264	8.26E-10	1.92	0.64	4.00	1.51	0.1678	3.46	4.45	4.50	-6.73	12.78	2.23	0.74	0.88
C, diamond	$Fd3m$	CB	IB	4100	1.00E-12	0.00	0.00	4.00	0.96	0.1750	3.50	5.48	5.81	-6.76	12.79	0.00	0.00	1.83
Ca	$P63mmc$	MB	MB	1115	2.90E+05	0.00	0.00	12.01	-	0.0236	1.54	0.00	-	15.91	1.46	0.00	0.00	0.41
Ca	$Im3m$	MB	MB	1115	2.90E+05	0.00	0.00	11.27	-	0.0237	1.58	0.00	-	15.75	1.27	0.00	0.00	-
CaO	$Fm3m$	IB	IB	2901	1.22E-15	2.35	1.18	6.04	2.95	0.0706	3.29	6.93	3.77	-6.49	11.11	1.48	0.74	0.46
CaS	$Fm3m$	IB	IB	2713	1.37E-06	2.37	1.19	6.04	2.52	0.0428	2.57	2.38	4.98	-4.25	9.32	1.46	0.73	0.43
CaSe	$Fm3m$	IB	IB	1153	5.48E-06	2.38	1.19	6.04	2.17	0.0377	3.74	2.08	5.61	-2.03	5.02	1.45	0.73	0.42
CaTe	$Fm3m$	IB	IB	1873	6.11E-05	2.42	1.21	6.04	1.87	0.0306	4.25	1.55	6.61	-2.48	7.42	1.42	0.71	0.42
CdS	$F43m$	CB	CB	1678	5.88E-06	2.24	1.12	4.00	1.89	0.0382	4.58	2.48	6.36	-4.75	6.20	0.85	0.43	1.16
CdSe	$F43m$	CB	CB	1537	1.25E-06	2.33	1.17	4.00	1.85	0.0333	5.30	1.73	7.42	-4.29	5.00	0.72	0.36	1.16
CdTe	$F43m$	CB	CB	1346	3.80E-04	2.33	1.17	4.00	1.87	0.0275	5.47	1.48	8.92	-1.38	-5.52	0.52	0.26	1.23
CsBr	$Fm3m$	IB	IB	909	2.23E-16	1.24	1.24	6.04	-	0.0197	3.49	7.30	3.36	-15.74	18.75	-	-	-
CsCl	$Fm3m$	IB	IB	919	1.68E-11	1.26	1.26	6.04	-	0.0227	3.18	4.85	3.16	-16.50	16.17	0.85	0.85	0.21
CsF	$Fm3m$	IB	IB	976	2.57E-12	1.30	1.30	6.04	3.60	0.0349	4.41	5.26	2.29	-15.82	16.48	0.88	0.88	0.27
CsF	$Pm3m$	IB	IB	976	1.17E-13	1.34	1.34	11.27	-	0.0394	4.97	5.93	2.55	-18.42	15.28	0.87	0.87	0.18
CsI	$Fm3m$	IB	IB	899	2.75E-14	1.22	1.22	6.04	-	0.0165	3.55	6.25	3.68	-16.55	17.81	-	-	-
Cu	$Fm3m$	MB	MB	1358	5.90E+05	0.00	0.00	12.02	-	0.0842	8.89	0.00	-	19.00	6.85	0.00	0.00	0.51
CuAu	$P4/mmm$	MB	MB	683	5.15E+05	0.00	0.00	11.84	-	0.0666	14.40	0.00	-	21.74	5.28	0.00	0.00	-
GaN	$P63mc$	CB	IB	2791	1.08E-08	2.69	0.90	4.00	1.61	0.0852	5.92	3.44	6.00	-5.51	11.43	1.56	0.52	1.20
GaP	$F43m$	CB	CB	1763	1.00E-14	2.20	0.73	4.00	1.23	0.0479	4.01	2.27	11.06	-9.98	4.32	0.81	0.27	1.42
GaAs	$F43m$	CB	CB	1506	1.00E-08	2.20	0.73	4.00	1.21	0.0421	5.05	1.42	14.52	-6.59	3.03	0.68	0.23	1.40
GaSb	$F43m$	CB	CB	976	2.70E+02	2.91	0.97	4.00	1.23	0.0333	5.29	0.77	23.20	5.40	-6.62	0.30	0.10	1.44
GaSe	$P63/mmc$	CB	CB	1230	3.30E-02	1.74	0.87	3.50	-	0.0357	4.41	1.24	6.84	1.75	-1.63	0.67	0.22	1.33
Ge	$Fd3m$	CB	CB	1211	3.30E-02	0.00	0.00	4.00	1.14	0.0418	5.04	0.66	16.00	1.15	-0.40	0.00	0.00	1.57
GeSe	$Pnma$	CB	IB	950	1.30E-06	1.96	0.98	3.11	1.41	0.0407	5.12	1.10	16.73	-3.92	-2.62	0.65	0.33	1.38

Table 11.18: Compounds and respective properties used for classification (Part 2). Melting temperature T_{m} in (K), conductivity σ in (S/cm), atomic density ρ_{A} in $(10^{-2}/\text{Å}^3)$, density ρ in (g/cm^3), band gap E_{G} in (eV). Properties mostly obtained from the literature[204].

Compound	Structure	Exp.	EMA	T_{m}	σ	Z^*	Z^*_+	ECoN	γ_{TO}	ρ_{A}	ρ	E_{G}	ϵ_∞	t-SNE X	t-SNE Y	TET	ET	ES
GeTe	$R3m$	MVB	MVB	988	5.00E+03	5.98	2.99	4.91	-	0.0355	5.91	0.10	44.10	-7.12	-24.86	0.38	0.19	1.11
Hf	$P63mmc$	MB	MB	2506	3.30E+04	0.00	0.00	11.96	-	0.0445	13.18	0.00	-	23.31	8.94	0.02	0.01	0.80
HgS	$F43m$	CB	IB	1148	2.10E-13	3.40	1.70	4.00	3.28	0.0369	7.13	2.03	24.60	-5.86	-3.66	0.55	0.28	1.34
HgSe	$F43m$	CB	CB	1070	1.80E-06	3.38	1.69	4.00	2.55	0.0324	7.52	0.06	19.92	-3.09	-14.75	0.40	0.20	1.32
HgTe	$F43m$	CB	CB	943	9.30E+02	3.30	1.65	4.00	2.00	0.0271	7.38	0.14	20.16	-3.12	-17.78	0.10	0.05	1.37
InAs	$F43m$	CB	CB	1211	5.00E+00	2.74	0.91	4.00	1.33	0.0336	5.29	0.35	16.34	3.16	-5.62	0.72	0.24	1.40
InP	$F43m$	CB	CB	1328	1.67E-07	2.60	0.87	4.00	1.33	0.0378	4.58	1.34	12.72	-5.14	0.85	0.88	0.29	1.42
InSb	$F43m$	CB	CB	798	5.00E+01	2.50	0.83	4.00	1.41	0.0274	5.38	0.18	21.34	6.99	-5.45	0.55	0.18	1.39
K	$Im3m$	MB	MB	337	1.40E+05	0.00	0.00	11.28	0.79	0.0136	0.88	0.00	-	15.58	0.41	0.00	0.00	-
KBr	$Fm3m$	IB	IB	1007	2.00E-10	1.14	1.14	6.04	2.99	0.0264	2.61	4.31	2.49	-16.36	15.47	0.83	0.83	0.22
KCl	$Fm3m$	IB	IB	1038	7.46E-12	1.16	1.16	6.04	2.75	0.0306	1.89	5.03	2.80	-15.69	15.95	0.84	0.84	0.21
KF	$Fm3m$	IB	IB	1131	8.79E-14	1.13	1.13	6.04	2.73	0.0501	2.42	6.00	1.95	-13.82	16.44	0.86	0.86	0.23
KI	$Fm3m$	IB	IB	949	1.64E-09	1.16	1.16	6.04	2.75	0.0216	2.97	3.85	2.80	-16.66	15.17	0.82	0.82	0.20
Li	$Im3m$	MB	MB	454	1.10E+05	0.00	0.00	11.27	-	0.0497	0.57	0.00	-	16.68	0.70	0.00	0.00	-
Mg	$P63mmc$	MB	MB	923	2.30E+05	0.00	0.00	12.01	-	0.0434	1.75	0.00	-	16.54	1.30	0.00	0.00	0.40
MgO	$Fm3m$	IB	IB	3433	4.09E-17	2.00	1.00	6.04	2.53	0.1037	3.47	7.67	3.23	-6.73	11.63	1.70	0.85	0.24
MgS	$Fm3m$	IB	IB	2200	8.36E-11	2.35	1.18	6.04	3.09	0.0560	2.02	4.50	5.44	-5.76	9.11	1.67	0.84	0.24
MgSe	$F43m$	IB	IB	1560	6.57E-10	1.91	0.96	4.00	1.48	0.0370	3.17	4.05	4.68	-9.02	7.86	1.59	0.80	0.41
MgTe	$F43m$	IB	IB	1300	8.56E-09	1.95	0.98	4.00	1.48	0.0290	3.65	3.49	5.55	-9.09	7.88	1.55	0.78	0.43
Na	$Im3m$	MB	MB	371	2.10E+05	0.00	0.00	11.27	-	0.0275	1.05	0.00	-	15.99	0.49	0.00	0.00	0.20
NaCl	$Fm3m$	IB	IB	1073	8.33E-12	1.10	1.10	6.02	2.43	0.0434	2.16	8.50	2.49	-13.69	18.17	0.87	0.87	0.14
NaBr	$Fm3m$	IB	IB	1028	8.00E-10	1.13	1.13	6.04	-	0.0365	3.12	7.00	2.61	-14.67	17.42	-	-	-
NaF	$Fm3m$	IB	IB	1268	5.61E-14	1.02	1.02	6.04	2.50	0.0772	2.69	6.09	1.86	-12.75	15.96	0.88	0.88	0.15
Nb	$Im3m$	MB	MB	2750	6.70E+04	0.00	0.00	11.27	-	0.0546	8.43	0.00	-	22.80	10.33	0.00	0.00	-
Ni	$Fm3m$	MB	MB	1728	1.40E+05	0.00	0.00	12.02	-	0.0927	9.03	0.00	-	18.40	8.12	0.00	0.00	0.60
NiPt	$P4/mmm$	MB	MB	1720	1.15E+05	0.00	0.00	11.96	-	0.0746	15.72	0.00	-	22.61	6.51	0.00	0.00	-
Pb	$Fm3m$	MB	MB	601	5.20E+04	0.00	0.00	12.02	-	0.0310	10.65	0.00	-	18.60	3.32	0.00	0.00	0.45
PbO	$P4nmm$	CB	IB	1908	4.00E-04	2.76	1.38	4.53	0.39	0.0463	8.58	2.42	6.06	-1.51	9.14	1.17	0.59	0.94
PbS	$Fm\bar{3}m$	MVB	MVB	1387	1.18E+02	4.37	2.18	6.04	19.97	0.0370	7.35	0.56	15.91	-1.09	-26.00	1.00	0.50	0.74
PbSe	$Fm\bar{3}m$	MVB	MVB	1352	2.40E+02	4.80	2.40	6.04	17.53	0.0334	7.94	0.47	20.21	-1.93	-26.47	0.86	0.43	0.76
PbTe	$Fm\bar{3}m$	MVB	MVB	1197	2.90E+03	5.76	2.88	6.04	21.39	0.0283	7.85	0.24	26.82	-4.35	-26.59	0.68	0.34	0.80
Pd	$Fm3m$	MB	MB	1828	1.00E+05	0.00	0.00	12.02	-	0.0647	11.44	0.00	-	21.43	7.66	0.00	0.00	0.62
RbBr	$Fm3m$	IB	IB	966	3.41E-10	1.16	1.16	6.04	2.70	0.0230	3.16	4.19	2.47	-16.66	15.45	0.84	0.84	0.23
RbBr	$Pm3m$	IB	IB	966	1.18E-09	1.24	1.24	11.27	-	0.0270	3.70	3.92	2.83	-18.30	15.16	0.83	0.83	0.17
RbCl	$Fm3m$	IB	IB	1020	1.77E-11	1.17	1.17	6.04	2.70	0.0266	2.67	4.84	2.28	-16.22	16.00	0.85	0.85	0.23
RbI	$Fm3m$	IB	IB	1577	8.63E-14	1.16	1.16	6.04	-	0.0190	3.36	6.00	2.59	-17.56	17.44	-	-	-
Sb2S3	$Pnma$	CB	IB	823	1.00E-08	3.40	1.42	3.15	-	0.0380	4.28	1.88	11.32	-4.95	-3.39	0.96	0.40	1.12
Sb2Se3	$Pnma$	CB	IB	884	4.00E-07	3.54	1.48	3.38	-	0.0335	5.34	1.17	15.74	-4.58	-3.97	0.93	0.32	1.52
$Sb2Te_3$	$R\bar{3}m$	MVB	MVB	893	2.30E+03	5.93	1.98	4.85	-	0.0294	6.12	0.11	37.69	-4.44	-19.89	0.39	0.13	1.24
Sc	$P63mmc$	MB	MB	1814	1.80E+04	0.00	0.00	11.89	0.27	0.0406	3.03	0.00	-	15.72	2.71	0.00	0.00	-
Si	$Fd3m$	CB	CB	1683	1.50E-08	0.00	0.00	4.00	0.99	0.0489	2.28	1.22	13.00	-9.39	1.75	0.00	0.00	1.68
SnO	$P4nmm$	CB	IB	2075	2.30E-02	2.96	1.48	4.59	5.68	0.0531	5.94	0.41	7.28	-1.89	9.35	1.27	0.64	0.89
SnS	$Pnma$	CB	IB	1129	3.40E-04	3.39	1.70	3.33	1.31	0.0391	4.90	1.16	14.06	-3.97	-5.03	0.97	0.49	1.12
SnSe	$Pnma$	CB	CB	1147	2.50E-05	3.49	1.75	3.72	1.97	0.0355	5.83	0.86	17.11	-4.01	-5.09	0.83	0.42	1.24
SnTe	$Fm3m$	MVB	MVB	1080	9.80E+03	6.65	3.33	6.04	5.83	0.0303	6.19	0.26	79.17	-6.84	-26.61	0.63	0.31	0.94
Sr	$Im3m$	MB	MB	1050	7.70E+04	0.00	0.00	11.27	-	0.0186	2.71	0.00	-	15.44	1.60	0.00	0.00	-
SrO	$Fm3m$	IB	IB	3643	3.08E-12	2.43	1.22	6.04	3.27	0.0567	4.88	5.22	3.77	-6.15	10.67	1.48	0.74	0.50
SrS	$Fm3m$	IB	IB	2558	7.96E-07	2.40	1.20	6.04	2.32	0.0359	3.57	2.50	4.62	-3.91	9.12	1.47	0.74	0.46
SrSe	$Fm3m$	IB	IB	1873	2.74E-06	2.40	1.20	6.04	2.32	0.0319	4.42	2.23	5.14	-2.80	7.48	1.45	0.73	0.46
SrTe	$Fm3m$	IB	IB	1766	2.36E-05	2.43	1.22	6.04	2.41	0.0263	4.70	1.76	5.92	-2.29	7.20	1.43	0.72	0.45

Table 11.19: Compounds and respective properties used for classification (Part 3). Melting temperature T_{m} in (K), conductivity σ in (S/cm), atomic density ρ_{A} in (10^{-2}/Å^3), density ρ in (g/cm^3), band gap E_{G} in (eV). Properties mostly obtained from the literature[204].

Compound	Structure	Exp.	EMA	T_{m}	σ	Z^*	Z^*_+	ECoN	γ_{TO}	ρ_{A}	ρ	E_{G}	ϵ_∞	t-SNE X	t-SNE Y	TET	ET	ES
Ta	$Im3m$	MB	MB	3290	7.70E+04	0.00	0.00	11.27	2.20	0.0548	16.46	0.00	-	24.43	8.57	0.00	0.00	0.98
Y	$P63mmc$	MB	MB	1799	1.80E+04	0.00	0.00	11.87	-	0.0304	4.45	0.00	-	15.80	2.77	0.00	0.00	-
ZnO	$F43m$	CB	IB	1159	1.00E-08	2.22	1.11	4.00	1.31	0.0806	5.45	3.44	4.00	-7.94	5.88	1.22	0.61	0.98
ZnS	$F43m$	CB	IB	1991	6.25E-07	2.00	1.00	4.00	1.80	0.0494	4.00	3.76	5.80	-5.62	7.75	0.88	0.44	1.17
ZnSe	$F43m$	CB	IB	1803	1.00E-05	2.09	1.05	4.00	1.79	0.0423	5.06	2.82	7.27	-5.02	6.73	0.67	0.35	1.17
ZnTe	$F43m$	CB	CB	1563	5.90E-06	2.08	1.04	4.00	1.77	0.0338	5.42	2.35	8.96	-4.61	5.49	0.51	0.25	1.22
Zr	$P63mmc$	MB	MB	2130	2.40E+04	0.00	0.00	11.99	-	0.0426	6.45	0.00	-	21.66	10.19	0.00	0.00	0.80
AlAu	$Pm3m$	MB	MB	1000	1.72E+05	0.00	0.00	11.28	-	0.0597	11.10	0.00	-	20.31	5.33	-	-	-
AlCu	$Pm3m$	MB	MB	1085	1.15E+05	0.00	0.00	11.28	-	0.0743	5.58	0.00	-	18.20	5.33	-	-	-
AlPd	$P213$	MB	MB	1918	1.42E+04	0.00	0.00	10.06	-	0.0676	7.49	0.00	-	19.98	9.54	-	-	-
AlPt	$P213$	MB	MB	1773	2.75E+04	0.00	0.00	10.25	-	0.0670	12.40	0.00	-	21.66	7.13	1.98	0.66	-
Co		MB	MB	1768	1.60E+05	0.00	0.00	12.02	2.10	0.0916	8.97	0.00	-	19.41	8.48	0.00	0.00	0.71
Cu	$Fm3m$	MB	MB	1358	5.90E+05	0.00	0.00	12.02	-	0.0842	8.89	0.00	-	19.00	6.85	0.00	0.00	0.51
CuZr	$Pm3m$	MB	MB	1225	1.33E+04	0.00	0.00	11.28	-	0.0570	7.33	0.00	-	18.85	5.03	-	-	-
Fe	-	MB	MB	1811	1.20E+05	0.00	0.00	11.28	1.83	0.0873	8.10	0.00	40.00	19.18	9.01	0.00	0.00	0.95
Ga	$Cmce$	MB	MB	302	3.70E+04	0.00	0.00	6.62	-	0.0496	5.10	0.00	25.00	9.26	-3.16	-	-	-
GaPd	$P213$	MB	MB	1305	6.70E+04	0.00	0.00	9.63	-	0.0657	9.61	0.00	-	19.69	5.74	-	-	-
GaPt	$P213$	MB	MB	1377	2.85E+04	0.00	0.00	9.99	-	0.0649	14.27	0.00	-	21.97	6.05	-	-	-
In	$R3m$	MB	MB	430	1.20E+05	0.00	0.00	11.94	-	0.0368	7.02	0.00	-	17.87	2.95	-	-	-
Ir	$Fm3m$	MB	MB	2739	2.12E+05	0.00	0.00	6.02	2.30	0.0687	21.93	0.00	50.00	24.54	7.33	-	-	-
La	$P63mmc$	MB	MB	1193	1.63E+04	0.00	0.00	12.00	0.70	0.0269	6.21	0.00	-	16.54	2.55	-	-	-
Mn	$Fm3m$	MB	MB	1519	6.94E+03	0.00	0.00	12.02	-	0.0925	8.44	0.00	-	18.16	7.61	0.00	0.00	0.91
Mo	$Im3m$	MB	MB	2896	1.87E+05	0.00	0.00	11.28	-	0.0629	10.02	0.00	-	23.34	10.15	-	-	-
NbC	$Fm3m$	MB	MB	3873	2.50E+04	0.00	0.00	6.04	1.71	0.0886	7.62	0.00	-	20.89	12.08	1.81	0.45	1.15
NbN	$Fm3m$	MB	MB	3273	1.25E+04	0.00	0.00	6.04	-	0.0914	8.05	0.00	-	20.84	12.04	1.69	0.56	1.00
Ni	$Fm3m$	MB	MB	1728	1.40E+05	0.00	0.00	12.02	-	0.0927	9.03	0.00	-	18.40	8.12	0.00	0.00	0.60
NiO	$Fm3m$	CB	IB	2230	1.00E-08	2.31	1.16	6.04	1.47	0.1077	6.62	0.50	5.72	-4.87	11.83	1.05	0.53	0.76
NiAl	$Pm3m$	MB	MB	1924	9.09E+04	0.00	0.00	11.27	8.45	0.0834	5.91	0.00	-	18.88	9.55	0.49	0.16	0.55
Pt	$Fm3m$	MB	MB	2041	9.52E+04	0.00	0.00	12.02	2.30	0.0636	20.60	0.00	50.00	23.81	6.71	-	-	-
Re	$P63mmc$	MB	MB	3458	5.18E+04	0.00	0.00	12.01	1.83	0.0664	20.53	0.00	-	24.62	7.95	-	-	-
Rh	$Fm3m$	MB	MB	2240	2.31E+05	0.00	0.00	12.02	-	0.0704	12.03	0.00	-	21.97	8.17	-	-	-
TiO	$Fm3m$	MB	MB	2037	6.45E+03	0.00	0.00	6.04	-	0.1007	5.38	0.00	-	18.57	10.47	1.52	0.76	0.69
Zn	-	MB	MB	693	1.79E+05	0.00	0.00	9.61	-	0.0643	6.98	0.00	-	18.35	4.57	0.00	0.00	0.61
In_3SbTe_2	-	MB	MB	915	1.00E+04	0.00	0.00	6.03	-	0.0335	6.69	0.00	-	8.69	-3.40	-	0.20	0.82
CdS:0.2-CdSe:0.8	$F43m$	CB	CB	1565	1.70E-06	2.31	1.16	4.00	1.86	0.0343	5.16	1.88	7.21	-4.35	5.22	0.75	0.37	1.16
CdS:0.4-CdSe:0.6	$F43m$	CB	CB	1594	2.32E-06	2.29	1.15	4.00	1.87	0.0353	5.01	2.03	7.00	-4.43	5.48	0.77	0.39	1.16
CdS:0.6-CdSe:0.4	$F43m$	CB	CB	1622	3.17E-06	2.28	1.14	4.00	1.87	0.0362	4.87	2.18	6.78	-4.53	5.75	0.80	0.40	1.16
CdS:0.8-CdSe:0.2	$F43m$	CB	CB	1650	4.31E-06	2.26	1.13	4.00	1.88	0.0372	4.72	2.33	6.57	-4.64	5.99	0.82	0.41	1.16
CdTe:0.2-HgTe:0.8	$F43m$	CB	CB	1024	4.91E+01	3.11	1.55	4.00	1.97	0.0272	7.00	0.41	17.91	-2.88	-17.05	0.18	0.09	1.34
CdTe:0.4-HgTe:0.6	$F43m$	CB	CB	1104	2.59E+00	2.91	1.46	4.00	1.95	0.0273	6.62	0.68	15.66	-2.54	-16.39	0.27	0.13	1.31
CdTe:0.6-HgTe:0.4	$F43m$	CB	CB	1185	1.37E-01	2.72	1.36	4.00	1.92	0.0273	6.23	0.94	13.42	-1.08	-6.45	0.35	0.18	1.29
CdTe:0.8-HgTe:0.2	$F43m$	CB	CB	1266	7.20E-03	2.52	1.26	4.00	1.90	0.0274	5.85	1.21	11.17	-1.24	-5.88	0.44	0.22	1.26
HgSe:0.2-HgTe:0.8	$F43m$	CB	CB	969	1.68E+01	3.32	1.66	4.00	2.11	0.0282	7.41	0.12	20.11	-3.07	-17.18	0.16	0.08	1.36
HgSe:0.4-HgTe:0.6	$F43m$	CB	CB	994	3.04E-01	3.33	1.67	4.00	2.22	0.0292	7.44	0.11	20.06	-3.09	-16.30	0.22	0.11	1.35
HgSe:0.6-HgTe:0.4	$F43m$	CB	CB	1019	5.50E-03	3.35	1.67	4.00	2.33	0.0303	7.46	0.09	20.02	-3.07	-15.51	0.28	0.14	1.34
HgSe:0.8-HgTe:0.2	$F43m$	CB	CB	1045	9.95E-05	3.36	1.68	4.00	2.44	0.0313	7.49	0.08	19.97	-3.08	-15.01	0.34	0.17	1.33
PbS:0.2-PbSe:0.8	$Fm\bar{3}m$	MVB	MVB	1359	2.08E+02	4.71	2.36	6.04	18.02	0.0341	7.82	0.49	19.35	-1.75	-26.26	0.89	0.44	0.76
PbS:0.4-PbSe:0.6	$Fm\bar{3}m$	MVB	MVB	1366	1.81E+02	4.63	2.31	6.04	18.51	0.0348	7.70	0.51	18.49	-1.50	-26.37	0.92	0.46	0.75
PbS:0.6-PbSe:0.4	$Fm\bar{3}m$	MVB	MVB	1373	1.57E+02	4.54	2.27	6.04	18.99	0.0356	7.59	0.52	17.63	-1.37	-26.10	0.94	0.47	0.75
PbS:0.8-PbSe:0.2	$Fm\bar{3}m$	MVB	MVB	1380	1.36E+02	4.45	2.23	6.04	19.48	0.0363	7.47	0.54	16.77	-1.20	-26.10	0.97	0.49	0.74

Table 11.20: Compounds and respective properties used for classification (Part 4). Melting temperature T_{m} in (K), conductivity σ in (S/cm), atomic density ρ_{A} in ($10^{-2}/\mathrm{\AA}^3$), density ρ in (g/cm^3), band gap E_{G} in (eV). Properties mostly obtained from the literature[204].

Compound	Structure	Exp.	EMA	T_{m}	σ	Z^*	Z^*_+	ECoN	γ_{TO}	ρ_{A}	ρ	E_{G}	ϵ_∞	t-SNE X	t-SNE Y	TET	ET	ES
PbTe:0.2-PbSe:0.8	$Fm\bar{3}m$	MVB	MVB	1321	3.95E+02	4.99	2.50	6.04	18.30	0.0324	7.92	0.42	21.53	-2.36	-26.38	0.82	0.41	0.77
PbTe:0.4-PbSe:0.6	$Fm\bar{3}m$	MVB	MVB	1290	6.50E+02	5.18	2.59	6.04	19.07	0.0314	7.90	0.38	22.85	-2.85	-26.39	0.79	0.39	0.78
PbTe:0.6-PbSe:0.4	$Fm\bar{3}m$	MVB	MVB	1259	1.07E+03	5.38	2.69	6.04	19.85	0.0303	7.89	0.33	24.18	-3.34	-26.41	0.75	0.38	0.79
PbTe:0.8-PbSe:0.2	$Fm\bar{3}m$	MVB	MVB	1228	1.76E+03	5.57	2.78	6.04	20.62	0.0293	7.87	0.29	25.50	-3.85	-26.54	0.72	0.36	0.80
PbTe:0.2-SnTe:0.8	$Fm\bar{3}m$	MVB	MVB	1104	7.68E+03	6.47	3.24	6.04	8.94	0.0299	6.52	0.26	68.70	-6.41	-26.55	0.64	0.32	0.92
PbTe:0.4-SnTe:0.6	$Fm\bar{3}m$	MVB	MVB	1127	6.02E+03	6.29	3.15	6.04	12.05	0.0295	6.85	0.25	58.23	-5.99	-26.53	0.65	0.32	0.89
PbTe:0.6-SnTe:0.4	$Fm\bar{3}m$	MVB	MVB	1150	4.72E+03	6.12	3.06	6.04	15.17	0.0291	7.19	0.25	47.76	-5.48	-26.52	0.66	0.33	0.86
PbTe:0.8-SnTe:0.2	$Fm\bar{3}m$	MVB	MVB	1174	3.70E+03	5.94	2.97	6.04	18.28	0.0287	7.52	0.24	37.29	-4.90	-26.55	0.67	0.33	0.83
GeTe:0.2-SnTe:0.8	$R3m$	MVB	MVB	1062	8.57E+03	6.52	3.26	5.81	-	0.0313	6.13	0.23	72.16	-6.92	-26.29	0.58	0.29	0.98
GeTe:0.4-SnTe:0.6	$R3m$	MVB	MVB	1043	7.49E+03	6.38	3.19	5.59	-	0.0324	6.08	0.20	65.14	-6.97	-25.94	0.53	0.26	1.01
GeTe:0.6-SnTe:0.4	$R3m$	MVB	MVB	1025	6.54E+03	6.25	3.12	5.36	-	0.0334	6.02	0.16	58.13	-7.03	-25.59	0.48	0.24	1.04
GeTe:0.8-SnTe:0.2	$R3m$	MVB	MVB	1006	5.72E+03	6.11	3.06	5.13	-	0.0345	5.97	0.13	51.11	-7.04	-25.21	0.43	0.21	1.07
$Sb2Te_3$:0.2-Bi_2Te_3:0.8	$R\bar{3}m$	MVB	MVB	861	8.47E+02	6.71	2.24	5.05	-	0.0284	7.21	0.13	55.54	-3.68	-21.30	0.63	0.21	1.17
$Sb2Te_3$:0.4-Bi_2Te_3:0.6	$R\bar{3}m$	MVB	MVB	869	1.09E+03	6.52	2.17	5.00	-	0.0286	6.94	0.12	51.08	-3.85	-20.89	0.57	0.19	1.19
$Sb2Te_3$:0.6-Bi_2Te_3:0.4	$R\bar{3}m$	MVB	MVB	877	1.40E+03	6.32	2.11	4.95	-	0.0289	6.66	0.12	46.61	-4.07	-20.53	0.51	0.17	1.21
$Sb2Te_3$:0.8-Bi_2Te_3:0.2	$R\bar{3}m$	MVB	MVB	885	1.79E+03	6.13	2.04	4.90	-	0.0291	6.39	0.11	42.15	-4.28	-20.16	0.45	0.15	1.23
Bi_2Se3:0.2-Bi_2Te_3:0.8	$R\bar{3}m$	MVB	MVB	879	7.17E+02	6.52	2.17	5.03	-	0.0291	7.42	0.15	51.43	-3.33	-20.90	0.75	0.25	1.15
Bi_2Se3:0.4-Bi_2Te_3:0.6	$R\bar{3}m$	MVB	MVB	905	7.79E+02	6.14	2.05	4.97	-	0.0300	7.35	0.18	42.86	-3.24	-20.22	0.82	0.27	1.15
Bi_2Se3:0.6-Bi_2Te_3:0.4	$R\bar{3}m$	MVB	MVB	931	8.47E+02	5.75	1.92	4.90	-	0.0310	7.29	0.20	34.30	-3.22	-19.55	0.88	0.30	1.14
Bi_2Se3:0.8-Bi_2Te_3:0.2	$R\bar{3}m$	MVB	MVB	957	9.20E+02	5.37	1.79	4.83	-	0.0319	7.22	0.23	25.73	-3.19	-18.92	0.95	0.32	1.14
PbTe:0.2-$AgSbTe_2$:0.8	$Fm\bar{3}m$	MVB	MVB	917	5.03E+02	4.59	2.30	6.02	6.70	0.0334	7.15	0.33	36.73	-2.65	-22.50	0.36	0.18	0.91
PbTe:0.4-$AgSbTe_2$:0.6	$Fm\bar{3}m$	MVB	MVB	987	7.80E+02	4.88	2.44	6.03	10.37	0.0321	7.33	0.31	34.25	-3.22	-23.60	0.44	0.22	0.88
PbTe:0.6-$AgSbTe_2$:0.4	$Fm\bar{3}m$	MVB	MVB	1057	1.21E+03	5.18	2.59	6.03	14.05	0.0309	7.50	0.28	31.78	-3.59	-24.93	0.52	0.26	0.86
PbTe:0.8-$AgSbTe_2$:0.2	$Fm\bar{3}m$	MVB	MVB	1127	1.87E+03	5.47	2.73	6.03	17.72	0.0296	7.68	0.26	29.30	-3.96	-25.88	0.60	0.30	0.83
GeTe:0.2-$AgSbTe_2$:0.8	$R3m$	MVB	MVB	875	5.61E+02	4.64	2.32	5.80	-	0.0349	6.77	0.30	40.19	-3.00	-22.35	0.30	0.15	0.97
GeTe:0.4-$AgSbTe_2$:0.6	$R3m$	MVB	MVB	903	9.70E+02	4.97	2.49	5.57	-	0.0350	6.55	0.25	41.17	-4.87	-23.02	0.32	0.16	1.00
GeTe:0.6-$AgSbTe_2$:0.4	$R3m$	MVB	MVB	932	1.68E+03	5.31	2.65	5.35	-	0.0352	6.34	0.20	42.14	-5.93	-23.66	0.34	0.17	1.04
GeTe:0.8-$AgSbTe_2$:0.2	$R3m$	MVB	MVB	960	2.89E+03	5.64	2.82	5.13	-	0.0353	6.12	0.15	43.12	-6.65	-24.30	0.36	0.18	1.07
$AgSbTe_2$:0.2-$AgBiTe_2$:0.8	$R3m$	MVB	MVB	832	9.85E+02	4.26	2.13	6.01	6.02	0.0336	7.73	0.20	53.88	-2.04	-21.20	0.34	0.17	0.91
$AgSbTe_2$:0.4-$AgBiTe_2$:0.6	$R3m$	MVB	MVB	836	7.47E+02	4.27	2.13	6.01	5.27	0.0339	7.54	0.24	50.22	-2.17	-21.41	0.33	0.16	0.92
$AgSbTe_2$:0.6-$AgBiTe_2$:0.4	$R3m$	MVB	MVB	840	5.66E+02	4.28	2.14	6.01	4.53	0.0341	7.36	0.27	46.55	-2.22	-21.62	0.31	0.16	0.92
$AgSbTe_2$:0.8-$AgBiTe_2$:0.2	$R3m$	MVB	MVB	843	4.29E+02	4.29	2.14	6.02	3.78	0.0344	7.17	0.31	42.88	-2.22	-21.82	0.30	0.15	0.93
GeTe:0.2-$Sb2Te_3$:0.8	$R3m$	MVB	MVB	912	2.69E+03	5.94	2.18	4.86	-	0.0306	6.08	0.11	38.97	-4.96	-20.87	0.39	0.14	1.22
GeTe:0.4-$Sb2Te_3$:0.6	$R3m$	MVB	MVB	931	3.14E+03	5.95	2.38	4.87	-	0.0318	6.04	0.11	40.25	-5.43	-22.00	0.39	0.15	1.19
GeTe:0.6-$Sb2Te_3$:0.4	$R3m$	MVB	MVB	950	3.66E+03	5.96	2.58	4.88	-	0.0331	5.99	0.10	41.54	-6.13	-23.15	0.38	0.17	1.16
GeTe:0.8-$Sb2Te_3$:0.2	$R3m$	MVB	MVB	969	4.28E+03	5.97	2.79	4.90	-	0.0343	5.95	0.10	42.82	-6.68	-24.09	0.38	0.18	1.13
SnTe:0.2-$Sb2Te_3$:0.8	$Fm3m$	MVB	MVB	930	3.07E+03	6.07	2.25	5.09	-	0.0296	6.13	0.14	45.99	-4.90	-21.38	0.44	0.17	1.18
SnTe:0.4-$Sb2Te_3$:0.6	$Fm3m$	MVB	MVB	968	4.11E+03	6.22	2.52	5.33	-	0.0298	6.15	0.17	54.28	-5.47	-22.99	0.49	0.20	1.12
SnTe:0.6-$Sb2Te_3$:0.4	$Fm3m$	MVB	MVB	1005	5.49E+03	6.36	2.79	5.56	-	0.0299	6.16	0.20	62.58	-6.10	-24.63	0.53	0.24	1.06
SnTe:0.8-$Sb2Te_3$:0.2	$Fm3m$	MVB	MVB	1043	7.33E+03	6.51	3.06	5.80	-	0.0301	6.18	0.23	70.87	-6.40	-25.75	0.58	0.28	1.00
PbTe:0.2-$Sb2Te_3$:0.8	$Fm\bar{3}m$	MVB	MVB	954	2.41E+03	5.90	2.16	5.09	-	0.0292	6.47	0.14	35.52	-4.48	-20.98	0.45	0.17	1.16
PbTe:0.4-$Sb2Te_3$:0.6	$Fm\bar{3}m$	MVB	MVB	1015	2.52E+03	5.86	2.34	5.33	-	0.0290	6.81	0.16	33.34	-4.30	-22.34	0.51	0.21	1.07
PbTe:0.6-$Sb2Te_3$:0.4	$Fm\bar{3}m$	MVB	MVB	1075	2.64E+03	5.83	2.52	5.56	-	0.0287	7.16	0.19	31.17	-4.01	-24.35	0.56	0.26	0.98
PbTe:0.8-$Sb2Te_3$:0.2	$Fm\bar{3}m$	MVB	MVB	1136	2.77E+03	5.79	2.70	5.80	-	0.0285	7.50	0.21	29.00	-4.06	-25.63	0.62	0.30	0.89
PbTe:0.2-Bi_2Te_3:0.8	$Fm\bar{3}m$	MVB	MVB	922	8.87E+02	6.68	2.42	5.29	5.47	0.0281	7.55	0.15	53.36	-3.71	-22.90	0.69	0.25	1.08
PbTe:0.4-Bi_2Te_3:0.6	$Fm\bar{3}m$	MVB	MVB	991	1.19E+03	6.45	2.53	5.48	9.45	0.0282	7.63	0.17	46.73	-3.76	-24.14	0.69	0.27	1.01
PbTe:0.6-Bi_2Te_3:0.4	$Fm\bar{3}m$	MVB	MVB	1059	1.60E+03	6.22	2.65	5.66	13.43	0.0282	7.70	0.20	40.09	-3.93	-25.13	0.68	0.30	0.94
PbTe:0.8-Bi_2Te_3:0.2	$Fm\bar{3}m$	MVB	MVB	1128	2.16E+03	5.99	2.76	5.85	17.41	0.0283	7.78	0.22	33.46	-4.18	-25.95	0.68	0.32	0.87
Ta:0.2-Nb:0.8	$Im3m$	MB	MB	2858	6.89E+04	0.00	0.00	11.27	-	0.0546	10.04	0.00	-	23.24	10.02	0.00	0.00	-
Ta:0.4-Nb:0.6	$Im3m$	MB	MB	2966	7.08E+04	0.00	0.00	11.27	-	0.0547	11.64	0.00	-	23.65	9.66	0.00	0.00	-

Table 11.21: Compounds and respective properties used for classification (Part 5). Melting temperature T_{m} in (K), conductivity σ in (S/cm), atomic density ρ_{A} in (10^{-2}/Å^3), density ρ in (g/cm^3), band gap E_{G} in (eV). Properties mostly obtained from the literature[204].

Compound	Structure	Exp.	EMA	T_{m}	σ	Z^*	Z^*_+	ECoN	γ_{TO}	ρ_{A}	ρ	E_{G}	ϵ_∞	t-SNE X	t-SNE Y	TET	ET	ES
Ta:0.6-Nb:0.4	$Im3m$	MB	MB	3074	7.28E+04	0.00	0.00	11.27	-	0.0547	13.25	0.00	-	23.99	9.29	0.00	0.00	-
Ta:0.8-Nb:0.2	$Im3m$	MB	MB	3182	7.49E+04	0.00	0.00	11.27	-	0.0548	14.85	0.00	-	24.25	8.92	0.00	0.00	-
Ni:0.2-Pd:0.8	$Fm3m$	MB	MB	1808	1.07E+05	0.00	0.00	12.02	-	0.0703	10.96	0.00	-	21.01	7.66	0.00	0.00	0.62
Ni:0.4-Pd:0.6	$Fm3m$	MB	MB	1788	1.14E+05	0.00	0.00	12.02	-	0.0759	10.48	0.00	-	20.54	7.74	0.00	0.00	0.61
Ni:0.6-Pd:0.4	$Fm3m$	MB	MB	1768	1.22E+05	0.00	0.00	12.02	-	0.0815	9.99	0.00	-	20.10	7.96	0.00	0.00	0.61
Ni:0.8-Pd:0.2	$Fm3m$	MB	MB	1748	1.31E+05	0.00	0.00	12.02	-	0.0871	9.51	0.00	-	19.70	8.01	0.00	0.00	0.60
Ni:0.2-Cu:0.8	$Fm3m$	MB	MB	1432	4.42E+05	0.00	0.00	12.02	-	0.0859	8.92	0.00	-	18.63	7.16	0.00	0.00	0.53
Ni:0.4-Cu:0.6	$Fm3m$	MB	MB	1506	3.32E+05	0.00	0.00	12.02	-	0.0876	8.95	0.00	-	19.21	7.36	0.00	0.00	0.54
Ni:0.6-Cu:0.4	$Fm3m$	MB	MB	1580	2.49E+05	0.00	0.00	12.02	-	0.0893	8.97	0.00	-	19.10	7.61	0.00	0.00	0.56
Ni:0.8-Cu:0.2	$Fm3m$	MB	MB	1654	1.87E+05	0.00	0.00	12.02	-	0.0910	9.00	0.00	-	18.99	7.83	0.00	0.00	0.58
Sr:0.2-Ca:0.8	$Im3m$	MB	MB	1102	2.22E+05	0.00	0.00	11.27	-	0.0227	1.81	0.00	-	15.74	1.37	0.00	0.00	-
Sr:0.4-Ca:0.6	$Im3m$	MB	MB	1089	1.71E+05	0.00	0.00	11.27	-	0.0217	2.03	0.00	-	15.46	1.28	0.00	0.00	-
Sr:0.6-Ca:0.4	$Im3m$	MB	MB	1076	1.31E+05	0.00	0.00	11.27	-	0.0206	2.26	0.00	-	15.41	1.37	0.00	0.00	-
Sr:0.8-Ca:0.2	$Im3m$	MB	MB	1063	1.00E+05	0.00	0.00	11.27	-	0.0196	2.48	0.00	-	15.49	1.58	0.00	0.00	-
Hf:0.2-Zr:0.8	$P63mmc$	MB	MB	2205	2.56E+04	0.00	0.00	11.98	-	0.0430	7.80	0.00	-	21.96	10.09	0.00	0.00	0.80
Hf:0.4-Zr:0.6	$P63mmc$	MB	MB	2280	2.73E+04	0.00	0.00	11.98	-	0.0434	9.14	0.00	-	22.28	9.84	0.01	0.00	0.80
Hf:0.6-Zr:0.4	$P63mmc$	MB	MB	2356	2.91E+04	0.00	0.00	11.97	-	0.0437	10.49	0.00	-	22.62	9.52	0.01	0.01	0.80
Hf:0.8-Zr:0.2	$P63mmc$	MB	MB	2431	3.10E+04	0.00	0.00	11.97	-	0.0441	11.83	0.00	-	22.97	9.20	0.02	0.01	0.80
RbBr:0.2-RbI:0.8	$Fm3m$	IB	IB	1455	4.52E-13	1.16	1.16	6.04	-	0.0198	3.32	5.64	2.57	-17.38	17.15	-	-	-
RbBr:0.4-RbI:0.6	$Fm3m$	IB	IB	1333	2.37E-12	1.16	1.16	6.04	-	0.0206	3.28	5.28	2.54	-17.16	16.77	-	-	-
RbBr:0.6-RbI:0.4	$Fm3m$	IB	IB	1210	1.24E-11	1.16	1.16	6.04	-	0.0214	3.24	4.92	2.52	-16.98	16.33	-	-	-
RbBr:0.8-RbI:0.2	$Fm3m$	IB	IB	1088	6.51E-11	1.16	1.16	6.04	-	0.0222	3.20	4.55	2.49	-16.82	15.87	-	-	-
NaBr:0.2-NaCl:0.8	$Fm3m$	IB	IB	1064	2.08E-11	1.11	1.11	6.02	-	0.0420	2.35	8.20	2.51	-13.72	17.87	-	-	-
NaBr:0.4-NaCl:0.6	$Fm3m$	IB	IB	1055	5.17E-11	1.11	1.11	6.03	-	0.0406	2.54	7.90	2.54	-14.13	18.11	-	-	-
NaBr:0.6-NaCl:0.4	$Fm3m$	IB	IB	1046	1.29E-10	1.12	1.12	6.03	-	0.0393	2.74	7.60	2.56	-14.31	17.87	-	-	-
NaBr:0.8-NaCl:0.2	$Fm3m$	IB	IB	1037	3.21E-10	1.12	1.12	6.03	-	0.0379	2.93	7.30	2.59	-14.48	17.62	-	-	-
NaCl:0.2-KCl:0.8	$Fm3m$	IB	IB	1045	7.63E-12	1.15	1.15	6.03	2.69	0.0332	1.94	5.72	2.74	-15.25	16.46	0.85	0.85	0.19
NaCl:0.4-KCl:0.6	$Fm3m$	IB	IB	1052	7.80E-12	1.14	1.14	6.03	2.62	0.0357	2.00	6.42	2.68	-14.79	16.86	0.85	0.85	0.18
NaCl:0.6-KCl:0.4	$Fm3m$	IB	IB	1059	7.97E-12	1.12	1.12	6.03	2.56	0.0383	2.05	7.11	2.61	-14.31	17.16	0.86	0.86	0.17
NaCl:0.8-KCl:0.2	$Fm3m$	IB	IB	1066	8.15E-12	1.11	1.11	6.02	2.49	0.0408	2.11	7.81	2.55	-13.83	17.52	0.86	0.86	0.15
CaS:0.2-CaSe:0.8	$Fm3m$	IB	IB	1465	4.15E-06	2.38	1.19	6.04	2.24	0.0387	3.51	2.14	5.48	-2.23	5.42	1.45	0.73	0.42
CaS:0.4-CaSe:0.6	$Fm3m$	IB	IB	1777	3.15E-06	2.38	1.19	6.04	2.31	0.0397	3.27	2.20	5.36	-3.27	7.25	1.45	0.73	0.43
CaS:0.6-CaSe:0.4	$Fm3m$	IB	IB	2089	2.39E-06	2.37	1.19	6.04	2.38	0.0408	3.04	2.26	5.23	-3.64	8.26	1.46	0.73	0.43
CaS:0.8-CaSe:0.2	$Fm3m$	IB	IB	2401	1.81E-06	2.37	1.19	6.04	2.45	0.0418	2.80	2.32	5.11	-4.04	8.97	1.46	0.73	0.43
CsCl:0.2-CsBr:0.8	$Fm3m$	IB	IB	911	2.11E-15	1.24	1.24	6.04	-	0.0203	3.43	6.81	3.32	-15.80	18.15	-	-	-
CsCl:0.4-CsBr:0.6	$Fm3m$	IB	IB	913	1.99E-14	1.25	1.25	6.04	-	0.0209	3.37	6.32	3.28	-15.98	17.68	-	-	-
CsCl:0.6-CsBr:0.4	$Fm3m$	IB	IB	915	1.88E-13	1.25	1.25	6.04	-	0.0215	3.30	5.83	3.24	-16.24	17.23	-	-	-
CsCl:0.8-CsBr:0.2	$Fm3m$	IB	IB	917	1.78E-12	1.26	1.26	6.04	-	0.0221	3.24	5.34	3.20	-16.36	16.77	-	-	-
CsBr:0.2-CsI:0.8	$Fm3m$	IB	IB	901	1.05E-14	1.22	1.22	6.04	-	0.0171	3.54	6.46	3.62	-16.48	18.02	-	-	-
CsBr:0.4-CsI:0.6	$Fm3m$	IB	IB	903	4.01E-15	1.23	1.23	6.04	-	0.0178	3.53	6.67	3.55	-16.50	18.33	-	-	-
CsBr:0.6-CsI:0.4	$Fm3m$	IB	IB	905	1.53E-15	1.23	1.23	6.04	-	0.0184	3.51	6.88	3.49	-16.26	18.52	-	-	-
CsBr:0.8-CsI:0.2	$Fm3m$	IB	IB	907	5.84E-16	1.24	1.24	6.04	-	0.0191	3.50	7.09	3.42	-15.86	18.59	-	-	-
Au:0.2-Co:0.8	$Fm3m$	MB	MB	1682	1.97E+05	0.00	0.00	12.02	-	0.0843	10.78	0.00	-	20.10	7.35	0.00	0.00	-
Au:0.4-Co:0.6	$Fm3m$	MB	MB	1596	2.42E+05	0.00	0.00	12.02	-	0.0770	12.59	0.00	-	21.17	6.83	0.00	0.00	-
Au:0.6-Co:0.4	$Fm3m$	MB	MB	1509	2.98E+05	0.00	0.00	12.02	-	0.0697	14.41	0.00	-	22.08	6.38	0.00	0.00	-
Au:0.8-Co:0.2	$Fm3m$	MB	MB	1423	3.66E+05	0.00	0.00	12.02	-	0.0624	16.22	0.00	-	22.68	6.02	0.00	0.00	-
Co:0.2-Ni:0.8	-	MB	MB	1736	1.44E+05	0.00	0.00	12.02	-	0.0925	9.02	0.00	-	18.45	8.53	0.00	0.00	0.62
Co:0.4-Ni:0.6	-	MB	MB	1744	1.48E+05	0.00	0.00	12.02	-	0.0923	9.01	0.00	-	18.74	8.55	0.00	0.00	0.64
Co:0.6-Ni:0.4	-	MB	MB	1752	1.52E+05	0.00	0.00	12.02	-	0.0921	8.99	0.00	-	19.07	8.29	0.00	0.00	0.67
Co:0.8-Ni:0.2	-	MB	MB	1760	1.56E+05	0.00	0.00	12.02	-	0.0918	8.98	0.00	-	19.37	8.43	0.00	0.00	0.69

Table 11.22: Compounds and respective properties used for classification (Part 6). Melting temperature T_{m} in (K), conductivity σ in (S/cm), atomic density ρ_{A} in (10^{-2}/Å^3), density ρ in (g/cm^3), band gap E_{G} in (eV). Properties mostly obtained from the literature[204].

Compound	Structure	Exp.	EMA	T_{m}	σ	Z^*	Z^*_+	ECoN	γ_{TO}	ρ_{A}	ρ	E_{G}	ϵ_∞	t-SNE X	t-SNE Y	TET	ET	ES
InP:0.2-GaP:0.8	$F43m$	CB	CB	1676	2.78E-13	2.28	0.76	4.00	1.25	0.0459	4.12	2.08	11.39	-9.77	3.88	0.82	0.27	1.42
InP:0.4-GaP:0.6	$F43m$	CB	CB	1589	7.75E-12	2.36	0.79	4.00	1.27	0.0439	4.24	1.90	11.72	-9.24	3.22	0.84	0.28	1.42
InP:0.6-GaP:0.4	$F43m$	CB	CB	1502	2.16E-10	2.44	0.81	4.00	1.29	0.0419	4.35	1.71	12.05	-8.20	2.59	0.85	0.28	1.42
InP:0.8-GaP:0.2	$F43m$	CB	CB	1415	6.00E-09	2.52	0.84	4.00	1.31	0.0398	4.47	1.53	12.39	-6.84	1.86	0.87	0.29	1.42
InAs:0.2-GaAs:0.8	$F43m$	CB	CB	1447	5.49E-07	2.31	0.77	4.00	1.23	0.0404	5.10	1.21	14.88	-4.61	2.07	0.69	0.23	1.40
InAs:0.4-GaAs:0.6	$F43m$	CB	CB	1388	3.02E-05	2.42	0.81	4.00	1.26	0.0387	5.15	0.99	15.25	-2.42	0.63	0.70	0.23	1.40
InAs:0.6-GaAs:0.4	$F43m$	CB	CB	1329	1.66E-03	2.52	0.84	4.00	1.28	0.0370	5.19	0.78	15.61	-0.10	-1.09	0.70	0.23	1.40
InAs:0.8-GaAs:0.2	$F43m$	CB	CB	1270	9.10E-02	2.63	0.88	4.00	1.31	0.0353	5.24	0.56	15.98	1.42	-3.84	0.71	0.24	1.40
InP:0.2-InAs:0.8	$F43m$	CB	CB	1235	1.60E-01	2.71	0.90	4.00	1.33	0.0344	5.15	0.55	15.62	1.86	-4.22	0.75	0.25	1.40
InP:0.4-InAs:0.6	$F43m$	CB	CB	1258	5.11E-03	2.69	0.90	4.00	1.33	0.0353	5.01	0.75	14.89	0.54	-2.23	0.78	0.26	1.41
InP:0.6-InAs:0.4	$F43m$	CB	CB	1281	1.63E-04	2.66	0.89	4.00	1.33	0.0361	4.86	0.94	14.17	-1.35	-1.03	0.82	0.27	1.41
InP:0.8-InAs:0.2	$F43m$	CB	CB	1305	5.22E-06	2.63	0.88	4.00	1.33	0.0370	4.72	1.14	13.44	-3.45	0.02	0.85	0.28	1.41
GaP:0.2-GaAs:0.8	$F43m$	CB	CB	1558	6.31E-10	2.20	0.73	4.00	1.21	0.0433	4.84	1.59	13.83	-7.70	3.40	0.71	0.24	1.40
GaP:0.4-GaAs:0.6	$F43m$	CB	CB	1609	3.98E-11	2.20	0.73	4.00	1.22	0.0444	4.63	1.76	13.14	-8.49	3.79	0.73	0.24	1.41
GaP:0.6-GaAs:0.4	$F43m$	CB	CB	1660	2.51E-12	2.20	0.73	4.00	1.22	0.0456	4.43	1.93	12.44	-9.12	4.09	0.76	0.25	1.41
GaP:0.8-GaAs:0.2	$F43m$	CB	CB	1712	1.58E-13	2.20	0.73	4.00	1.23	0.0467	4.22	2.10	11.75	-9.60	4.27	0.78	0.26	1.42
InSb:0.2-GaSb:0.8	$F43m$	CB	CB	941	1.93E+02	2.83	0.94	4.00	1.27	0.0321	5.31	0.65	22.83	5.72	-6.42	0.35	0.12	1.43
InSb:0.4-GaSb:0.6	$F43m$	CB	CB	905	1.38E+02	2.75	0.92	4.00	1.30	0.0309	5.33	0.53	22.46	6.07	-6.23	0.40	0.13	1.42
InSb:0.6-GaSb:0.4	$F43m$	CB	CB	869	9.82E+01	2.66	0.89	4.00	1.34	0.0298	5.34	0.42	22.08	6.41	-6.04	0.45	0.15	1.41
InSb:0.8-GaSb:0.2	$F43m$	CB	CB	834	7.01E+01	2.58	0.86	4.00	1.37	0.0286	5.36	0.30	21.71	6.74	-5.80	0.50	0.17	1.40
InP:0.2-InSb:0.8	$F43m$	CB	CB	904	1.01E+00	2.52	0.84	4.00	1.39	0.0295	5.22	0.41	19.62	5.87	-4.55	0.62	0.21	1.39
InP:0.4-InSb:0.6	$F43m$	CB	CB	1010	2.03E-02	2.54	0.85	4.00	1.38	0.0316	5.06	0.64	17.89	3.98	-3.16	0.68	0.23	1.40
InP:0.6-InSb:0.4	$F43m$	CB	CB	1116	4.10E-04	2.56	0.85	4.00	1.36	0.0337	4.90	0.88	16.17	-0.58	-2.72	0.75	0.25	1.40
InP:0.8-InSb:0.2	$F43m$	CB	CB	1222	8.28E-06	2.58	0.86	4.00	1.34	0.0357	4.74	1.11	14.44	-3.05	-0.98	0.81	0.27	1.41
GaP:0.2-GaSb:0.8	$F43m$	CB	CB	1134	1.40E-01	2.77	0.92	4.00	1.23	0.0362	5.03	1.07	20.77	2.43	-3.15	0.40	0.13	1.43
GaP:0.4-GaSb:0.6	$F43m$	CB	CB	1291	7.22E-05	2.63	0.88	4.00	1.23	0.0391	4.78	1.37	18.34	-2.50	-0.49	0.50	0.17	1.43
GaP:0.6-GaSb:0.4	$F43m$	CB	CB	1448	3.74E-08	2.48	0.83	4.00	1.23	0.0421	4.52	1.67	15.92	-6.50	2.06	0.61	0.20	1.43
GaP:0.8-GaSb:0.2	$F43m$	CB	CB	1606	1.93E-11	2.34	0.78	4.00	1.23	0.0450	4.27	1.97	13.49	-8.98	3.56	0.71	0.24	1.43
InAs:0.2-InSb:0.8	$F43m$	CB	CB	881	3.15E+01	2.55	0.85	4.00	1.39	0.0286	5.36	0.21	20.34	6.54	-5.23	0.58	0.19	1.39
InAs:0.4-InSb:0.6	$F43m$	CB	CB	963	1.99E+01	2.60	0.87	4.00	1.38	0.0299	5.34	0.25	19.34	5.86	-5.43	0.62	0.21	1.39
InAs:0.6-InSb:0.4	$F43m$	CB	CB	1046	1.26E+01	2.64	0.88	4.00	1.36	0.0311	5.33	0.28	18.34	4.87	-5.83	0.65	0.22	1.40
InAs:0.8-InSb:0.2	$F43m$	CB	CB	1129	7.92E+00	2.69	0.90	4.00	1.35	0.0324	5.31	0.32	17.34	4.04	-5.81	0.69	0.23	1.40
GaAs:0.2-GaSb:0.8	$F43m$	CB	CB	1082	2.21E+00	2.77	0.92	4.00	1.23	0.0351	5.24	0.90	21.46	3.58	-4.55	0.38	0.13	1.43
GaAs:0.4-GaSb:0.6	$F43m$	CB	CB	1188	1.81E-02	2.63	0.88	4.00	1.22	0.0368	5.19	1.03	19.73	1.35	-2.52	0.45	0.15	1.42
GaAs:0.6-GaSb:0.4	$F43m$	CB	CB	1294	1.49E-04	2.48	0.83	4.00	1.22	0.0386	5.15	1.16	17.99	-1.75	-0.53	0.53	0.18	1.41
GaAs:0.8-GaSb:0.2	$F43m$	CB	CB	1400	1.22E-06	2.34	0.78	4.00	1.21	0.0403	5.10	1.29	16.26	-4.31	1.56	0.60	0.20	1.41

Table 11.23: Compounds and respective properties used for classification (Part 7). Melting temperature T_{m} in (K), conductivity σ in (S/cm), atomic density ρ_{A} in (10^{-2}/Å^3), density ρ in (g/cm^3), band gap E_{G} in (eV). Properties mostly obtained from the literature[204].

Compound	Structure	Exp.	EMA	T_{m}	σ	Z^*	Z^*_+	ECoN	γ_{TO}	ρ_{A}	ρ	E_{G}	ϵ_∞	t-SNE X	t-SNE Y	TET	ET	ES
(InP:0.2-GaP:0.8):0.2 (InAs:0.2-GaAs:0.8):0.8	$F43m$	CB	CB	1493	3.03E-08	2.30	0.77	4.00	1.24	0.0415	4.90	1.38	14.19	-6.16	2.69	0.72	0.24	1.40
(InP:0.2-GaP:0.8):0.4 (InAs:0.2-GaAs:0.8):0.6	$F43m$	CB	CB	1539	1.67E-09	2.30	0.77	4.00	1.24	0.0426	4.71	1.56	13.49	-7.46	3.08	0.74	0.25	1.41
(InP:0.2-GaP:0.8):0.6 (InAs:0.2-GaAs:0.8):0.4	$F43m$	CB	CB	1585	9.18E-11	2.29	0.76	4.00	1.24	0.0437	4.51	1.73	12.79	-8.40	3.44	0.77	0.26	1.41
(InP:0.2-GaP:0.8):0.8 (InAs:0.2-GaAs:0.8):0.2	$F43m$	CB	CB	1630	5.05E-12	2.29	0.76	4.00	1.25	0.0448	4.32	1.91	12.09	-9.36	3.45	0.80	0.27	1.42
(InP:0.4-GaP:0.6):0.2 (InAs:0.4-GaAs:0.6):0.8	$F43m$	CB	CB	1428	1.45E-06	2.41	0.80	4.00	1.26	0.0397	4.96	1.17	14.54	-4.11	1.60	0.72	0.24	1.40
(InP:0.4-GaP:0.6):0.4 (InAs:0.4-GaAs:0.6):0.6	$F43m$	CB	CB	1469	6.97E-08	2.39	0.80	4.00	1.26	0.0408	4.78	1.35	13.84	-5.76	2.38	0.75	0.25	1.41
(InP:0.4-GaP:0.6):0.6 (InAs:0.4-GaAs:0.6):0.4	$F43m$	CB	CB	1509	3.35E-09	2.38	0.79	4.00	1.26	0.0418	4.60	1.54	13.13	-7.22	2.78	0.78	0.26	1.41
(InP:0.4-GaP:0.6):0.8 (InAs:0.4-GaAs:0.6):0.2	$F43m$	CB	CB	1549	1.61E-10	2.37	0.79	4.00	1.27	0.0428	4.42	1.72	12.43	-8.35	3.01	0.81	0.27	1.42
(InP:0.6-GaP:0.4):0.2 (InAs:0.6-GaAs:0.4):0.8	$F43m$	CB	CB	1364	6.95E-05	2.51	0.84	4.00	1.28	0.0380	5.03	0.96	14.90	-1.89	0.13	0.73	0.24	1.40
(InP:0.6-GaP:0.4):0.4 (InAs:0.6-GaAs:0.4):0.6	$F43m$	CB	CB	1398	2.92E-06	2.49	0.83	4.00	1.28	0.0389	4.86	1.15	14.19	-3.70	1.14	0.76	0.25	1.41
(InP:0.6-GaP:0.4):0.6 (InAs:0.6-GaAs:0.4):0.4	$F43m$	CB	CB	1433	1.22E-07	2.48	0.83	4.00	1.29	0.0399	4.69	1.34	13.48	-5.42	2.05	0.79	0.26	1.41
(InP:0.6-GaP:0.4):0.8 (InAs:0.6-GaAs:0.4):0.2	$F43m$	CB	CB	1468	5.14E-09	2.46	0.82	4.00	1.29	0.0409	4.52	1.53	12.77	-7.00	2.39	0.82	0.27	1.42
(InP:0.8-GaP:0.2):0.2 (InAs:0.8-GaAs:0.2):0.8	$F43m$	CB	CB	1299	3.33E-03	2.61	0.87	4.00	1.31	0.0362	5.09	0.76	15.26	0.27	-1.72	0.74	0.25	1.40
(InP:0.8-GaP:0.2):0.4 (InAs:0.8-GaAs:0.2):0.6	$F43m$	CB	CB	1328	1.22E-04	2.59	0.86	4.00	1.31	0.0371	4.93	0.95	14.54	-1.48	-0.37	0.77	0.26	1.41
(InP:0.8-GaP:0.2):0.6 (InAs:0.8-GaAs:0.2):0.4	$F43m$	CB	CB	1357	4.47E-06	2.57	0.86	4.00	1.31	0.0380	4.78	1.14	13.82	-3.49	0.62	0.80	0.27	1.41
(InP:0.8-GaP:0.2):0.8 (InAs:0.8-GaAs:0.2):0.2	$F43m$	CB	CB	1386	1.64E-07	2.54	0.85	4.00	1.31	0.0389	4.62	1.33	13.10	-5.22	1.46	0.84	0.28	1.41
(InP:0.2-GaP:0.8):0.2 (InSb:0.2-GaSb:0.8):0.8	$F43m$	CB	CB	1088	2.07E-01	2.72	0.91	4.00	1.26	0.0349	5.07	0.94	20.54	3.11	-3.57	0.44	0.15	1.43
(InP:0.2-GaP:0.8):0.4 (InSb:0.2-GaSb:0.8):0.6	$F43m$	CB	CB	1235	2.23E-04	2.61	0.87	4.00	1.26	0.0376	4.83	1.22	18.25	-1.75	-1.28	0.54	0.18	1.42
(InP:0.2-GaP:0.8):0.6 (InSb:0.2-GaSb:0.8):0.4	$F43m$	CB	CB	1382	2.40E-07	2.50	0.83	4.00	1.26	0.0404	4.60	1.51	15.97	-5.53	1.42	0.63	0.21	1.42
(InP:0.2-GaP:0.8):0.8 (InSb:0.2-GaSb:0.8):0.2	$F43m$	CB	CB	1529	2.59E-10	2.39	0.80	4.00	1.25	0.0431	4.36	1.80	13.68	-8.38	2.76	0.73	0.24	1.42
(InP:0.4-GaP:0.6):0.2 (InSb:0.4-GaSb:0.6):0.8	$F43m$	CB	CB	1042	3.08E-01	2.67	0.89	4.00	1.30	0.0335	5.11	0.81	20.31	3.84	-3.91	0.49	0.16	1.42
(InP:0.4-GaP:0.6):0.4 (InSb:0.4-GaSb:0.6):0.6	$F43m$	CB	CB	1179	6.90E-04	2.59	0.86	4.00	1.29	0.0361	4.89	1.08	18.16	-0.83	-2.06	0.58	0.19	1.42
(InP:0.4-GaP:0.6):0.6 (InSb:0.4-GaSb:0.6):0.4	$F43m$	CB	CB	1315	1.54E-06	2.52	0.84	4.00	1.28	0.0387	4.67	1.35	16.02	-4.28	0.54	0.66	0.22	1.42
(InP:0.4-GaP:0.6):0.8 (InSb:0.4-GaSb:0.6):0.2	$F43m$	CB	CB	1452	3.46E-09	2.44	0.81	4.00	1.28	0.0413	4.46	1.63	13.87	-7.37	2.18	0.75	0.25	1.42
(InP:0.6-GaP:0.4):0.2 (InSb:0.6-GaSb:0.4):0.8	$F43m$	CB	CB	996	4.57E-01	2.62	0.87	4.00	1.33	0.0322	5.15	0.68	20.08	4.59	-4.12	0.53	0.18	1.41
(InP:0.6-GaP:0.4):0.4 (InSb:0.6-GaSb:0.4):0.6	$F43m$	CB	CB	1122	2.13E-03	2.58	0.86	4.00	1.32	0.0346	4.95	0.93	18.07	0.16	-2.81	0.61	0.20	1.41
(InP:0.6-GaP:0.4):0.6 (InSb:0.6-GaSb:0.4):0.4	$F43m$	CB	CB	1249	9.93E-06	2.53	0.84	4.00	1.31	0.0370	4.75	1.19	16.07	-3.17	-0.52	0.69	0.23	1.41
(InP:0.6-GaP:0.4):0.8 (InSb:0.6-GaSb:0.4):0.2	$F43m$	CB	CB	1376	4.63E-08	2.49	0.83	4.00	1.30	0.0394	4.55	1.45	14.06	-5.98	1.51	0.77	0.26	1.42

Table 11.24: Compounds and respective properties used for classification (Part 8). Melting temperature T_{m} in (K), conductivity σ in (S/cm), atomic density ρ_{A} in (10^{-2}/Å^3), density ρ in (g/cm^3), band gap E_{G} in (eV). Properties mostly obtained from the literature[204].

Compound	Structure	Exp.	EMA	T_{m}	σ	Z^*	Z^*_+	ECoN	γ_{TO}	ρ_{A}	ρ	E_{G}	ϵ_∞	t-SNE X	t-SNE Y	TET	ET	ES
(InP:0.8-GaP:0.2):0.2 (InSb:0.8-GaSb:0.2):0.8	$F43m$	CB	CB	950	6.79E-01	2.57	0.86	4.00	1.36	0.0308	5.18	0.54	19.85	5.26	-4.39	0.57	0.19	1.40
(InP:0.8-GaP:0.2):0.4 (InSb:0.8-GaSb:0.2):0.6	$F43m$	CB	CB	1066	6.58E-03	2.56	0.85	4.00	1.35	0.0331	5.00	0.79	17.98	2.71	-2.67	0.65	0.22	1.41
(InP:0.8-GaP:0.2):0.6 (InSb:0.8-GaSb:0.2):0.4	$F43m$	CB	CB	1183	6.38E-05	2.55	0.85	4.00	1.33	0.0353	4.82	1.03	16.12	-1.93	-1.74	0.72	0.24	1.41
(InP:0.8-GaP:0.2):0.8 (InSb:0.8-GaSb:0.2):0.2	$F43m$	CB	CB	1299	6.19E-07	2.53	0.84	4.00	1.32	0.0376	4.65	1.28	14.25	-4.53	0.43	0.79	0.26	1.41
(InAs:0.2-GaAs:0.8):0.2 (InSb:0.2-GaSb:0.8):0.8	$F43m$	CB	CB	1042	3.77E+00	2.72	0.91	4.00	1.26	0.0338	5.27	0.76	21.24	4.26	-4.90	0.42	0.14	1.42
(InAs:0.2-GaAs:0.8):0.4 (InSb:0.2-GaSb:0.8):0.6	$F43m$	CB	CB	1143	7.36E-02	2.62	0.87	4.00	1.25	0.0354	5.22	0.87	19.65	2.22	-3.39	0.49	0.16	1.42
(InAs:0.2-GaAs:0.8):0.6 (InSb:0.2-GaSb:0.8):0.4	$F43m$	CB	CB	1245	1.44E-03	2.52	0.84	4.00	1.25	0.0371	5.18	0.98	18.06	-0.28	-1.63	0.55	0.18	1.41
(InAs:0.2-GaAs:0.8):0.8 (InSb:0.2-GaSb:0.8):0.2	$F43m$	CB	CB	1346	2.81E-05	2.41	0.80	4.00	1.24	0.0387	5.14	1.10	16.47	-2.51	0.30	0.62	0.21	1.41
(InAs:0.4-GaAs:0.6):0.2 (InSb:0.4-GaSb:0.6):0.8	$F43m$	CB	CB	1002	6.41E+00	2.68	0.89	4.00	1.29	0.0325	5.29	0.63	21.01	4.83	-5.26	0.46	0.15	1.41
(InAs:0.4-GaAs:0.6):0.4 (InSb:0.4-GaSb:0.6):0.6	$F43m$	CB	CB	1098	2.98E-01	2.61	0.87	4.00	1.28	0.0340	5.25	0.72	19.57	3.28	-4.14	0.52	0.17	1.41
(InAs:0.4-GaAs:0.6):0.6 (InSb:0.4-GaSb:0.6):0.4	$F43m$	CB	CB	1195	1.39E-02	2.55	0.85	4.00	1.28	0.0356	5.22	0.81	18.13	1.24	-2.84	0.58	0.19	1.41
(InAs:0.4-GaAs:0.6):0.8 (InSb:0.4-GaSb:0.6):0.2	$F43m$	CB	CB	1292	6.48E-04	2.48	0.83	4.00	1.27	0.0371	5.18	0.90	16.69	-0.62	-1.03	0.64	0.21	1.40
(InAs:0.6-GaAs:0.4):0.2 (InSb:0.6-GaSb:0.4):0.8	$F43m$	CB	CB	961	1.09E+01	2.64	0.88	4.00	1.33	0.0312	5.31	0.49	20.79	5.42	-5.49	0.50	0.17	1.41
(InAs:0.6-GaAs:0.4):0.4 (InSb:0.6-GaSb:0.4):0.6	$F43m$	CB	CB	1053	1.21E+00	2.61	0.87	4.00	1.32	0.0327	5.28	0.56	19.50	4.30	-4.68	0.55	0.18	1.40
(InAs:0.6-GaAs:0.4):0.6 (InSb:0.6-GaSb:0.4):0.4	$F43m$	CB	CB	1145	1.34E-01	2.58	0.86	4.00	1.30	0.0341	5.25	0.63	18.20	2.60	-4.07	0.60	0.20	1.40
(InAs:0.6-GaAs:0.4):0.8 (InSb:0.6-GaSb:0.4):0.2	$F43m$	CB	CB	1237	1.49E-02	2.55	0.85	4.00	1.29	0.0356	5.22	0.71	16.91	1.01	-2.92	0.65	0.22	1.40
(InAs:0.8-GaAs:0.2):0.2 (InSb:0.8-GaSb:0.2):0.8	$F43m$	CB	CB	921	1.85E+01	2.59	0.86	4.00	1.36	0.0299	5.34	0.35	20.56	6.05	-5.29	0.54	0.18	1.40
(InAs:0.8-GaAs:0.2):0.4 (InSb:0.8-GaSb:0.2):0.6	$F43m$	CB	CB	1008	4.91E+00	2.60	0.87	4.00	1.35	0.0313	5.31	0.40	19.42	5.12	-5.14	0.58	0.19	1.40
(InAs:0.8-GaAs:0.2):0.6 (InSb:0.8-GaSb:0.2):0.4	$F43m$	CB	CB	1096	1.30E+00	2.61	0.87	4.00	1.33	0.0326	5.29	0.46	18.27	3.84	-5.13	0.63	0.21	1.40
(InAs:0.8-GaAs:0.2):0.8 (InSb:0.8-GaSb:0.2):0.2	$F43m$	CB	CB	1183	3.44E-01	2.62	0.87	4.00	1.32	0.0340	5.27	0.51	17.12	2.57	-4.67	0.67	0.22	1.40

Chapter 12

Data Preservation and Compute Time Acknowledgements

Data Preservation

> Directive 01 is to ensure all data acquired is preserved and passed along.
>
> *IF Prototype LQ-84i (Wolf)*

All calculation results ($\approx$ 7700) are stored on the servers of the 1. Institute of Physics (IA), RWTH Aachen University. Backups of the data are copied to servers of the RWTH IT Center in regular intervals.
Project DIO, including *JoJo* (see chapter 7.5), is stored on the RWTH Gitlab at:

git.rwth-aachen.de/carl-friedrich.schoen/project_dio

The interactive map (see chapter 10) is available at:

git.rwth-aachen.de/carl-friedrich.schoen/interactive-map

As both projects contain commercial code for which a corresponding license is required, access must first be granted by contacting the 1. Institute of Physics (IA), RWTH Aachen University.

Compute Time Acknowledgements

The author of this thesis gratefully acknowledges the computational ressources provided by the RWTH Aachen and the Forschungszentrum Jülich. The following compute projects have been utilized:

- rwth0508, rwth0662, p0020115, p0020357, jara0229, jara0236

The author furthermore wants to thank the staff of the RWTH IT Center for their continuous technical support.

List of Publications

This chapter contains the full list of publications in reverse chronological order to which the author of this thesis contributed:

1. **Nonresonant optomechanical phase control**
   Jiaojian Shi, Yijing Huang, Christian Heide, Burak Guzelturk, Carl-Friedrich Schön, Haowei Xu, Yuki Kobayashi, Andrew F. May, Pooja Donthi Reddy, Duan Luo, Eamonn Hughes, Kunal Mukherjee, Mariano Trigo, Ju Li, Jian Zhou, Shambhu Ghimire, Matthias Wuttig, David A. Reis, Aaron M. Lindenberg
   In submission (Reference [91])

2. **Quantum Materials that defy Zachariasen's Conjecture about Structure and Bonding in Glasses**
   Jean-Yves Raty, Christophe Bichara, Carl-Friedrich Schön, Carlo Gatti, Matthias Wuttig
   In submission (Reference [205])

3. **Doping by Design: Enhanced Thermoelectric Performance of GeSe Alloys Through Metavalent Bonding**
   Yuan Yu, Chongjian Zhou, Tanmoy Ghosh, Carl-Friedrich Schön, Yiming Zhou, Sophia Wahl, Mohit Raghuwanshi, Peter Kerres, Christophe Bellin, Abhay Shukla, Oana Cojocaru-Mirédin, Matthias Wuttig
   Advanced Materials, Volume 35, Issue 19, 2023 (Reference [95])

4. **Metavalent Bonding in Layered Phase-Change Memory Materials**
   Wei Zhang, Hangming Zhang, Suyang Sun, Xiaozhe Wang, Zhewen Lu, Xudong Wang, Jiang-Jing Wang, Chunlin Jia, Carl-Friedrich Schön, Riccardo Mazzarello, En Ma, Matthias Wuttig. Metavalent Bonding in Layered Phase-Change Memory Materials
   Advanced Science, Volume 10, Issue 15, 2023 (Reference [206])

5. **Revisiting the Nature of Chemical Bonding in Chalcogenides to Explain and Design their Properties**
   Matthias Wuttig, Carl-Friedrich Schön, Jakob Lötfering, Pavlo Golub, Carlo Gatti, Jean-Yves Raty
   Advanced Materials, Volume 35, Issue 20, 2023 (Reference [3])

6. **Classification of Properties and their Relation to Chemical Bonding: Essential Steps toward the Inverse Design of Functional Materials**
   Carl-Friedrich Schön, Steffen van Bergerem, Christian Mattes, Aakash Yadav, Martin Grohe, Leif Kobbelt, Matthias Wuttig
   Science Advances, Volume 8, Issue 47, 2022 (Reference [2])

7. **Scaling and Confinement in Ultrathin Chalcogenide Films as Exemplified by GeTe**
   Peter Kerres, Yiming Zhou, Hetal Vaishnav, Mohit Raghuwanshi, Jiangjing Wang, Maria Häser, Marc Pohlmann, Yudong Cheng, Carl-Friedrich Schön, Thomas Jansen, Christophe Bellin, Daniel E Bürgler, Abdur Rehman Jalil, Christoph Ringkamp, Hugo Kowalczyk, Claus M Schneider, Abhay Shukla, Matthias Wuttig
   small, Volume 18, Issue 21, 2022 (Reference [82])

8. **Halide Perovskites: Advanced Photovoltaic Materials Empowered by a Unique Bonding Mechanism**
   Matthias Wuttig, Carl-Friedrich Schön, Mathias Schumacher, John Robertson, Pavlo Golub, Eric Bousquet, Carlo Gatti, Jean-Yves Raty
   Advanced Functional Materials, Volume 32, Issue 2, 2022 (Reference [47])

9. **Metavalent Bonding in Crystalline Solids: How does it collapse?**
   Ludovica Guarneri, Stefan Jakobs, Alexander von Hoegen, Stefan Maier, Ming Xu, Min Zhu, Sophia Wahl, Christian Teichrib, Yiming Zhou, Oana Cojocaru-Mirédin, Mohit Raghuwanshi, Carl-Friedrich Schön, Marc Drögeler, Christoph Stampfer, Ricardo P S M Lobo, Andrea Piarristeguy, Annie Pradel, Jean-Yves Raty, Matthias Wuttig
   Advanced Materials, Volume 33, Issue 39, 2021 (Reference [23])

10. **How to identify Lone Pairs, van der Waals gaps and Metavalent bonding using charge and pair density methods: From Elemental Chalcogens to Lead Chalcogenides and Phase Change Materials**
    Jean-Yves Raty, Carlo Gatti, Carl-Friedrich Schön, Matthias Wuttig
    Physica Status Solidi – RRL, Volume 15, Issue 11, 2021 (Reference [158])

11. **In_3SbTe_2 as a programmable nanophotonics material platform for the infrared**
    Andreas Heßler, Sophia Wahl, Till Leuteritz, Antonios Antonopoulos, Christina Stergianou, Carl-Friedrich Schön, Lukas Naumann, Niklas Eicker, Martin Lewin, Tobias W W Maß, Matthias Wuttig, Stefan Linden, Thomas Taubner
    Nature Communications, Volume 12, Issue 1, 2021 (Reference [83])

12. **Discovering Electron-Transfer-Driven Changes in Chemical Bonding in Lead Chalcogenides**
    Stefan Maier, Simon Steinberg, Yudong Cheng, Carl-Friedrich Schön, Mathias Schumacher, Riccardo Mazzarello, Pavlo Golub, Ryky Nelson, Oana Cojocaru-Mirédin, Jean-Yves Raty, Matthias Wuttig
    Advanced Materials, Volume 32, Issue 49, 2020 (Reference [24])

List of Figures

Bibliography

[1] William Shockley and Hans J Queisser. Detailed Balance Limit of Efficiency of p-n Junction Solar Cells. *Journal of Applied Physics*, 32(3):510–519, mar 1961. Cited on page 3.

[2] Carl-Friedrich Schön, Steffen van Bergerem, Christian Mattes, Aakash Yadav, Martin Grohe, Leif Kobbelt, and Matthias Wuttig. Classification of properties and their relation to chemical bonding: Essential steps toward the inverse design of functional materials. *Science Advances*, 8(47):eade0828, nov 2022. Cited on pages 4, 6, 8, 14, 16, 17, 20, 22, 24, 25, 26 and 201.

[3] Matthias Wuttig, Carl-Friedrich Schön, Jakob Lötfering, Pavlo Golub, Carlo Gatti, and Jean-Yves Raty. Revisiting the Nature of Chemical Bonding in Chalcogenides to Explain and Design their Properties. *Advanced Materials*, 35(20):2208485, may 2023. Cited on pages 4, 13, 15, 29, 31, 32, 34, 35, 36, 37, 39, 40, 41, 43, 45, 46, 47, 111, 166, 181, 184 and 201.

[4] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer New York, 2006. Cited on page 6.

[5] Matthias Wuttig, Volker L Deringer, Xavier Gonze, Christophe Bichara, and Jean-Yves Raty. Incipient Metals: Functional Materials with a Unique Bonding Mechanism. *Advanced Materials*, 30(51):1803777, dec 2018. Cited on pages 7, 12 and 56.

[6] L Pauling, Cornell University, and Cornell University Press. *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry*. George Fisher Baker Non-Resident Lecture Series. Cornell University Press, 1960. Cited on pages 10, 14 and 30.

[7] Gilbert N Lewis. THE ATOM AND THE MOLECULE. *Journal of the American Chemical Society*, 38(4):762–785, apr 1916. Cited on page 10.

[8] Frank Rioux. The Covalent Bond Clarified Through the Use of the Virial Theorem. In *Quantum Tutorials*. LibreTexts (College of Saint Benedict/Saint John's University), 2023. Cited on page 11.

[9] Klaus Ruedenberg. The Physical Nature of the Chemical Bond. *Reviews of Modern Physics*, 34(2):326–376, apr 1962. Cited on page 11.

[10] Bart J Kooi and Matthias Wuttig. Chalcogenides by Design: Functionality through Metavalent Bonding and Confinement. *Advanced Materials*, 32(21):1908302, may 2020. Cited on pages 12, 24, 27, 67, 165 and 166.

[11] Irving Langmuir. THE ARRANGEMENT OF ELECTRONS IN ATOMS AND MOLECULES. *Journal of the American Chemical Society*, 41(6):868–934, jun 1919. Cited on page 14.

[12] Laurens Van Der Maaten and Geoffrey Hinton. Visualizing Data using t-SNE. *Journal of Machine Learning Research*, 9:2579–2605, 2008. Cited on page 21.

[13] Geoffrey Hinton and Sam Roweis. Stochastic Neighbor Embedding, 2023. Cited on page 21.

[14] Jean-Yves Raty, Mathias Schumacher, Pavlo Golub, Volker L Deringer, Carlo Gatti, and Matthias Wuttig. A Quantum-Mechanical Map for Bonding and Properties in Solids. *Advanced Materials*, 31(3):1806280, jan 2019. Cited on page 24.

[15] Richard F. Bader. Molecular Fragments or Chemical Bonds ? *Acc. Chem. Res.*, 5(1):34–40, 1974. Cited on pages 24, 109 and 110.

[16] R F W Bader. Atoms in Molecules. *Acc. Chem. Res.*, pages 9–15, 1985.

[17] Richard F. Bader. *Atoms in Molecules: A Quantum Theory*. Clarendon Press, 1990. Cited on pages 30 and 110.

[18] Richard F W Bader. A Quantum Theory of Molecular Structure and Its Applications. *Chem. Rev.*, 91:893–928, 1991. Cited on pages 24 and 109.

[19] Matthias Wuttig and Noboru Yamada. Phase-change materials for rewriteable data storage. *Nature Materials*, 6(11):824–832, 2007. Cited on page 26.

[20] Jian Zhou, Shunhong Zhang, and Ju Li. Normal-to-topological insulator martensitic phase transition in group-IV monochalcogenides driven by light. *NPG Asia Materials*, 12(1):2, 2020. Cited on page 71.

[21] Simone Raoux, Wojciech Wełnic, and Daniele Ielmini. Phase Change Materials and Their Application to Nonvolatile Memories. *Chemical Reviews*, 110(1):240–267, jan 2010.

[22] Xiao-Liang Qi and Shou-Cheng Zhang. Topological insulators and superconductors. *Reviews of Modern Physics*, 83(4):1057–1110, oct 2011. Cited on page 26.

[23] Ludovica Guarneri, Stefan Jakobs, Alexander von Hoegen, Stefan Maier, Ming Xu, Min Zhu, Sophia Wahl, Christian Teichrib, Yiming Zhou, Oana Cojocaru-Mirédin, Mohit Raghuwanshi, Carl-Friedrich Schön, Marc Drögeler, Christoph Stampfer, Ricardo P S M Lobo, Andrea Piarristeguy, Annie Pradel, Jean-Yves Raty, and Matthias Wuttig. Metavalent Bonding in Crystalline Solids: How Does It Collapse? *Advanced Materials*, 33(39):2102356, oct 2021. Cited on pages 27, 38, 39 and 202.

[24] Stefan Maier, Simon Steinberg, Yudong Cheng, Carl-Friedrich Schön, Mathias Schumacher, Riccardo Mazzarello, Pavlo Golub, Ryky Nelson, Oana Cojocaru-Mirédin, Jean-Yves Raty, and Matthias Wuttig. Discovering Electron-Transfer-Driven Changes in Chemical Bonding in Lead Chalcogenides (PbX, where X = Te, Se, S, O). *Advanced Materials*, 32(49):2005533, dec 2020. Cited on pages 27, 43, 46, 69, 70, 173, 174, 181 and 202.

[25] John C Slater. Note on Molecular Structure. *Physical Review*, 41(2):255–257, jul 1932. Cited on page 29.

[26] J H Van Vleck and Albert Sherman. The Quantum Theory of Valence. *Reviews of Modern Physics*, 7(3):167–228, jul 1935.

[27] Sason Shaik and Philippe Hiberty, C. *Reviews in Computational Chemistry*. John Wiley & Sons, Ltd, 2004. Cited on page 29.

[28] Kostas Gavroglu and Ana Simões. *Neither Physics nor Chemistry*. The MIT Press, 2011. Cited on page 30.

[29] Roald Hoffmann. *Solids and Surfaces: A Chemist's View of Bonding in Extended Structures*. Wiley-VCH, 1991. Cited on page 30.

[30] Philip Ball. Beyond the bond. *Nature*, 469(7328):26–28, 2011. Cited on page 30.

[31] Tae Hoon Lee and Stephen R Elliott. Chemical Bonding in Chalcogenides: The Concept of Multicenter Hyperbonding. *Advanced Materials*, 32(28):2000340, jul 2020. Cited on page 31.

[32] Jan Hempelmann, Peter C Müller, Christina Ertural, and Richard Dronskowski. The Orbital Origins of Chemical Bonding in Ge-Sb-Te Phase-Change Materials**. *Angewandte Chemie International Edition*, 61(17):e202115778, apr 2022. Cited on page 31.

[33] J.D. Lee. *Concise Inorganic Chemistry*. Blackwell Publishers, 5 edition, 1999. Cited on page 31.

[34] J. McMurry. *Organic Chemistry 5th Ed.* Ceneage Learning Indi Pvt Limited, 2000.

[35] M.A. Fox and J.K. Whitesell. *Organic Chemistry.* Jones and Bartlett Publishers, 2004. Cited on page 31.

[36] K B Wiberg. Application of the pople-santry-segal CNDO method to the cyclopropylcarbinyl and cyclobutyl cation and to bicyclobutane. *Tetrahedron*, 24(3):1083–1096, 1968. Cited on page 34.

[37] I Mayer. Charge, bond order and valence in the AB initio SCF theory. *Chemical Physics Letters*, 97(3):270–274, 1983. Cited on page 34.

[38] Peter C Müller, Christina Ertural, Jan Hempelmann, and Richard Dronskowski. Crystal Orbital Bond Index: Covalent Bond Orders in Solids. *The Journal of Physical Chemistry C*, 125(14):7959–7970, apr 2021. Cited on page 34.

[39] Stefan Maintz, Volker L Deringer, Andrei L Tchougréeff, and Richard Dronskowski. LOBSTER: A tool to extract chemical bonding from plane-wave based DFT., apr 2016. Cited on pages 34 and 150.

[40] Enrico Fermi. *Nuclear Physics.* Univ. of Chicago Press, 1974. Cited on page 44.

[41] Peter B Littlewood. The crystal structure of IV-VI compounds. II. A microscopic model for cubic/rhombohedral materials. *Journal of Physics C: Solid State Physics*, 13:4875–4892, 1980. Cited on page 45.

[42] Matthias Wuttig and Noboru Yamada. Phase-change materials for rewriteable data storage. *Nature Materials*, 2007. Cited on page 46.

[43] T. Siegrist, P. Merkelbach, and M. Wuttig. Challenges on the Path to a Universal Storage Device. *Annual Review of Condensed Matter Physics*, 2012. Cited on page 46.

[44] G Bruns, P Merkelbach, C Schlockermann, M Salinga, M Wuttig, T D Happ, J B Philipp, and M Kund. Nanosecond switching in GeTe phase change memory cells. *Applied Physics Letters*, 95(4):43108, jul 2009. Cited on page 46.

[45] D Loke, T H Lee, W J Wang, L P Shi, R Zhao, Y C Yeo, T C Chong, and S R Elliott. Breaking the speed limits of phase-change memory. *Science (New York, N.Y.)*, 336(6088):1566–1569, jun 2012. Cited on page 46.

[46] Christoph Persch, Maximilian J Müller, Aakash Yadav, Julian Pries, Natalie Honné, Peter Kerres, Shuai Wei, Hajime Tanaka, Paolo Fantini, Enrico Varesi, Fabio Pellizzer, and Matthias Wuttig. The potential of chemical bonding to design crystallization and vitrification kinetics. *Nature communications*, 12(1):4978, aug 2021. Cited on page 47.

[47] Matthias Wuttig, Carl-Friedrich Schön, Mathias Schumacher, John Robertson, Pavlo Golub, Eric Bousquet, Carlo Gatti, and Jean-Yves Raty. Halide Perovskites: Advanced Photovoltaic Materials Empowered by a Unique Bonding Mechanism. *Advanced Functional Materials*, 32(2):2110166, jan 2022. Cited on pages 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 77, 185, 186 and 202.

[48] Kenji Uchino. Glory of piezoelectric perovskites. *Science and Technology of Advanced Materials*, 16(4):46001, jul 2015. Cited on page 48.

[49] Nadege Courjal, Maria-Pilar Bernal, Alexis Caspar, Gwenn Ulliac, Florent Bassignot, Ludovic Gauthier-Manuel, and Miguel Suarez. Lithium Niobate Optical Waveguides and Microwaveguides. In *Emerging Waveguide Technology.* IntechOpen.com, aug 2018. Cited on page 48.

[50] S Pamir Alpay, Joseph Mantese, Susan Trolier-McKinstry, Qiming Zhang, and Roger W Whatmore. Next-generation electrocaloric and pyroelectric materials for solid-state electrothermal energy interconversion. *MRS Bulletin*, 39(12):1099–1111, 2014. Cited on page 48.

[51] D B Mitzi, S Wang, C A Feild, C A Chess, and A M Guloy. Conducting Layered Organic-inorganic Halides Containing <110>-Oriented Perovskite Sheets. *Science (New York, N.Y.)*, 267(5203):1473–1476, mar 1995. Cited on page 48.

[52] D B Mitzi, K Chondroudis, and C R Kagan. Organic-inorganic electronics. *IBM Journal of Research and Development*, 45(1):29–45, 2001.

[53] Akihiro Kojima, Kenjiro Teshima, Yasuo Shirai, and Tsutomu Miyasaka. Organometal Halide Perovskites as Visible-Light Sensitizers for Photovoltaic Cells. *Journal of the American Chemical Society*, 131(17):6050–6051, may 2009.

[54] Hui-Seon Kim, Chang-Ryul Lee, Jeong-Hyeok Im, Ki-Beom Lee, Thomas Moehl, Arianna Marchioro, Soo-Jin Moon, Robin Humphry-Baker, Jun-Ho Yum, Jacques E Moser, Michael Grätzel, and Nam-Gyu Park. Lead Iodide Perovskite Sensitized All-Solid-State Submicron Thin Film Mesoscopic Solar Cell with Efficiency Exceeding 9%. *Scientific Reports*, 2(1):591, 2012.

[55] Mingzhen Liu, Michael B Johnston, and Henry J Snaith. Efficient planar heterojunction perovskite solar cells by vapour deposition. *Nature*, 501(7467):395–398, sep 2013.

[56] Martin A Green, Anita Ho-Baillie, and Henry J Snaith. The emergence of perovskite solar cells. *Nature Photonics*, 8(7):506–514, 2014.

[57] Michael Kulbak, David Cahen, and Gary Hodes. How Important Is the Organic Part of Lead Halide Perovskite Photovoltaic Cells? Efficient CsPbBr3 Cells. *The Journal of Physical Chemistry Letters*, 6(13):2452–2456, jul 2015.

[58] Henry J Snaith. Present status and future prospects of perovskite photovoltaics. *Nature Materials*, 17(5):372–376, 2018.

[59] Li Na Quan, F Pelayo García de Arquer, Randy P Sabatini, and Edward H Sargent. Perovskites for Light Emission. *Advanced Materials*, 30(45):1801996, nov 2018. Cited on page 48.

[60] David A Egger, Achintya Bera, David Cahen, Gary Hodes, Thomas Kirchartz, Leeor Kronik, Robert Lovrincic, Andrew M Rappe, David R Reichman, and Omer Yaffe. What Remains Unexplained about the Properties of Halide Perovskites? *Advanced Materials*, 30(20):1800691, may 2018. Cited on page 48.

[61] Douglas H Fabini, Ram Seshadri, and Mercouri G Kanatzidis. The underappreciated lone pair in halide perovskites underpins their unusual properties , jun 2020. Cited on page 48.

[62] Jianwei Sun, Adrienn Ruzsinszky, and John P. Perdew. Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Physical Review Letters*, 115(3):36402, jul 2015. Cited on pages 50, 51 and 52.

[63] Aliaksandr V Krukau, Oleg A Vydrov, Artur F Izmaylov, and Gustavo E Scuseria. Influence of the exchange screening parameter on the performance of screened hybrid functionals. *The Journal of Chemical Physics*, 125(22):224106, dec 2006. Cited on pages 51, 52, 53 and 54.

[64] Ann-Katrin U Michel, Andreas Heßler, Sebastian Meyer, Julian Pries, Yuan Yu, Thomas Kalix, Martin Lewin, Julian Hanss, Angela De Rose, Tobias W W Maß, Matthias Wuttig, Dmitry N Chigrin, and Thomas Taubner. Advanced Optical Programming of Individual Meta-Atoms Beyond the Effective Medium Approach. *Advanced Materials*, 31(29):1901033, jul 2019. Cited on page 57.

[65] A C Ferreira, S Paofai, A Létoublon, J Ollivier, S Raymond, B Hehlen, B Rufflé, S Cordier, C Katan, J Even, and P Bourges. Direct evidence of weakly dispersed and strongly anharmonic optical phonons in hybrid perovskites. *Communications Physics*, 3(1):48, 2020.

[66] T Lanigan-Atkins, X He, M J Krogstad, D M Pajerowski, D L Abernathy, Guangyong N M N Xu, Zhijun Xu, D-Y Chung, M G Kanatzidis, S Rosenkranz, R Osborn, and O Delaire. Two-dimensional overdamped fluctuations of the soft perovskite lattice in CsPbBr(3). *Nature materials*, 20(7):977–983, jul 2021. Cited on page 57.

[67] Kostiantyn Shportko, Stephan Kremers, Michael Woda, Dominic Lencer, John Robertson, and Matthias Wuttig. Resonant bonding in crystalline phase-change materials. *Nature Materials*, 2008. Cited on page 57.

[68] Alexander Kiehn. Analysis of Property Trends in Two-Dimensional Materials via Quantum-Chemical Bonding Descriptors, 2022. Cited on pages 58, 60, 63, 64, 187, 188 and 189.

[69] Ruining Wang, Felix R L Lange, Stefano Cecchi, Michael Hanke, Matthias Wuttig, and Raffaella Calarco. 2D or Not 2D: Strain Tuning in Weakly Coupled Heterostructures. *Advanced Functional Materials*, 28(14):1705901, apr 2018. Cited on page 58.

[70] Antonio M Mio, Philipp M Konze, Alexander Meledin, Michael Küpers, Marc Pohlmann, Marvin Kaminski, Richard Dronskowski, Joachim Mayer, and Matthias Wuttig. Impact of Bonding on the Stacking Defects in Layered Chalcogenides. *Advanced Functional Materials*, 29(37):1902332, sep 2019. Cited on page 61.

[71] Wouter Mortelmans, Karel De Smet, Ruishen Meng, Michel Houssa, Stefan De Gendt, Marc Heyns, and Clement Merckling. Role of Stronger Interlayer van der Waals Coupling in Twin-Free Molecular Beam Epitaxy of 2D Chalcogenides. *Advanced Materials Interfaces*, 8(13):2100438, jul 2021.

[72] Pradeep R Varadwaj, Arpita Varadwaj, Helder M Marques, and Koichi Yamashita. Chalcogen Bonding in the Molecular Dimers of WCh2 (Ch = S, Se, Te): On the Basic Understanding of the Local Interfacial and Interlayer Bonding Environment in 2D Layered Tungsten Dichalcogenides, 2022.

[73] Pradeep R Varadwaj, Helder M Marques, Arpita Varadwaj, and Koichi Yamashita. Chalcogen $\cdots$ Chalcogen Bonding in Molybdenum Disulfide, Molybdenum Diselenide and Molybdenum Ditelluride Dimers as Prototypes for a Basic Understanding of the Local Interfacial Chemical Bonding Environment in 2D Layered Transition Metal Dichalcogenides, 2022. Cited on page 58.

[74] Kamal Choudhary, Irina Kalish, Ryan Beams, and Francesca Tavazza. High-throughput Identification and Characterization of Two-dimensional Materials using Density functional theory. *Scientific Reports*, 7(1):5179, 2017. Cited on page 58.

[75] Michael Ashton, Joshua Paul, Susan B Sinnott, and Richard G Hennig. Topology-Scaling Identification of Layered Solids and Stable Exfoliated 2D Materials. *Physical Review Letters*, 118(10):106101, mar 2017. Cited on pages 58, 59 and 60.

[76] Prashun Gorai, Eric S Toberer, and Vladan Stevanović. Computational identification of promising thermoelectric materials among known quasi-2D binary compounds. *Journal of Materials Chemistry A*, 4(28):11110–11116, 2016. Cited on pages 58, 60 and 168.

[77] Gowoon Cheon, Karel-Alexander N Duerloo, Austin D Sendek, Chase Porter, Yuan Chen, and Evan J Reed. Data Mining for New Two- and One-Dimensional Weakly Bonded Solids and Lattice-Commensurate Heterostructures. *Nano Letters*, 17(3):1915–1923, mar 2017. Cited on pages 59 and 60.

[78] Nicolas Mounet, Marco Gibertini, Philippe Schwaller, Davide Campi, Andrius Merkys, Antimo Marrazzo, Thibault Sohier, Ivano Eligio Castelli, Andrea Cepellotti, Giovanni Pizzi, and Nicola Marzari. Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds. *Nature nanotechnology*, 13(3):246–252, mar 2018. Cited on pages 59 and 60.

[79] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of Chemical Physics*, 132(15):154104, apr 2010. Cited on pages 62 and 162.

[80] Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, and Kristin A Persson. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1(1):11002, jul 2013. Cited on pages 63, 79 and 161.

[81] Jason M Munro, Katherine Latimer, Matthew K Horton, Shyam Dwaraknath, and Kristin A Persson. An improved symmetry-based approach to reciprocal space path selection in band structure calculations. *npj Computational Materials*, 6(1):112, 2020. Cited on page 63.

[82] Peter Kerres, Yiming Zhou, Hetal Vaishnav, Mohit Raghuwanshi, Jiangjing Wang, Maria Häser, Marc Pohlmann, Yudong Cheng, Carl-Friedrich Schön, Thomas Jansen, Christophe Bellin, Daniel E Bürgler, Abdur Rehman Jalil, Christoph Ringkamp, Hugo Kowalczyk, Claus M Schneider, Abhay Shukla, and Matthias Wuttig. Scaling and Confinement in Ultrathin Chalcogenide Films as Exemplified by GeTe. *Small*, 18(21):2201753, may 2022. Cited on pages 65, 66, 67 and 202.

[83] Andreas Heßler, Sophia Wahl, Till Leuteritz, Antonios Antonopoulos, Christina Stergianou, Carl-Friedrich Schön, Lukas Naumann, Niklas Eicker, Martin Lewin, Tobias W W Maß, Matthias Wuttig, Stefan Linden, and Thomas Taubner. In3SbTe2 as a programmable nanophotonics material platform for the infrared. *Nature Communications*, 12(1):924, 2021. Cited on pages 68, 69 and 202.

[84] Personal Correspondence with Christian Teichrib, RWTH Aachen University, 2023. Cited on page 68.

[85] Chhatrasal Gayner, Kamal K Kar, and Woochul Kim. Recent progress and futuristic development of PbSe thermoelectric materials and devices. *Materials Today Energy*, 9:359–376, 2018. Cited on page 69.

[86] Gangjian Tan, Michihiro Ohta, and Mercouri G Kanatzidis. Thermoelectric power generation: from new materials to devices. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 377(2152):20180450, jul 2019.

[87] Yixuan Wu, Pengfei Nan, Zhiwei Chen, Zezhu Zeng, Ruiheng Liu, Hongliang Dong, Li Xie, Youwei Xiao, Zhiqiang Chen, Hongkai Gu, Wen Li, Yue Chen, Binghui Ge, and Yanzhong Pei. Thermoelectric Enhancements in PbTe Alloys Due to Dislocation-Induced Strains and Converged Bands. *Advanced Science*, 7(12):1902628, jun 2020. Cited on page 69.

[88] Min Zhu, Oana Cojocaru-Mirédin, Antonio M Mio, Jens Keutgen, Michael Küpers, Yuan Yu, Ju-Young Cho, Richard Dronskowski, and Matthias Wuttig. Unique Bond Breaking in Crystalline Phase Change Materials and the Quest for Metavalent Bonding. *Advanced Materials*, 30(18):1706735, may 2018. Cited on page 70.

[89] Yudong Cheng, Oana Cojocaru-Mirédin, Jens Keutgen, Yuan Yu, Michael Küpers, Mathias Schumacher, Pavlo Golub, Jean-Yves Raty, Richard Dronskowski, and Matthias Wuttig. Understanding the Structure and Properties of Sesqui-Chalcogenides (i.e., V(2) VI(3) or Pn(2) Ch(3) (Pn = Pnictogen, Ch = Chalcogen) Compounds) from a Bonding Perspective. *Advanced materials (Deerfield Beach, Fla.)*, 31(43):e1904316, oct 2019.

[90] Yuan Yu, Matteo Cagnoni, Oana Cojocaru-Mirédin, and Matthias Wuttig. Chalcogenide Thermoelectrics Empowered by an Unconventional Bonding Mechanism. *Advanced Functional Materials*, 30(8):1904862, feb 2020. Cited on page 70.

[91] Jiaojian Shi, Yijing Huang, Christian Heide, Burak Guzelturk, Carl-Friedrich Schön, Haowei Xu, Yuki Kobayashi, Andrew F. May, Pooja Donthi Reddy, Duan Luo, Eamonn Hughes, Kunal Mukherjee, Mariano Trigo, Ju Li, Jian Zhou, Shambhu Ghimire, Matthias Wuttig, David A. Reis, and Aaron M. Lindenberg. Nonresonant optomechanical phase control [Manuscript submitted for publication], 2023. Cited on pages 71, 72, 73 and 201.

[92] Jian Zhou, Haowei Xu, Yifei Li, R Jaramillo, and Ju Li. Opto-Mechanics Driven Fast Martensitic Transition in Two-Dimensional Materials. *Nano Letters*, 18(12):7794–7800, dec 2018. Cited on page 71.

[93] Alan Lai and Christopher A. Schuh. Direct Electric-Field Induced Phase Transformation in Paraelectric Zirconia via Electrical Susceptibility Mismatch. *Physical Review Letters*, 126(1):15701, jan 2021. Cited on page 71.

[94] Raagya Arora, Umesh V Waghmare, and C N R Rao. Metavalent Bonding Origins of Unusual Properties of Group IV Chalcogenides. *Advanced materials (Deerfield Beach, Fla.)*, 35(7):e2208724, feb 2023. Cited on page 77.

[95] Yuan Yu, Chongjian Zhou, Tanmoy Ghosh, Carl-Friedrich Schön, Yiming Zhou, Sophia Wahl, Mohit Raghuwanshi, Peter Kerres, Christophe Bellin, Abhay Shukla, Oana Cojocaru-Mirédin, and Matthias Wuttig. Doping by Design: Enhanced Thermoelectric Performance of GeSe Alloys Through Metavalent Bonding. *Advanced materials (Deerfield Beach, Fla.)*, 35(19):e2300893, may 2023. Cited on pages 77 and 201.

[96] June Gunn Lee. *Computational Material Science: An Introduction*. CRC Press, 2 edition, 2016. Cited on pages 83, 87, 92 and 100.

[97] David S. Sholl and Janice A. Steckel. *DENSITY FUNCTIONAL THEORY*. John Wiley & Sons, 2009.

[98] Eberhard Engel and Reiner M. Dreizler. *Density Functional Theory Theoretical and Mathematical Physics*. Springer-Verlag GmbH, 2011.

[99] F Bechstedt. *Many-Body Approach to Electronic Excitations: Concepts and Applications*. Springer Series in Solid-State Sciences. Springer Berlin Heidelberg, 2014.

[100] Lev Kantorovich. *Quantum Theory of the Solid State: An Introduction*. Springer Dordrecht, 1 edition, 2004.

[101] Gerd Czycholl. *Theoretische Festkörperphysik*. Springer Berlin, Heidelberg, 2008. Cited on pages 87, 88, 100 and 102.

[102] H. Ibach and H. Lüth. *Solid-State Physics*. Springer Berlin, Heidelberg, 4 edition, 2009. Cited on pages 83 and 86.

[103] W H Zachariasen. The Atomic Arrangement In Glass. *Journal of the American Chemical Society*, 54(10):3841–3851, oct 1932. Cited on page 84.

[104] C A Angell. Formation of Glasses from Liquids and Biopolymers. *Science*, 267(5206):1924–1935, mar 1995.

[105] Pablo G Debenedetti and Frank H Stillinger. Supercooled liquids and the glass transition. *Nature*, 410(6825):259–267, 2001. Cited on page 84.

[106] Rudolf Gross and Achim Marx. *Festkörperphysik*. Oldenbourg Verlag München, 2012. Cited on page 86.

[107] Wolfram Koch and Max C. Holthausen. *A Chemist's Guide to Density Functional Theory*. Wiley-VCH, 2 edition, 2001. Cited on pages 87 and 91.

[108] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 1964. Cited on page 87.

[109] W. Kohn and L. J. Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review*, 385(1951), 1965. Cited on page 88.

[110] Gordon J Miller, Yuemei Zhang, and Frank R Wagner. Chemical Bonding in Solids. In *Handbook of Solid State Chemistry*, pages 405–489. Wiley-VCH, aug 2017. Cited on page 88.

[111] P. E. Blöchl. CP-PAW Hands-On Course on First-Principles Calculations, 2021. Cited on page 90.

[112] Mel Levy. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proceedings of the National Academy of Sciences*, 76(12):6062–6065, dec 1979. Cited on page 91.

[113] Elliott H Lieb. Density functionals for coulomb systems. *International Journal of Quantum Chemistry*, 24(3):243–277, sep 1983. Cited on page 90.

[114] Richard M. Martin. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, 2nd edition, 2020. Cited on page 94.

[115] Veronika Brázdová and David R. Bowler. *Atomistic Computer Simulations: A Practical Guide*. Wiley-VCH, 2013. Cited on page 96.

[116] David R. Bowler. An introduction to pseudopotentials, 2015. Cited on pages 94 and 96.

[117] Giuseppe Grosso and Giuseppe Pastori Parravicini. Band Theory. In *Solid State Physics*, chapter Band Theor. Academic Press, 2014. Cited on page 94.

[118] G. B. Bachelet, D. R. Hamann, and M. Schlüter. Pseudopotentials that work: From H to Pu. *Physical Review B*, 26(8), 1982. Cited on page 95.

[119] David Vanderbilt. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Physical Review B*, 41(11):7892–7895, 1990. Cited on page 96.

[120] Dassault Systemes. CASTEP GUIDE, Materials Studio 2020, 2020. Cited on page 96.

[121] Carsten Rostgaard. The Projector Augmented-wave Method. *arxiv.org*, 2009. Cited on pages 96 and 97.

[122] J Enkovaara, C Rostgaard, J J Mortensen, J Chen, M Dułak, L Ferrighi, J Gavnholt, C Glinsvad, V Haikola, H A Hansen, H H Kristoffersen, M Kuisma, A H Larsen, L Lehtovaara, M Ljungberg, O Lopez-Acevedo, P G Moses, J Ojanen, T Olsen, V Petzold, N A Romero, J Stausholm-Møller, M Strange, G A Tritsaris, M Vanin, M Walter, B Hammer, H Häkkinen, G K H Madsen, R M Nieminen, J K Nørskov, M Puska, T T Rantala, J Schiøtz, K S Thygesen, and K W Jacobsen. Electronic structure calculations with GPAW: a real-space implementation of the

projector augmented-wave method. *Journal of Physics: Condensed Matter*, 22(25):253202, 2010. Cited on page 96.

[123] P. E. Blöchl. The projector augmented-wave method. *Physical Review B*, 50(24), 1994. Cited on page 97.

[124] Adam Kiejna, Georg Kresse, Jutta Rogal, Abir De Sarkar, Karsten Reuter, and Matthias Scheffler. Comparison of the full-potential and frozen-core approximation approaches to density-functional calculations of surfaces. *Physical Review B*, pages 6–9, 2006.

[125] J J Mortensen, L B Hansen, and K W Jacobsen. Real-space grid implementation of the projector augmented wave method. *Physical Review B*, pages 1–11, 2005. Cited on page 97.

[126] Philip J Hasnip, Keith Refson, Matt I J Probert, Jonathan R Yates, Stewart J Clark, and Chris J Pickard. Density functional theory in the solid state. *Molecular Physics*, 2014. Cited on page 98.

[127] Narbe Mardirossian and Martin Head-gordon. Thirty years of density functional theory in computational chemistry : an overview and extensive assessment of 200 density functionals. *Molecular Physics*, 8976, 2017. Cited on pages 98 and 99.

[128] S. H. Vosko, L Wilk, and M Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Canadian Journal of Physics*, 1980. Cited on page 98.

[129] John P Perdew and Yue Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Physical Review B*, 45(23):244–249, 1992. Cited on page 98.

[130] John P Perdew and Alex Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Physical Review B*, 23(10), 1981. Cited on page 98.

[131] D. M. Ceperley and B .J. Alder. Ground State of the Electron Gas by a Stochastic Method f. *Physical Review Letters*, 45(7), 1980. Cited on page 98.

[132] A. D. Becke. Density functional calculations of bond energies. *The Journal of Chemical Physics*, 4524(84), 1986. Cited on page 99.

[133] Frank Herman, John P Van Dyke, and Irene B Ortenburger. Improved Statistical Exchange Approximation for Inhomogeneous Many-Electron Systems. *Physical Review Letters*, 22(16):807–811, apr 1969. Cited on page 99.

[134] John P Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized Gradient Approximation Made Simple. *Physical Review Letters*, 77(18):3865–3868, 1996. Cited on page 99.

[135] I Csonka, Oleg A Vydrov, Gustavo E Scuseria, Lucian A Constantin, John P Perdew, Adrienn Ruzsinszky, Xiaolan Zhou, and Kieron Burke. Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces. *Physical Review Letters*, 136406(April):1–4, 2008. Cited on page 99.

[136] John P Perdew, J. A. Chevary, S. H. Vosko, Koblar A Jackson, Mark R Pederson, D J Singh, and Carlos Fiolhais. Atoms, molecules, s olids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Physical Review B*, 46(11), 1992. Cited on page 99.

[137] John Singleton. *Band Theory and Electronic Properties of Solids*. Oxford University Press, 2001. Cited on page 99.

[138] M Kuisma, J Ojanen, J Enkovaara, and T T Rantala. Kohn-Sham potential with discontinuity for band gap materials. *Physical Review B*, 82:1–7, 2010. Cited on page 100.

[139] M Städele, J A Majewski, and P Vogl. Exact Kohn-Sham Exchange Potential in Semiconductors. *Physical Review Letters*, 79(11):2089–2092, 1997.

[140] M Städele, M Moukara, J A Majewski, P Vogl, and A Görling. Exact exchange Kohn-Sham formalism applied to semiconductors ¨. *Physical Review B*, 59(15):29–33, 1999. Cited on page 100.

[141] N. W. Ashcroft and N. D. Mermin. *Solid State Physics*. Brooks/Cole, 1976. Cited on page 100.

[142] Hendrik Bruus and Karsten Flensberg. *Many-body Quantum Theory in Condensed Matter Physics: An Introduction*. Oxford University Press, 2004. Cited on page 102.

[143] Torsten Fließbach. Elektrodynamik: Lehrbuch zur Theoretischen Physik II. In *Elektrodynamik: Lehrbuch zur Theoretischen Physik II*, chapter Elektrodyn. Springer Berlin Heidelberg, 2008. Cited on page 102.

[144] M Gajdoš, K Hummer, G Kresse, J Furthmüller, and F Bechstedt. Linear optical properties in the projector-augmented wave methodology. *Physical Review B*, 73(4):45112, jan 2006. Cited on pages 102 and 103.

[145] Andrea Benassi, Andrea Ferretti, and Carlo Cavazzoni. PWSCF's epsilon.x user's manual. Cited on page 104.

[146] Theresa Julia Zielinski, Erica Harvey, Robert Sweeney, and David M Hanson. Quantum States of Atoms and Molecules. *Journal of Chemical Education*, 82(12):1880, dec 2005. Cited on pages 104, 105 and 106.

[147] Matteo Cagnoni. *Interplay Between Chemical Bonding, Band Structure and Charge Transport in Thermoelectric Chalcogenides*. PhD thesis, RWTH Aachen University, 2019. Cited on page 106.

[148] Jean Felix Dushimineza. Bonding investigation of d-metal alloys AlPd/Pt and GaPd/Pt by DFT and QTAIM methods, 2021. Cited on page 111.

[149] A. Otero-de-la Roza, Ángel Martín Pendás, and Erin R Johnson. Quantitative Electron Delocalization in Solids from Maximally Localized Wannier Functions. *Journal of Chemical Theory and Computation*, 14(9), 2018. Cited on pages 111, 115 and 116.

[150] A. Gallo-Bueno, M Kohout, E Francisco, and Á. Martín Pendás. Localization and Delocalization in Solids from Electron Distribution Functions. *Journal of Chemical Theory and Computation*, 18, 2022. Cited on page 111.

[151] Robert L Fulton and Stacey T Mixon. Comparison of Covalent Bond Indices and Sharing Indices. *J. Phys. Chem.*, 97(3):7530–7534, 1993. Cited on page 112.

[152] János G Ángyán, Michel Loos, and István Mayer. Covalent Bond Orders and Atomic Valence Indices in the Topological Theory of Atoms in Molecules. *J. Phys. Chem*, 98:5244–5248, 1994. Cited on page 112.

[153] Robert Ponec. Electron pairing and chemical bonds . Chemical structure , valences and structural similarities from the analysis of the Fermi holes. *Journal of Mathematical Chemistry*, 21:323–333, 1997. Cited on page 113.

[154] Robert Ponec. Electron pairing and chemical bonds . Molecular structure from the analysis of pair densities and related quantities. *Journal of Mathematical Chemistry*, 23:85–103, 1998. Cited on page 113.

[155] Wiliiam L Luken. Properties of the Fermi Hole in Molecules. *Croatica Chemica Acta*, 57(6):1283–1294, 1984. Cited on page 113.

[156] E Francisco, A Martín Pendás, and M A Blanco. A connection between domain-averaged Fermi hole orbitals and electron number distribution functions in real space. *The Journal of Chemical Physics*, 131(12):124125, sep 2009. Cited on page 113.

[157] J Cioslowski. Isopycnic Orbital Transformations and Localization of Natural Orbitals. *International Journal of Quantum Chemistry*, 38(S24), 1990. Cited on page 113.

[158] Jean-Yves Raty, Carlo Gatti, Carl-Friedrich Schön, and Matthias Wuttig. How to Identify Lone Pairs, Van der Waals Gaps, and Metavalent Bonding Using Charge and Pair Density Methods: From Elemental Chalcogens to Lead Chalcogenides and Phase-Change Materials. *physica status solidi (RRL) – Rapid Research Letters*, 15(11):2000534, nov 2021. Cited on pages 114 and 202.

[159] Nicola Marzari, Arash A Mostofi, Jonathan R Yates, Ivo Souza, and David Vanderbilt. Maximally localized Wannier functions: Theory and applications. *Reviews of Modern Physics*, 84(4):1419–1475, oct 2012. Cited on pages 115 and 116.

[160] J. C. Slater and W. Shockley. Optical Absorption by the Alkali Halides. *Physical Review*, 50, 1936. Cited on page 115.

[161] Gregory H. Wannier. The Structure of Electronic Excitation Levels in Insulating Crystals. *Physical Review*, 52, 1937. Cited on page 115.

[162] Nicola Marzari and David Vanderbilt. Maximally localized generalized Wannier functions for composite energy bands. *Physical Review B*, 56(20):847–865, 1997. Cited on page 116.

[163] Arash A. Mostofi, Jonathan R. Yates, Giovanni Pizzi, Young-Su Lee, Ivo Souza, David Vanderbilt, and Nicola Marzari. An updated version of wannier90: A tool for obtaining maximally-localised Wannier functions. *Computer Physics Communications*, 185(8), 2014. Cited on page 116.

[164] Denis Rafael Nacbar, Allan Victor Ribeiro, and Alexys Bruno-Alfonso. A Simple Geometrical Path Towards Hybrid Orbitals. *Materials Research*, 17(6):1474–1476, 2014. Cited on page 116.

[165] Paolo Giannozzi and The QUANTUM ESPRESSO Foundation. Quantum Espresso, 2022. Cited on pages 119 and 121.

[166] Paolo Giannozzi, Stefano Baroni, Nicola Bonini, Matteo Calandra, Roberto Car, Carlo Cavazzoni, Davide Ceresoli, Guido L Chiarotti, Matteo Cococcioni, Ismaila Dabo, Andrea Dal Corso, Stefano De Gironcoli, Stefano Fabris, Guido Fratesi, Ralph Gebauer, Uwe Gerstmann, Christos Gougoussis, Anton Kokalj, Michele Lazzeri, Layla Martin-samos, Nicola Marzari, Francesco Mauri, Riccardo Mazzarello, Stefano Paolini, Alfredo Pasquarello, Lorenzo Paulatto, Carlo Sbraccia, Alexander Smogunov, and Paolo Umari. QUANTUM ESPRESSO : a modular and open-source software project for quantum simulations of materials. *Journal of Physics: Condensed Matter*, 21, 2009.

[167] P Giannozzi, O Andreussi, T Brumme, O Bunau, and M Buongiorno. Advanced capabilities for materials modelling with Quantum ESPRESSO. *Journal of Physics: Condensed Matter*, 29, 2017.

[168] Paolo Giannozzi, Pietro Bonfà, Davide Brunato, Roberto Car, Ivan Carnimeo, and Stefano De Gironcoli. Quantum ESPRESSO toward the exascale. *The Journal of Chemical Physics*, 152, 2020. Cited on page 119.

[169] A. Otero-de-la Roza, Erin R Johnson, and Víctor Luaña. Critic2: A program for real-space analysis of quantum chemical interactions in solids. *Computer Physics Communications*, 185(3), 2014. Cited on page 119.

[170] A. Otero-de-la Roza, M.A. Blanco, Ángel Martín Pendás, and Víctor Luaña. Critic: a new program for the topological analysis of solid-state electron densities. *Computer Physics Communications*, 180(1), 2009. Cited on page 119.

[171] The QUANTUM ESPRESSO Foundation. Quantum Espresso Input Parameters, 2022. Cited on page 121.

[172] A. Otero-de-la Roza. Delocalization Indices in Solids, 2022. Cited on page 122.

[173] Ben G. Streetman and Sanjay Banerjee. *Solid State electronic Devices*. New Jersey: Prentice Hall, 5th edition, 2000. Cited on page 129.

[174] Xavier Gonze, Bernard Amadon, Gabriel Antonius, Frédéric Arnardi, Lucas Baguet, Jean-Michel Beuken, Jordan Bieder, François Bottin, Johann Bouchet, Eric Bousquet, Nils Brouwer, Fabien Bruneval, Guillaume Brunin, Théo Cavignac, Jean-Baptiste Charraud, Wei Chen, Michel Côté, Stefaan Cottenier, Jules Denier, Grégory Geneste, Philippe Ghosez, Matteo Giantomassi, Yannick Gillet, Olivier Gingras, Donald R Hamann, Geoffroy Hautier, Xu He, Nicole Helbig, Natalie Holzwarth, Yongchao Jia, François Jollet, William Lafargue-Dit-Hauret, Kurt Lejaeghere, Miguel A L Marques, Alexandre Martin, Cyril Martins, Henrique P C Miranda, Francesco Naccarato, Kristin Persson, Guido Petretto, Valentin Planes, Yann Pouillon, Sergei Prokhorenko, Fabio Ricci, Gian-Marco Rignanese, Aldo H Romero, Michael Marcus Schmitt, Marc Torrent, Michiel J van Setten, Benoit Van Troeye, Matthieu J Verstraete, Gilles Zérah, and Josef W Zwanziger. The Abinitproject: Impact, environment and recent developments. *Computer Physics Communications*, 248:107042, 2020. Cited on page 135.

[175] Aldo H Romero, Douglas C Allan, Bernard Amadon, Gabriel Antonius, Thomas Applencourt, Lucas Baguet, Jordan Bieder, François Bottin, Johann Bouchet, Eric Bousquet, Fabien Bruneval, Guillaume Brunin, Damien Caliste, Michel Côté, Jules Denier, Cyrus Dreyer, Philippe Ghosez, Matteo Giantomassi, Yannick Gillet, Olivier Gingras, Donald R Hamann, Geoffroy Hautier, François Jollet, Gérald Jomard, Alexandre Martin, Henrique P C Miranda, Francesco Naccarato, Guido Petretto, Nicholas A Pike, Valentin Planes, Sergei Prokhorenko, Tonatiuh Rangel, Fabio Ricci, Gian-Marco Rignanese, Miquel Royo, Massimiliano Stengel, Marc Torrent, Michiel J van Setten, Benoit Van Troeye, Matthieu J Verstraete, Julia Wiktor, Josef W Zwanziger, and Xavier Gonze. ABINIT: Overview and focus on selected capabilities. *The Journal of Chemical Physics*, 152(12):124102, mar 2020. Cited on page 135.

[176] Miroslav Kohout. DGrid. Cited on page 135.

[177] The Abinit Group. Abinit Input Variables, 2022. Cited on pages 137, 138 and 149.

[178] N A W Holzwarth, A R Tackett, and G E Matthews. A Projector Augmented Wave (PAW) code for electronic structure calculations, Part I: atompaw for generating atom-centered functions. *Computer Physics Communications*, 135(3):329–347, 2001. Cited on page 143.

[179] Kitware Inc. ParaView. Cited on page 147.

[180] G Kresse and J Hafner. Ab initio molecular dynamics for liquid metals. *Physical Review B*, 47(1):558–561, jan 1993. Cited on page 150.

[181] G Kresse and J Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science*, 6(1):15–50, 1996.

[182] G Kresse and J Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B*, 54(16):11169–11186, oct 1996.

[183] Universität Wien. Vasp.at. Cited on pages 150 and 152.

[184] Stefan Maintz, Volker L. Deringer, Andrei L. Tchougréeff, and Richard Dronskowski. Lobster. Cited on page 150.

[185] Universität Wien. VASP Wiki. Cited on pages 151 and 152.

[186] Davide Sangalli, J A Berger, Claudio Attaccalite, Myrta Grüning, and Pina Romaniello. Optical properties of periodic systems within the current-current response framework: Pitfalls and remedies. *Physical Review B*, 95(15):155203, apr 2017. Cited on page 157.

[187] Davide Sangalli, Andrea Marini, and Alberto Debernardi. Pseudopotential-based first-principles approach to the magneto-optical Kerr effect: From metals to the inclusion of local fields and excitonic effects. *Physical Review B*, 86(12):125139, sep 2012. Cited on page 157.

[188] Yoyo Hinuma, Giovanni Pizzi, Yu Kumagai, Fumiyasu Oba, and Isao Tanaka. Band structure diagram paths based on crystallography. *Computational Materials Science*, 128:140–184, 2017. Cited on page 160.

[189] Shyue Ping Ong, William Davidson Richards, Anubhav Jain, Geoffroy Hautier, Michael Kocher, Shreyas Cholia, Dan Gunter, Vincent L Chevrier, Kristin A Persson, and Gerbrand Ceder. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science*, 68:314–319, 2013. Cited on page 160.

[190] Alexandr Fonari and Christophe Sutton. Effective Mass Calculator, 2012. Cited on page 160.

[191] Oleg Rubel, Fabien Tran, Xavier Rocquefelte, and Peter Blaha. Perturbation approach to ab initio effective mass calculations. *Computer Physics Communications*, 261:107648, 2021. Cited on page 161.

[192] A. E. van Arkel. Molecules and crystals in inorganic chemistry. *ACS Publications*, 1956. Cited on page 165.

[193] J. A. A. Ketelaar. Chemical Constitution. An introduction to the theory of the chemical bond. *ACS Publications*, 1958. Cited on page 165.

[194] P B Littlewood. The crystal structure of IV-VI compounds. I. Classification and description. *Journal of Physics C: Solid State Physics*, 13(26):4855, 1980. Cited on page 166.

[195] Maximilian J Müller, Aakash Yadav, Christoph Persch, Sophia Wahl, Felix Hoff, and Matthias Wuttig. Tailoring Crystallization Kinetics of Chalcogenides for Photonic Applications. *Advanced Electronic Materials*, 8(8):2100974, aug 2022. Cited on pages 170, 171 and 172.

[196] Hirohiko Araki. JoJo's Bizarre Adventure: Stone Ocean. *Weekly Shōnen Jump*, 2000. Cited on page 170.

[197] Hirohiko Araki. JoJo's Bizarre Adventure: Golden Wind. *Weekly Shōnen Jump*, 1995. Cited on page 171.

[198] Hirohiko Araki. JoJo's Bizarre Adventure: Diamond Is Unbreakable. *Weekly Shōnen Jump*, 1992. Cited on page 172.

[199] Hirohiko Araki. JoJo's Bizarre Adventure: Battle Tendency. *Weekly Shōnen Jump*, 1987. Cited on page 173.

[200] Hirohiko Araki. JoJo's Bizarre Adventure: Phantom Blood. *Weekly Shōnen Jump*, 1987. Cited on page 174.

[201] Hirohiko Araki. JoJo's Bizarre Adventure: Stardust Crusaders. *Weekly Shōnen Jump*, 1989. Cited on page 175.

[202] Hirohiko Araki. JoJo's Bizarre Adventure: Steel Ball Run. *Weekly Shōnen Jump*, 2004. Cited on page 176.

[203] Unity Technologies. Unity3D, 2023. Cited on page 177.

[204] Springer. SpringerMaterials. Cited on pages 190, 191, 192, 193, 194, 195, 196 and 197.

[205] Matthias Wuttig, Carl-Friedrich Schön, Dasol Kim, Pavlo Golub, Carlo Gatti, and Jean-Yves Raty. Metavalent versus Hypervalent Bonding: Is there a chance for reconciliation?, 2023. Cited on page 201.

[206] Wei Zhang, Hangming Zhang, Suyang Sun, Xiaozhe Wang, Zhewen Lu, Xudong Wang, Jiang-Jing Wang, Chunlin Jia, Carl-Friedrich Schön, Riccardo Mazzarello, En Ma, and Matthias Wuttig. Metavalent Bonding in Layered Phase-Change Memory Materials (Adv. Sci. 15/2023). *Advanced Science*, 10(15):2370094, may 2023. Cited on page 201.

Danksagung

Ich möchte all denjenigen danken, die mich in meiner Zeit am 1. Physikalischen Institut (IA) begleitet und unterstützt haben.

Der erste und größte Dank gebührt Prof. Dr. Matthias Wuttig. Seine fachliche Kompetenz und sein ansteckender Enthusiasmus für die Physik, Chemie und Materialwissenschaften haben diese Arbeit erst ermöglicht. Die Zusammenarbeit mit Prof. Dr. Wuttig hat mir zudem die Möglichkeit gegeben, auch Kompetenzen außerhalb der eigentlichen Forschung aufzubauen: Webauftritte programmieren, Vorlesungen filmen und schneiden, als Synchronsprecher die Vorlesung neu vertonen, am Tag der Physik das Institut im Livestream promoten, sowie das Designen von Grafiken und Journal Covers. Auch für die Aufenthalte und Konferenzen in Dresden, Regensburg, Berchtesgaden, Trieste, Boston, Oxford und Singapur möchte ich meinen Dank aussprechen. Vielen Dank für alles, Matthias!

Als Zweites danke ich Prof. Dr. Stefan Blügel für die Übernahme der Zweitkorrektur dieser Dissertation, sowie seine Unterstützung von theoretischer Seite.

Weiterhin möchte ich meinen Dank Dr. Jean-Yves Raty aussprechen. Ohne seine fachliche und technische Unterstützung wäre diese Arbeit nicht möglich gewesen. Aber auch die Zusammenarbeit an sich hat mir immer große Freude bereitet. Sein Humor hat auch die kryptischste Fehlermeldung erträglich gemacht.

Ein weiterer Dank gilt Prof. Dr. Kedar Hippalgaonkar, der es mir ermöglicht hat, mit seiner Arbeitsgruppe in Singapur zu forschen.

Zudem danke ich meinen Studenten, Jean Felix Dushimineza, Caghan Ünlüer, Alexander Kiehn und Diego Leonardo Zepeda-Hickstein sowie meinem Bonusstudenten Tim Bartsch, für die spannende Zusammenarbeit und die Projekte, die wir gemeinsam bearbeiten konnten.

Ein großer Dank gilt Dr. Matthias Dück, welcher meine Bachelor- und Masterarbeit betreut hat, und mich durch sein Vorbild dazu motiviert hat zu promovieren. Auch darüber hinaus hat er nie aufgehört, mich zu unterstützen.

Für das Korrekturlesen dieser Arbeit danke ich Maximilian Müller, Ramon Pfeiffer und Thorben Frahm, Alexander Kiehn, Carlo Gatti sowie James Turner für sprachliches und grammatikalisches Feedback.

Ich danke allen Mitgliedern des 1. Physikalischen Institut (IA), ehemaligen und aktuellen, die meine Zeit hier unvergesslich gemacht haben. Insbesondere Christian Teichrib, der es fast vier Jahre mit mir in einem Büro ausgehalten hat (ohne sich über meine viel zu laute Tastatur zu beschweren); Peter Kerres für all die gemeinsamen Aktivitäten innerhalb und außerhalb des Instituts; Felix Hoff für die Gesellschaft, die er mir stets auf der morgendlichen Fahrt zum Institut geleistet hat; Julian Mertens für die gemeinsame Zeit "beim Flitschen"; Sophia Wahl für all die zusätzlichen Kaffeepausen; Niklas Junker dafür, dass es ihm nie leid wurde, mich beim Kickern zu besiegen; Lisa Schäfer dafür, dass sie seit meines ersten Tages eine Konstante des Instituts war; Maria Häser für die gemeinsamen Besuche am "Snackomaten"; Jan Köttgen für seine Tipps zur Kaffeeoptimierung, sowie allen, die mich stets in die Mensa begleitet haben: Carolin Peterson, Franka Braun, Dasol Kim, Christian Stenz, Ramon Pfeiffer und Thomas Schmidt.

Großer Dank gilt auch Renate Breuer, Sarah Schlenter, Mara Kallweit, Oliver Lehmann, Stephan Hermes und Ralf Detemple für die logistische, organisatorische und technische Unterstützung am Institut.

Meinen Kommilitonen und Freunden Alexander Pauls, Christian von Byern (geb. Käseberg), Christian Teichrib, Thorben Frahm, Tom Görtzen und insbesondere Ellen Kim danke ich für all die gemeinsame Zeit während meines Studiums.

Mein letzter und wichtigster Dank gilt meiner Familie, die mich unentwegt während meines Studiums unterstützt haben, und ohne welche diese Arbeit nie hätte entstehen können.

Map Island: Rumor has it that it comes to live under the right conditions... Nevertheless, vacation. Finally. The island's shape seems somehow familiar however...

Eidesstattliche Erklärung

Carl-Friedrich Schön

erklärt hiermit, dass diese Dissertation und die darin dargelegten Inhalte die eigenen sind und selbstständig, als Ergebnis der eigenen originären Forschung, generiert wurden.

Hiermit erkläre ich an Eides statt

1. Diese Arbeit wurde vollständig oder größtenteils in der Phase als Doktorand dieser Fakultät und Universität angefertigt;

2. Sofern irgendein Bestandteil dieser Dissertation zuvor für einen akademischen Abschluss oder eine andere Qualifikation an dieser oder einer anderen Institution verwendet wurde, wurde dies klar angezeigt;

3. Wenn immer andere eigene- oder Veröffentlichungen Dritter herangezogen wurden, wurden diese klar benannt;

4. Wenn aus anderen eigenen- oder Veröffentlichungen Dritter zitiert wurde, wurde stets die Quelle hierfür angegeben. Diese Dissertation ist vollständig meine eigene Arbeit, mit der Ausnahme solcher Zitate;

5. Alle wesentlichen Quellen von Unterstützung wurden benannt;

6. Wenn immer ein Teil dieser Dissertation auf der Zusammenarbeit mit anderen basiert, wurde von mir klar gekennzeichnet, was von anderen und was von mir selbst erarbeitet wurde;

7. Kein Teil dieser Arbeit wurde vor deren Einreichung veröffentlicht.

Aachen, den 12.06.2023