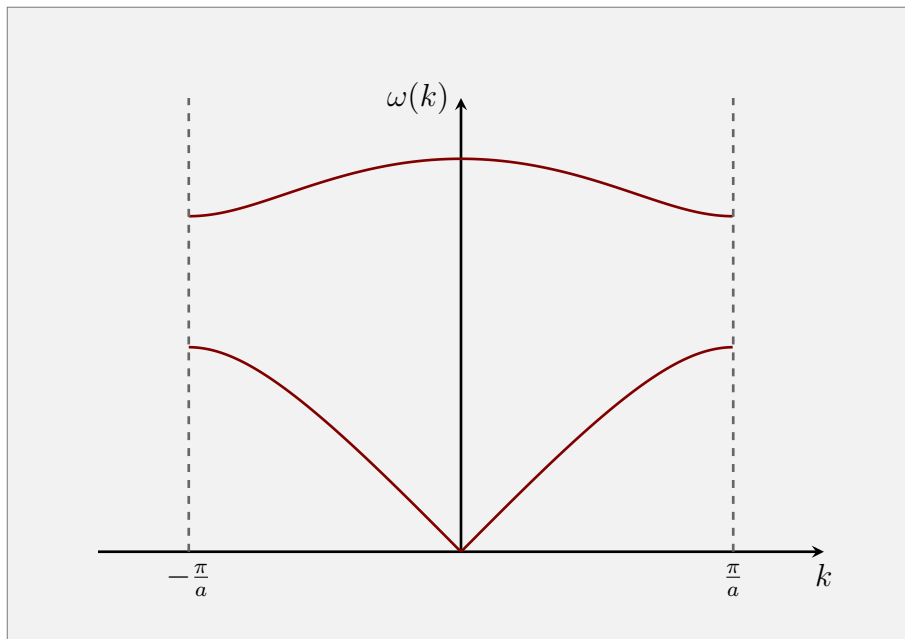


SOLID STATE THEORY

A short introduction



This page intentionally left blank.

Preface

Solid state theory is the description of matter and its macroscopic properties from a quantum mechanical point of view. These notes give a short introduction on how to approach the description of solids from quantum mechanics. They orient on a short introductory lecture course and follow [Czy16, chapters 1-6.2] mostly.

Some parts of these notes were created, at the time I was learning the corresponding concepts. So be aware, that there may not only be the usual typos, but possibly wrong statements. In that sense, read with caution. However, nothing presented here is new, and usually well covered in textbooks.

This page intentionally left blank.

Contents

Preface	i
1 Lattice structure	1
1.1 Bravais lattice	1
1.1.1 Definition of lattices	1
1.1.2 Primitive and unit cells	3
1.1.3 Wigner-Seitz cells and neighbors	4
1.1.4 Symmetry groups	5
1.1.5 Atomic basis	6
1.2 The reciprocal lattice and periodic functions	7
2 Separation of lattice and electron dynamics	10
2.1 General solid state Hamiltonian	10
2.2 Born-Oppenheimer approximation	12
3 Lattice vibrations	14
3.1 Harmonic approximation	14
3.2 Periodic boundary conditions	16
3.2.1 Translation operator	16
3.2.2 Bloch's theorem	17
3.2.3 Periodic boundary conditions	18
3.2.4 Diagonalization of the dynamical matrix	19
3.2.5 Ion Hamiltonian	21
3.3 Phonons	21
3.3.1 Phonon branches	22
3.3.2 Thermodynamics of phonons	22
4 Non-interacting electrons in solids	29
4.1 Electrons in a periodic potential	29
4.1.1 Application of Bloch's theorem	30
4.1.2 Fourier transformation of the Schrödinger equation	31
4.2 Nearly free electrons	32
4.3 The effective mass	35
4.3.1 The Rayleigh-Schrödinger approximation	35
4.3.2 Application of the Rayleigh-Schrödinger approximation	36
4.3.3 The effective mass tensor	38
4.4 The tight binding model	38
4.4.1 Terminology and idea behind the tight binding model	39
4.4.2 No overlap	39
4.4.3 General approximation for the overlap case	40
4.4.4 Wannier functions	43
4.4.5 Example: d -dimensional cubic lattice with nearest neighbor interaction	44
4.5 Density of electron states and Fermi surface	44
4.6 Thermodynamics of electrons in a solid	46
4.6.1 Sommerfeld expansion	47

4.6.2	The chemical potential	49
4.6.3	The specific heat	50
5	Outlook: Electron-electron interaction	51
5.1	The electron Hamiltonian in second quantization	51
5.2	Single band Hubbard model	52
5.3	Outlook	53
	Index	54
	Bibliography	55

Lattice structure

In the solid state, matter can have three structures. A crystalline structure, a poly crystalline structure and a non-crystalline structure. In a crystal, the constituents, e.g. atoms, molecules or groups of atoms and molecules, form a periodic structure, that spans the whole solid. A poly crystal is composed of many crystals. As the name suggests, non-crystals have no periodicity. Since most materials form crystals in the solid state, this chapter introduces the mathematical concepts to describe crystals.

1.1 Bravais lattice

1.1.1 Definition of lattices

Although the physics is confined to dimensions less than three, it can be useful, to keep the definitions general.

Definition 1.1.1.

An n -dimensional **Bravais lattice** is a point set $\mathbb{L} \subset E_n$, that is a subset of the n -dimensional Euclidean space E_n ,¹ that is periodic. This means, that there are n linearly independent vectors $v_j \in \mathbb{R}^n$, such that for any point $\sigma \in \mathbb{L}$, it holds that

$$\mathbb{L} := \left\{ p \in E_n \mid p = \sigma + \sum_{j=1}^n m_j \mathbf{v}_j, \quad m_j \in \mathbb{Z} \right\} .$$

Here we call the vectors $\{\mathbf{v}_j\}$ **primitive vectors**. The linear combination over \mathbb{Z} are called **lattice vectors**

$$L := \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle_{\mathbb{Z}} := \left\{ \mathbf{V} \in \mathbb{R}^n \mid \mathbf{V} = \sum_{j=1}^n m_j \mathbf{v}_j, \quad m_j \in \mathbb{Z} \right\} .$$

Remark 1.1.2.

If one fixes a single point $\sigma \in \mathbb{L}$ as the origin, the concept of affine spaces is not needed, and one can reduce E_n to \mathbb{R}^n . Then, the points p can be identified with

¹Here, the affine space is meant. Put simply, points and vectors as translations are differentiated. In fact, lattices can be defined on any affine space.

the vectors $p \equiv \sum_{j=1}^n m_j \mathbf{v}_j$ and thus:

$$\mathbb{L} \equiv L := \left\{ v \in \mathbb{R}^n \mid v = \sum_{j=1}^n m_j \mathbf{v}_j, \quad m_j \in \mathbb{Z} \right\} .$$

For that reason, L will denote the point set, as well as the lattice vectors.

The definition does not specify the choice of primitive vectors. Indeed, they are not

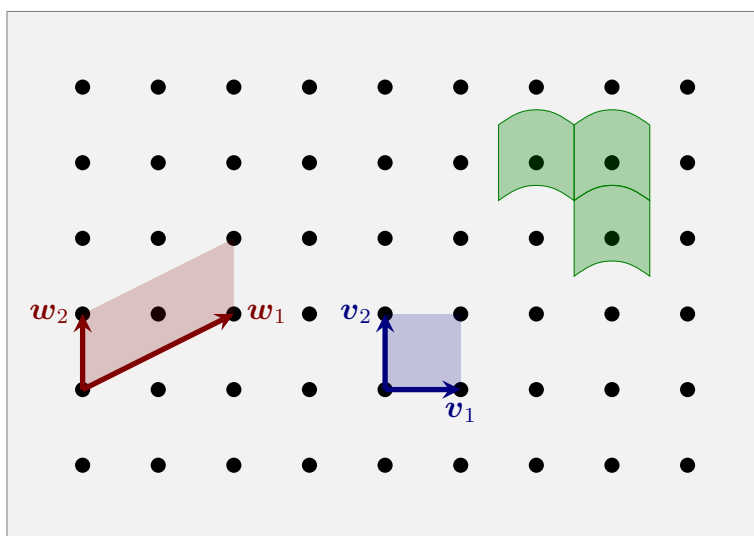


Figure 1.1: Example of a 2-dimensional lattice, with the choice of two different sets of generating vectors, with unit cells. Also primitive cells, that are not unit cells (green).

unique, and it requires conventions to fix them. This can be observed in figure 1.1. Yet, because of the definition, the linear span over \mathbb{Z}

$$\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle_{\mathbb{Z}}$$

does not depend on the choice of primitive vectors, since (suggestive notation):

$$\sigma + \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle_{\mathbb{Z}} = \sigma + L = \sigma + \langle \mathbf{w}_1, \dots, \mathbf{w}_n \rangle_{\mathbb{Z}} .$$

Definition 1.1.3.

A **translation vector** \mathbf{T} of a Bravais lattice L is an element of the linear span of a choice of (then any choice) primitive vectors \mathbf{v}_j :

$$\mathbf{T} \in \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle_{\mathbb{Z}} .$$

Fixing a choice of primitive vectors, a translation is determined by the coefficients $m_j \in \mathbb{Z}$, such that one writes:

$$\mathbf{T}_{\vec{m}} = \sum_{j=1}^n m_j \mathbf{v}_j .$$

Remark 1.1.4 (Notation).

In these notes, vectors are highlighted by boldface \mathbf{T} . However, to combine some objects m_j notationally, we write $\vec{m} = (m_1, \dots, m_n)$ and $\vec{m} \cdot \vec{\ell} = \sum_j m_j \ell_j$ etc.

1.1.2 Primitive and unit cells

The definition of translation, in the sense of lattices, allows to define primitive cells.²

Definition 1.1.5.

A **primitive cell** is a cell that satisfies the following conditions:

- i) It contains only one point.
- ii) It has minimal volume.
- iii) Upon lattice translations, it fills the whole lattice volume (tiling).

An example of a primitive cell, that is not a unit cell, can be found in figure 1.1. Unit cells are defined as follows:

Definition 1.1.6.

The parallelepiped spanned by the primitive vectors is called **unit cell**.

The volume of the unit cell is given by

$$V = \det(\mathbf{v}_1, \dots, \mathbf{v}_n) .$$

Another characteristic of a unit cell is the number of points, it contains. Again, considering figure 1.1 shows, that not all unit cells are primitive. Considering the primitive vectors $\{\mathbf{v}_i\}$ one could think, that the unit cell contains either 4 or no point. However, both is wrong, showing the need of a formal definition.

Definition 1.1.7.

A point p lies with a proportion of x in a unit cell C , if a for all $\varepsilon > 0$, the proportion of the volume of the ε -ball $B_\varepsilon(p)$ around p , that lies in C is x :

$$\text{Vol}(C \cap B_\varepsilon(p)) = x \cdot \text{Vol}(B_\varepsilon(p)) .$$

This definition is seems abstract at first, but becomes a matter of counting in dimensions less than 3.

Example 1.1.8.

If the unit cell is a cuboid in 3 dimensions, the following holds:

$$\begin{aligned} \text{point at corner:} & \quad x = 1/8 , \\ \text{point at edge:} & \quad x = 1/4 , \\ \text{point at surface:} & \quad x = 1/2 , \end{aligned}$$

²Here, the naming convention is somewhat odd. As will be seen, the cells, created by primitive vectors are nor called primitive cells, but unit cells. Primitive cells are defined by translational properties.

point in the interior: $x = 1$.

In case of a 2 dimensional rectangle, one starts from $1/4$. Hence for $\{\mathbf{v}_j\}$ in figure 1.1, there are $4 \times \frac{1}{4} = 1$ points in the unit cell. Hence, it is also a primitive cell.

1.1.3 Wigner-Seitz cells and neighbors

Primitive cells are not unique, as has been demonstrated by figure 1.1. Only their volume is unique, fixed by the density of the points in E_n . One approach, for a unique definition, is to discretize the notion of ε -ball:

Definition 1.1.9.

Let $d(\cdot, \cdot): E_n \times E_n \rightarrow \mathbb{R}$ denote the Euclidean distance of two points in E_n . Then, the **Wigner-Seitz cell** of a lattice point $\sigma \in L$ is defined as follows:

$$C(\sigma) := \{p \in E_n \mid d(p, \sigma) \leq d(p, q) \quad \forall q \in L \setminus \{\sigma\}\} .$$

It is worthwhile to reformulate the definition in terms of vector spaces. Fix the lattice point $\sigma \in L$ as origin and primitive vectors $\{\mathbf{v}_j\}$. In this conventional language, the definition of the Wigner-Seitz cell of a lattice point $\mathbf{R}_{\vec{m}} \in \mathbb{R}^n$ is given by:

$$C(\mathbf{R}_{\vec{m}}) := \{\mathbf{r} \in \mathbb{R}^n \mid \|\mathbf{r} - \mathbf{R}_{\vec{m}}\| \leq \|\mathbf{r} - \mathbf{R}_{\vec{n}}\| \quad \forall \vec{n} \neq \vec{m} \in \mathbb{Z}^n\} .$$

Example 1.1.10 (Construction of Wigner-Seitz cells).

The Wigner-Seitz cell of $\sigma \in L$ can be constructed as follows:

1. Connect the σ with neighboring points by straight lines.
2. Draw the orthogonal planes at half the length of the connecting lines.
3. The surface of the union of these planes, as viewed from the perspective of σ , is the boundary of the Wigner-Seitz cell.

The meaning of “from the perspective of σ ” is necessary, as only the innermost planes are needed. Compare to figure 1.2.

The meaning of neighboring point is not precise enough. For that reason, one defines the concept of nearest neighbor. The name already gives a definition, but we still define it formally:

Definition 1.1.11.

Let $p \in L$ be a lattice point. Another lattice point q is called **nearest neighbor**, if

$$\|p - q\| \leq \|p - r\| \quad \forall r \in L .$$

There can be more than one nearest neighbor, hence the use of less than equal. The

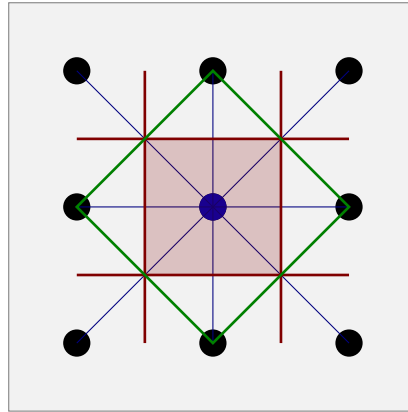


Figure 1.2: Construction of a Wigner-Seitz cell. The green lines do not contribute. They define the second Wigner-Seitz cell

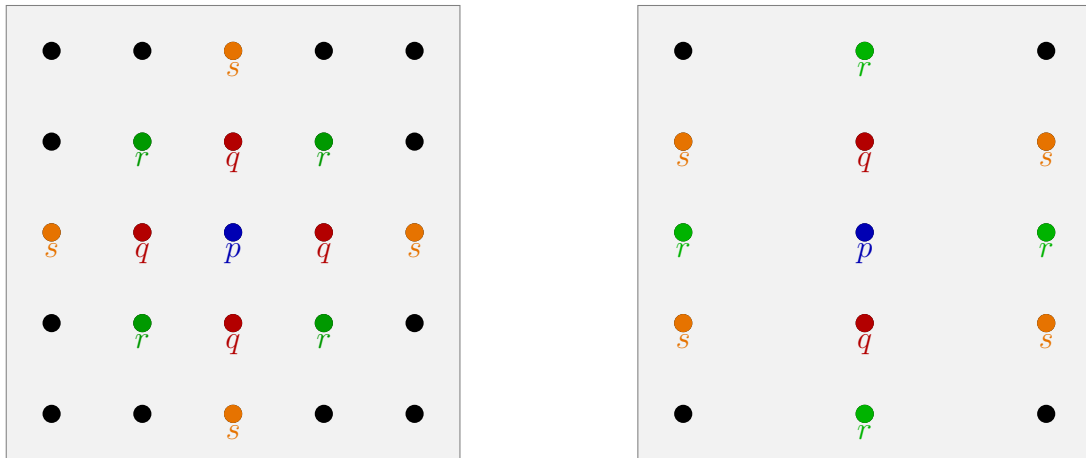


Figure 1.3: The nearest neighbors q , second nearest neighbors r and third nearest neighbors s of p , of different lattices

number of nearest neighbors is called **coordination number**. One defines the **second nearest neighbors**, by removing the nearest neighbors, and finding the next nearest neighbors. Repeating leads to the concept of **n -th nearest neighbors**.

1.1.4 Symmetry groups

Bravais lattices can be categorized by their symmetries. A symmetry is a space transformation, that leaves the lattice invariant. There are three symmetries, that come to mind:

Translation by a vector $\mathbf{t} \in \mathbb{R}^n$.

Rotation around a point $\sigma \in L$.

Space inversion at a point $\sigma \in L$ w.r.t. an axis A .

One can formalize these intuitive transformations by considering groups and their actions on E_n .

The translation group $T \cong \mathbb{R}^n$ acts on E_n by $\mathbf{t} \circ p = p + \mathbf{t}$ for all $\mathbf{t} \in T$. The symmetry subgroup of T for a lattice, denoted by $L(T)$ here, is by construction the linear span of some primitive vectors over \mathbb{Z} :

$$L(T) = \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle_{\mathbb{Z}} .$$

Rotations and space inversions together form a group, called orthogonal group $O(n) := O(\mathbb{R}^n)$. Fixing a point $\sigma \in L$, every other point $p = \sigma + \mathbf{r}$ of E_n can be identified with the vector $\mathbf{r} \in \mathbb{R}^n$. Then, $O(n)$ act on E_n , by $R \circ p = \sigma + R\mathbf{r}$, where the action of $O(n)$ on \mathbb{R}^n is understood. The symmetry group $L(O(n))$ is a subgroup of $O(n)$, that depends on the choice of lattice L .

To obtain a single group, that describes the symmetries of the lattice, one would proceed by considering the semidirect product of the translation symmetry group $L(T)$ and $L(O(n))$.

1.1.5 Atomic basis

The number of Bravais lattices is finite. In fact, there are only 14 Bravais lattices in three dimensions. However, there are much more crystalline structures in nature, e.g. the diamond structure or the NaCl structure. In two dimensions, the honeycomb structure is also no Bravais lattice. This means, that an additional concept apart from Bravais lattices is needed.

Definition 1.1.12.

The **atomic position vectors** are a set of vectors $\{\mathbf{b}_j\}_{j=1}^k$, that specify the positions \mathbf{q}_j of the k atoms, w.r.t. Bravais lattice positions p :

$$\mathbf{q}_j = p + \mathbf{b}_j .$$

If the atomic position is only $\mathbf{b}_1 = 0$, one obtains the original Bravais lattice. If no position vector is 0, no point of the Bravais lattice is in the crystalline structure. Choosing an origin σ , the lattice for a crystalline structure is modeled by the primitive vectors $\{\mathbf{v}_j\}$ and the atomic position vectors $\{\mathbf{b}_k\}$.

An example for a crystalline structure in two dimensions, that is no Bravais lattice, is the Honeycomb lattice (see figure 1.4). It is defined by the primitive vectors

$$\mathbf{v}_1 = a \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{v}_2 = a \begin{pmatrix} \cos(\frac{2\pi}{3}) \\ \sin(\frac{2\pi}{3}) \end{pmatrix} ,$$

where a is the lattice constant, and the position vectors:

$$\mathbf{b}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{b}_2 = a \frac{\sqrt{3}}{3} \begin{pmatrix} \cos(\frac{\pi}{3}) \\ \sin(\frac{\pi}{3}) \end{pmatrix} .$$

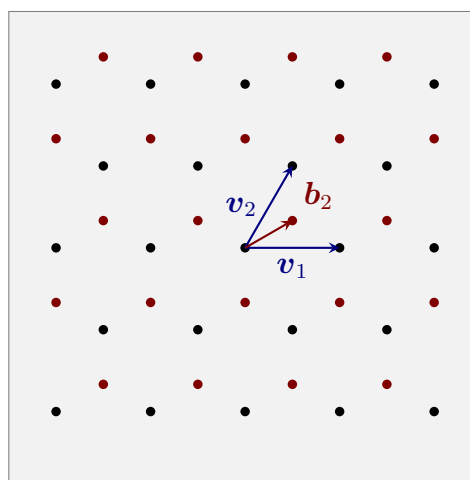


Figure 1.4: Honeycomb lattice.

The symmetry considerations of subsection 1.1.4 also apply to crystalline structures, providing a tool for further classification.

1.2 The reciprocal lattice and periodic functions

Let L be a Bravais lattice and $\{\mathbf{v}_j\}$ the primitive lattice vectors. Let $f: E_n \rightarrow \mathbb{C}$ be a function with the same periodicity as the lattice:

$$f(\mathbf{r} + \mathbf{R}_{\vec{m}}) = f(\mathbf{r}) \quad \forall \mathbf{r} \in E_n, \mathbf{R}_{\vec{m}} \in \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle_{\mathbb{Z}}.$$

Fix an origin σ , such that E_n can be identified with \mathbb{R}^n and choose the basis $\{\frac{\mathbf{v}_j}{\|\mathbf{v}_j\|}\}$:

$$\mathbf{r} = \sigma + \sum_{j=1}^n r_j \frac{\mathbf{v}_j}{\|\mathbf{v}_j\|}.$$

There is a function $\mathfrak{f}: \mathbb{R}^n \rightarrow \mathbb{C}$, that is the coordinate representation of f :

$$\mathfrak{f}(r_1, \dots, r_n) = f(\mathbf{r}) = f\left(\sigma + r_j \frac{\mathbf{v}_j}{\|\mathbf{v}_j\|}\right).$$

This coordinate representation, has the periodicity $\|\mathbf{v}_j\|$ in the j -th component:

$$\mathfrak{f}(r_1, \dots, r_j + \|\mathbf{v}_j\|, \dots, r_n) = \mathfrak{f}(r_1, \dots, r_j, \dots, r_n).$$

Then, the Fourier transformation is well defined:

$$f(\mathbf{r}) = \sum_{2\pi i \vec{\ell} \in \mathbb{Z}^n} \tilde{\mathfrak{f}}_{\vec{\ell}} \exp\left(2\pi i \sum_{j=1}^n \ell_j \frac{r_j}{\|\mathbf{v}_j\|}\right).$$

with

$$\tilde{\mathfrak{f}}_{\vec{\ell}} = \frac{1}{\|\mathbf{v}_1\|} \int_0^{\|\mathbf{v}_1\|} \dots \frac{1}{\|\mathbf{v}_n\|} \int_0^{\|\mathbf{v}_n\|} \mathfrak{f}(r_1, \dots, r_n) \exp\left(2\pi i \sum_{j=1}^n \ell_j \frac{r_j}{\|\mathbf{v}_j\|}\right) dr_1 \dots dr_n.$$

Next, we observe, that the term in the exponential can be written as:

$$2\pi i \sum_{j=1}^n \ell_j \frac{r_j}{\|\mathbf{v}_j\|} = i \left\langle \sum_j \ell_j \mathbf{g}_j, \sum_j \frac{r_j}{\|\mathbf{v}_j\|} \mathbf{v}_j \right\rangle \quad \text{iff} \quad \langle \mathbf{g}_i, \mathbf{v}_j \rangle = 2\pi \delta_{ij}.$$

This leads to the following definition:

Definition 1.2.1.

Let L be a lattice with primitive vectors $\{\mathbf{v}_j\}$. The **reciprocal lattice (for σ)** is defined by

$$\tilde{L} := \langle \mathbf{g}_1, \dots, \mathbf{g}_n \rangle_{\mathbb{Z}}$$

where $\mathbf{g}_j \in \mathbb{R}^n$ are defined by:

$$\langle \mathbf{g}_i, \mathbf{v}_j \rangle = 2\pi \delta_{ij}.$$

The name lattice is somewhat misleading, as the lattice is not a lattice in real space, as the points depend on the choice of origin σ . In fact, for every Bravais lattice point $p \in L$, one defines a separate reciprocal lattice. Similar to the concept of vector bundles.

Theorem 1.2.2 (Reciprocal lattice, alternative definition).

For the reciprocal lattice, it holds that:

$$\tilde{L} = \{ \mathbf{G} \in \mathbb{R}^n \mid \langle \mathbf{G}, \mathbf{R} \rangle \in 2\pi\mathbb{Z} \quad \forall \mathbf{R} \in L \} .$$

Proof 1.2.3.

“ \subset ”:

Choosing primitive vectors $\{\mathbf{v}_j\}$ of L , the lattice vector \mathbf{R} can be written as $\mathbf{R} = \sum_{i=1}^n m_i \mathbf{v}_i$, thus:

$$\begin{aligned} \langle \mathbf{G}, \mathbf{R} \rangle &= \left\langle \sum_j \ell_j \mathbf{g}_j, \sum_i m_i \mathbf{v}_i \right\rangle = \sum_{i,j} \ell_j m_i \langle \mathbf{g}_j, \mathbf{v}_i \rangle \\ &= 2\pi \sum_{i,j} \ell_j m_i \delta_{ij} = 2\pi \sum_j \ell_j m_j \in 2\pi\mathbb{Z} . \end{aligned}$$

“ \supset ”:

Let \mathbf{G} be as assumed. Since $\{\mathbf{g}_i\}$ are a basis, there are $x_i \in \mathbb{R}$, such that $\mathbf{G} = \sum_i x_i \mathbf{g}_i$. It remains to show, that $x_i \in \mathbb{Z}$. By assumption, $\langle \mathbf{G}, \mathbf{R} \rangle \in 2\pi\mathbb{Z}$ for all $\mathbf{R} \in L$, especially for $\mathbf{R} = \mathbf{v}_j$. But then:

$$2\pi\mathbb{Z} \ni \langle \mathbf{G}, \mathbf{v}_j \rangle = \sum_i x_i \langle \mathbf{g}_i, \mathbf{v}_j \rangle = 2\pi \sum_i x_i \delta_{ij} = 2\pi x_j .$$

□

Corollary 1.2.4.

The reciprocal lattice for a point $\sigma \in L$ does not depend on the choice of primitive vectors.

Remark 1.2.5.

In three dimensions, there is a standard construction for the reciprocal basis vectors:

$$\mathbf{g}_i = \varepsilon_{ijk} \frac{\mathbf{v}_j \times \mathbf{v}_k}{|\det(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)|} .$$

It still remains, to check that the Fourier series for f still satisfies the periodicity conditions. Using the definition of the reciprocal lattice, we can write:

$$\begin{aligned} f(\mathbf{r}) &= \sum_{\tilde{\ell} \in \mathbb{Z}^n} \tilde{f}_{\tilde{\ell}} \exp \left(i \left\langle \mathbf{G}_{\tilde{\ell}}, \sum_j \frac{r_j}{\|\mathbf{v}_j\|} \mathbf{v}_j \right\rangle \right) . \\ \Rightarrow \quad f(\mathbf{r} + \mathbf{R}) &= \sum_{\tilde{\ell} \in \mathbb{Z}^n} \tilde{f}_{\tilde{\ell}} \exp \left(i \left\langle \mathbf{G}_{\tilde{\ell}}, \mathbf{R} + \sum_j \frac{r_j}{\|\mathbf{v}_j\|} \mathbf{v}_j \right\rangle \right) \end{aligned}$$

$$\begin{aligned}
&= \sum_{\vec{\ell} \in \mathbb{Z}^n} \tilde{f}_{\vec{\ell}} \exp \left(i \left\langle \mathbf{G}_{\vec{\ell}}, \sum_j \frac{r_j}{\|\mathbf{v}_j\|} \mathbf{v}_j \right\rangle \right) e^{i \langle \mathbf{G}_{\vec{\ell}}, \mathbf{R} \rangle} \\
&= f(\mathbf{r}) ,
\end{aligned}$$

since $\langle \mathbf{G}_{\vec{\ell}}, \mathbf{R} \rangle = 2\pi N$ with $N \in \mathbb{Z}$ and $e^{2\pi N} = 1$.

Definition 1.2.6.

The (first) **Brillouin zone** is the Wigner-Seitz cell of $p = 0 \in \tilde{L}$, where \tilde{L} is the reciprocal lattice.

Remark 1.2.7 (On the proper notion of momentum).

The Fourier transformation is associated with momentum in quantum mechanics. However, since force is a 1-form (dual vector), so is momentum. Hence, the natural Fourier transformation uses the dual vector space (not needing a scalar product). Then, $g_i \in (\mathbb{R}^n)^*$ are the dual basis of \mathbf{v}_j , up to a factor 2π :

$$g_i(\mathbf{v}_j) = 2\pi \delta_{ij} .$$

In fact, if vectors have dimension length, dual vectors have the dimension $\frac{1}{\text{length}}$, which is in accordance to the standard construction from remark 1.2.5.

2

Separation of lattice and electron dynamics

Describing solids as quantum systems, one is interested in the Hamiltonian. Neglecting relativistic effects, the Hamiltonian is well known and fairly simple to be written down. However, because of the huge number of particles, not only analytical solutions, but also numerical solutions are impossible. For that reason, perturbative methods are necessary.

2.1 General solid state Hamiltonian

A solid consists of ions (atoms with bounded electrons) with relatively fixed positions and “free” valence electrons. There are N_a ions, with individual mass M_j and charge $Z_j e$. The N_e valence electrons each have the electron mass m_e and charge $-e$. From charge conservation it follows that:

$$0 = N_e \cdot (-e) + \sum_{j=1}^{N_a} Z_j \cdot e \quad \Rightarrow \quad N_e = \sum_{j=1}^{N_a} Z_j$$

The goal of this section is to find the Hamiltonian H for the stationary Schrödinger equation

$$H\psi(\vec{\mathbf{R}}, \vec{\mathbf{r}}) = E\psi(\vec{\mathbf{R}}, \vec{\mathbf{r}}) .$$

The wave function depends on the positions of the ions $\vec{\mathbf{R}} = \{\mathbf{R}_1, \dots, \mathbf{R}_{N_a}\}$ and on the positions of the valence electrons $\vec{\mathbf{r}} = \{\mathbf{r}_1, \dots, \mathbf{r}_{N_e}\}$.

Notation 2.1.1.

The μ -th component of \mathbf{r}_j is denoted by r_j^μ . The same holds for R_j^μ .

Remark 2.1.2.

It is common practice, to choose an origin and coordinates, when considering wave functions. Hence Bravais lattice points and Bravais vectors can be identified from here on.

Kinetic energy

With the separation of lattice and electron dynamics in mind, we split the kinetic energy:

$$T = T_a + T_e$$

where

$$T_a = \sum_{j=1}^{N_a} \frac{\mathbf{P}_j^2}{2M_j} = \sum_{j=1}^{N_a} \frac{-\hbar^2}{2M_j} \Delta_{\mathbf{R}_j} \quad \text{and} \quad T_e = \frac{-\hbar^2}{2m_e} \sum_{j=1}^{N_e} \Delta_{\mathbf{r}_j} .$$

Here, the Laplace operator $\Delta_{\mathbf{r}_j}$ acts on \mathbf{r}_j and $\Delta_{\mathbf{R}_j}$ acts on \mathbf{R}_j .

Potential energy

The potential energy can be split in three parts, the interaction of the ions with other ions V_{aa} , the interaction of the valence electrons with other electrons V_{ee} and the cross interaction V_{ea} :

$$V = V_{aa} + V_{aa} + V_{ea} .$$

The ion-ion interaction V_{aa} , electron-electron interaction V_{ee} and ion-electron interaction V_{ea} are:

$$V_{aa} = \sum_{j < k} \frac{Z_j Z_k e^2}{\|\mathbf{R}_j - \mathbf{R}_k\|} , \quad V_{ee} = \sum_{j < k} \frac{e^2}{\|\mathbf{r}_j - \mathbf{r}_k\|} , \quad V_{ea} = - \sum_{j,k} \frac{Z_j e^2}{\|\mathbf{r}_j - \mathbf{R}_k\|} .$$

The full Hamiltonian

Neglecting relativistic effects, the full **solid Hamiltonian** is

$$H = T_a + T_e + V_{aa} + V_{ee} + V_{ea} .$$

To separate the electron dynamics from the lattice dynamics, the Hamiltonian is rewritten in **atomic units**:

Bohr radius: Consider dimensionless length $\tilde{\mathbf{r}}_j, \tilde{\mathbf{R}}_j = \frac{\mathbf{r}_j}{a_0}, \frac{\mathbf{R}_j}{a_0}$, where $a_0 = \frac{\hbar^2}{m_e e^2}$ is the Bohr radius.

Energy: Consider dimensionless energies $\tilde{H} = \frac{H}{E_0}, \tilde{E} = \frac{E}{E_0}$, where $E_0 = \frac{e^2}{a_0} = \frac{m_e e^4}{\hbar^2}$.

Choosing a new length scale is a coordinate change. The partial derivatives transform as follows:

$$\begin{aligned} \frac{\partial}{\partial r_j^\mu} &= \frac{1}{a_0} \frac{\partial}{\partial \tilde{r}_j^\mu} \quad \text{and} \quad \frac{\partial}{\partial R_j^\mu} = \frac{1}{a_0} \frac{\partial}{\partial \tilde{R}_j^\mu} \\ \Rightarrow \quad \Delta_{\mathbf{r}_j} &= \frac{1}{a_0^2} \Delta_{\tilde{\mathbf{r}}_j} \quad \text{and} \quad \Delta_{\mathbf{R}_j} = \frac{1}{a_0^2} \Delta_{\tilde{\mathbf{R}}_j} . \end{aligned}$$

Plugging all in, leads to:

$$\tilde{V}_{aa} = \sum_{j < k} \frac{Z_j Z_k}{\|\tilde{\mathbf{R}}_j - \tilde{\mathbf{R}}_k\|} , \quad \tilde{V}_{ea} = - \sum_{j,k} \frac{Z_j}{\|\tilde{\mathbf{R}}_j - \tilde{\mathbf{r}}_k\|} , \quad \tilde{V}_{ee} = \sum_{j < k} \frac{1}{\|\tilde{\mathbf{r}}_j - \tilde{\mathbf{r}}_k\|}$$

$$\tilde{T}_e = -\frac{1}{2} \sum_{j=1}^{N_e} \Delta_{\tilde{\mathbf{r}}_j} \quad \text{and} \quad \boxed{\tilde{T}_a = -\frac{1}{2} \sum_{j=1}^{N_a} \frac{m_e}{M_j} \Delta_{\tilde{\mathbf{R}}_j}} . \quad (2.1)$$

Hence the solid Hamiltonian in atomic units reads:

$$\begin{aligned} \widetilde{H} = & \underbrace{\frac{1}{2} \sum_{j=1}^{N_e} -\Delta_{\widetilde{\mathbf{r}}_j} + \sum_{j<k} \frac{1}{\|\widetilde{\mathbf{r}}_j - \widetilde{\mathbf{r}}_k\|}}_{\text{electrons}} + \underbrace{\frac{1}{2} \sum_{j=1}^{N_a} \frac{m_e}{M_j} (-\Delta_{\widetilde{\mathbf{R}}_j}) + \sum_{j<k} \frac{Z_j Z_k}{\|\widetilde{\mathbf{R}}_j - \widetilde{\mathbf{R}}_k\|}}_{\text{ions}} \\ & + \underbrace{\sum_{j,k} \frac{Z_j}{\|\widetilde{\mathbf{R}}_j - \widetilde{\mathbf{r}}_k\|}}_{\text{interaction}} . \end{aligned}$$

In the following, we will drop the tilde notation, as we will stay in atomic units.

2.2 Born-Oppenheimer approximation

To decouple the electrons from the lattice, we observe the coefficient $\frac{m_e}{M_j}$ for the kinetic energy of the ions in (2.1). Even for the hydrogen atom, it is of magnitude $\approx 10^{-3}$. This suggests to treat the kinetic energy of the ions as perturbation, defining H_0 by:

$$H = H_0 + T_a ,$$

$$H_0 = T_e + V_{ee} + V_{aa} + V_{ea} .$$

The stationary Schrödinger equation for H_0 does not involve partial derivatives of $\vec{\mathbf{R}}$, as T_a is excluded. This means, that H_0 describes the solid, with fixed ion positions. Assume now, the stationary Schrödinger equation has been solved for H_0 :

$$H_0 \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}}) = \mathcal{E}_\alpha(\vec{\mathbf{R}}) \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}}) .$$

Here, α is a shorthand notation for the quantum numbers of the electrons. Note that the ion positions $\vec{\mathbf{R}}$ are parameters in this equation. For each ion configuration $\vec{\mathbf{R}}$, the wave functions $\{\phi_\alpha\}$ are assumed to be a Hilber basis, such that:

$$\psi(\vec{\mathbf{r}}, \vec{\mathbf{R}}) = \sum_{\alpha} \chi_\alpha(\vec{\mathbf{R}}) \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}}) .$$

Plugging in, into the full equation leads to:

$$\begin{aligned} (H - E)\psi(\vec{\mathbf{r}}, \vec{\mathbf{R}}) &= (H_0 + T_a - E) \sum_{\alpha} \chi_\alpha(\vec{\mathbf{R}}) \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}}) \\ &= \sum_{\alpha} \chi_\alpha(\vec{\mathbf{R}}) H_0 \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}}) + T_a (\chi_\alpha(\vec{\mathbf{R}}) \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}})) \\ &\quad - E \chi_\alpha(\vec{\mathbf{R}}) \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}}) \\ &= \sum_{\alpha} (\mathcal{E}_\alpha(\vec{\mathbf{R}}) - E) \chi_\alpha(\vec{\mathbf{R}}) \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}}) + T_a (\chi_\alpha(\vec{\mathbf{R}}) \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}})) \\ &= 0 . \end{aligned}$$

Plugging T_a from (2.1) in, it follows that:

$$T_a \chi_\alpha(\vec{\mathbf{R}}) \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}}) = [T_a \chi_\alpha(\vec{\mathbf{R}})] \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}}) + \chi_\alpha(\vec{\mathbf{R}}) T_a \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}})$$

$$\begin{aligned}
& - \sum_{j=1}^{N_a} \sum_{k=1}^n \frac{m_e}{M_j} \left(\frac{\partial}{\partial R_j^\mu} \chi_\alpha(\vec{\mathbf{R}}) \right) \left(\frac{\partial}{\partial R_j^\mu} \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}}) \right) \\
& =: [T_a \chi_\alpha(\vec{\mathbf{R}})] \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}}) + B_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}}) .
\end{aligned}$$

Hence:

$$0 = \sum_{\alpha} \left(\mathcal{E}_\alpha(\vec{\mathbf{R}}) \chi_\alpha(\vec{\mathbf{R}}) + T_a \chi_\alpha(\vec{\mathbf{R}}) - E \chi_\alpha(\vec{\mathbf{R}}) \right) \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}}) + B_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}}) . \quad (2.2)$$

Since $\{\phi_\alpha\}$ are a Hilbert basis ($\vec{\mathbf{R}}$ is a parameter), it holds that

$$\int d\vec{\mathbf{r}} \overline{\phi_\beta(\vec{\mathbf{r}}, \vec{\mathbf{R}})} \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}}) = \delta_{\alpha\beta} ,$$

where $d\vec{\mathbf{r}}$ is a shorthand notation for $dr_{11} dr_{12} \dots dr_{N_e n}$. Thus, acting with the integration operator $\int d\vec{\mathbf{r}} \overline{\phi_\beta(\vec{\mathbf{r}}, \vec{\mathbf{R}})}$ on both sides of (2.2) leads to¹

$$0 = \left(\mathcal{E}_\beta(\vec{\mathbf{R}}) \chi_\beta(\vec{\mathbf{R}}) + T_a \chi_\beta(\vec{\mathbf{R}}) - E \chi_\beta(\vec{\mathbf{R}}) \right) + \sum_{\alpha} A_{\alpha\beta}(\vec{\mathbf{R}}) ,$$

where $A_{\alpha\beta}(\vec{\mathbf{R}}) = \int d\vec{\mathbf{r}} \overline{\phi_\beta(\vec{\mathbf{r}}, \vec{\mathbf{R}})} B_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}})$.

Definition 2.2.1.

The **Born-Oppenheimer approximation** is to neglect the terms $A_{\alpha\beta}(\vec{\mathbf{R}})$. One obtains a sort of Schrödinger equation

$$\left(T_a + \mathcal{E}_\beta(\vec{\mathbf{R}}) \right) \chi_\beta(\vec{\mathbf{R}}) = E \chi_\beta(\vec{\mathbf{R}})$$

for $\chi_\beta(\vec{\mathbf{R}})$ with effective potential $\mathcal{E}_\beta(\vec{\mathbf{R}})$.

Remark 2.2.2 (Summary).

1. Obtain $\mathcal{E}_\alpha(\vec{\mathbf{R}})$ by solving

$$H_0 \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}}) = \mathcal{E}_\alpha(\vec{\mathbf{R}}) \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}}) .$$

2. Find the development coefficients $\chi_\alpha(\vec{\mathbf{R}})$ by solving

$$\left(T_a + \mathcal{E}_\beta(\vec{\mathbf{R}}) \right) \chi_\alpha(\vec{\mathbf{R}}) + \sum_{\gamma} A_{\gamma\alpha}(\vec{\mathbf{R}}) = E \chi_\alpha(\vec{\mathbf{R}})$$

or neglect $\sum_{\gamma} A_{\gamma\alpha}(\vec{\mathbf{R}})$ for the Born-Oppenheimer approximation.

3. The wave function for the full equation $H\psi = E\psi$ is given by:

$$\psi(\vec{\mathbf{r}}, \vec{\mathbf{R}}) = \sum_{\alpha} \chi_\alpha(\vec{\mathbf{R}}) \phi_\alpha(\vec{\mathbf{r}}, \vec{\mathbf{R}}) .$$

¹ $\left(\mathcal{E}_\alpha(\vec{\mathbf{R}}) \chi_\alpha(\vec{\mathbf{R}}) + T_a \chi_\alpha(\vec{\mathbf{R}}) - E \chi_\alpha(\vec{\mathbf{R}}) \right)$ does not depend on $\vec{\mathbf{r}}$.

3

Lattice vibrations

In this chapter, an approximation by harmonic oscillators and the translational symmetries of the solid are used, to find a solution for Schrödinger equations that describe the ions in a solid. This leads to the definition of phonons, quasi particles that correspond to lattice vibrations that are quantized, because of the finite size of real solids.

3.1 Harmonic approximation

We consider the Schrödinger equation for the ions

$$H\psi(\vec{\mathbf{R}}) = E\psi(\vec{\mathbf{R}}) .$$

Note, that for $M_j = \frac{m_e}{h^2 M_j}$ and $V(\vec{\mathbf{R}}) = \mathcal{E}_\alpha(\vec{\mathbf{R}})$ one obtains the equation of the Born-Oppenheimer approximation from the generic Hamiltonian. The general problem of the large number of particles remains, such that one needs further approximations.

Start with the classical Hamilton function

$$H(\vec{\mathbf{P}}, \vec{\mathbf{R}}) = \sum_{j=1}^{N_a} \frac{\mathbf{P}_j^2}{2M_j} + V(\vec{\mathbf{R}}) .$$

Assume, that the potential V has a global minimum in $\vec{\mathfrak{R}}$, that is also a local minimum. This means that

$$\left. \frac{\partial}{\partial R_j^\mu} V(\vec{\mathbf{R}}) \right|_{\vec{\mathbf{R}}=\vec{\mathfrak{R}}} = 0 .$$

Hence the Taylor expansion of lowest non-trivial order (**harmonic approximation**) reads ($\mathbf{u}_j = \mathbf{R}_j - \mathfrak{R}_j$):

$$V(\vec{\mathbf{R}}) \approx V(\vec{\mathfrak{R}}) + \frac{1}{2} \sum_{j,\mu;\ell,\nu} \Phi_{j\ell}^{\mu\nu} u_j^\mu u_\ell^\nu ,$$

where $\Phi_{j\ell}^{\mu\nu}$ are the components of the Hessian matrix:

$$\Phi_{j\ell}^{\mu\nu} = \left. \frac{\partial^2}{\partial R_j^\mu \partial R_\ell^\nu} V(\vec{\mathbf{R}}) \right|_{\vec{\mathbf{R}}=\vec{\mathfrak{R}}} .$$

By the freedom of potential energy, we can choose the classical zero point energy to be

$$V(\vec{\mathfrak{R}}) = 0 \quad \Rightarrow \quad V(\vec{\mathbf{R}}) = \frac{1}{2} \sum_{j,\mu;\ell,\nu} \Phi_{j\ell}^{\mu\nu} u_j^\mu u_\ell^\nu .$$

Definition 3.1.1.

The **dynamical matrix** $D_{j\ell}^{\mu\nu}$ is defined by

$$D_{j\ell}^{\mu\nu} := \frac{1}{\sqrt{M_j M_\ell}} \Phi_{j\ell}^{\mu\nu} ,$$

where M_j is the mass of the j -th ion.

By the condition, that $\vec{\mathfrak{R}}$ is a local minimum, it follows, that $\Phi_{j\ell}^{\mu\nu}$ and hence $D_{j\ell}^{\mu\nu}$ are positive definite $n \cdot N_e \times n \cdot N_e$ matrices. The reason, to define the dynamical matrix, is that the Hamilton function in harmonic approximation can be written as

$$H(\vec{P}, \vec{v}) = \frac{1}{2} \vec{P}'^T \mathbb{1} \vec{P}' + \frac{1}{2} \vec{u}'^T D \vec{u}' ,$$

where the tuples \vec{P}' and \vec{u}' are defined as follows:

$$\vec{P}' = \left(\frac{1}{\sqrt{M_1}} P_1^1, \frac{1}{\sqrt{M_1}} P_1^2, \dots, \frac{1}{\sqrt{M_{N_a}}} P_{N_a}^n \right) ,$$

$$\vec{u}' = \left(\sqrt{M_1} u_1^1, \sqrt{M_1} u_1^2, \dots, \sqrt{M_{N_a}} u_{N_a}^n \right) .$$

It is useful, to change the index conventions for the rest of this section, using only one index j , running from 1 to $n \cdot N_a$.

For real, $C^3(\mathbb{R}^{n \cdot N_e})$ potentials, the dynamical matrix is real and symmetric because of the symmetry of partial derivatives. Then, by the principal axis theorem, there is an orthogonal matrix C , i.e. $C^T = C^{-1}$, such that

$$C D C^T = \Omega \quad \text{with} \quad \Omega = \text{diag}(\omega_1^2, \dots, \omega_{n \cdot N_e}^2) .$$

Defining $\vec{p} = C \vec{P}'$ and $\vec{u} = C \vec{u}'$, we observe that:

$$\begin{aligned} H(\vec{P}, \vec{u}) &= \frac{1}{2} \vec{P}'^T \mathbb{1} \vec{P}' + \frac{1}{2} \vec{u}'^T D \vec{u}' \\ &= \frac{1}{2} \vec{P}'^T C^T C \vec{P}' + \frac{1}{2} \vec{u}'^T C^T C D C^T C \vec{u}' \\ &= \boxed{\frac{1}{2} \vec{p}^T \mathbb{1} \vec{p} + \frac{1}{2} \vec{u}^T \Omega \vec{u} = H(\vec{p}, \vec{u})} \end{aligned} \quad (3.1)$$

Passing to quantum mechanics (but keeping the linear indexing j for the tuples \vec{P} and \vec{v}), it holds that¹

$$[v_j, v_\ell] = 0 = [P_j, P_\ell] \quad \text{and} \quad [v_j, P_\ell] = i\hbar \delta_{j\ell}$$

We calculate:

$$\begin{aligned} [\mathbf{u}_j, \mathbf{p}_\ell] &= \left[\sum_\alpha C_{j\alpha} \sqrt{M_\alpha} v_\alpha, \sum_\beta C_{\ell\beta} \frac{1}{\sqrt{M_\beta}} P_\beta \right] \\ &= \sum_{\alpha\beta} C_{j\alpha} C_{\ell\beta} \sqrt{\frac{M_\alpha}{M_\beta}} \underbrace{[v_\alpha, P_\beta]}_{=i\hbar \delta_{\alpha\beta}} \end{aligned}$$

¹In fact, one needs to use the commutator relations for P_j and R_j . But since $v_j = R_j - \mathfrak{R}_j$ and the commutator is bilinear, the relations carry over to v_j .

$$= i\hbar \sum_{\alpha} C_{j\alpha} C_{\ell\alpha} = i\hbar \sum_{\alpha} C_{j\alpha} C_{\alpha\ell}^T = i\hbar (\mathbb{1})_{j\ell} = i\hbar \delta_{j\ell} .$$

The commutators $[\mathbf{u}_j, \mathbf{u}_{\ell}] = 0 = [\mathbf{p}_j, \mathbf{p}_{\ell}]$ can be calculated as before. Writing (3.1) in coordinates:

$$H(\vec{\mathbf{p}}, \vec{\mathbf{u}}) = \frac{1}{2} \sum_{j=1}^{n \cdot N_a} \mathbf{p}_j^2 + \omega_j^2 \mathbf{u}_j^2 = \sum_j H_j ,$$

we observe, that $H_{j\mu}$ is a harmonic oscillator with mass $m = 1$. Since \mathbf{u}_j and \mathbf{p}_j satisfy the canonical commutation relations, the standard approach for the harmonic oscillator can be applied here:

$$b_j = \sqrt{\frac{\omega_j}{2\hbar}} \mathbf{u}_j + i \sqrt{\frac{1}{2\hbar\omega_j}} \mathbf{p}_j .$$

$$\Rightarrow H_j = \hbar\omega_j (b_j^{\dagger} b_j + \frac{1}{2}) \quad \Rightarrow \quad \boxed{H = \sum_{j=1}^{n \cdot N_a} \hbar\omega_j (b_j^{\dagger} b_j + \frac{1}{2})} . \quad (3.2)$$

Remark 3.1.2.

The operators b_j^{\dagger}, b_j are bosonic creation/annihilation operators, satisfying the canonical commutation relations:

$$[b_j, b_{\ell}] = 0 = [b_j^{\dagger}, b_{\ell}^{\dagger}] \quad \text{and} \quad [b_j, b_{\ell}^{\dagger}] = \delta_{j\ell} .$$

In fact, we have obtained the second quantized Hamiltonian, by the means of canonical quantization.

3.2 Periodic boundary conditions

Here, we deviate from [Czy16] and use the Bloch theorem to characterize the eigen vectors /states of the ion Hamiltonian, as in [Pru13].

3.2.1 Translation operator

So far, we have made no restrictions for the equilibrium positions $\vec{\mathfrak{R}}$. However, since we want to describe solids that have a crystalline structure, the equilibrium positions of the ions form a crystal. The positions of the ions in the excited state are $\vec{\mathbf{R}} = \vec{\mathfrak{R}} + \vec{\mathbf{u}}$, i.e. oscillations around the equilibrium positions.

To use translational invariance, we need to extend the structure throughout the whole space, setting $N_a \rightarrow \infty$. Adopting for the crystalline structure, we change the notation again: We write:

$$\mathfrak{R}_{j,\alpha}^{\mu} = \mathfrak{R}_j^{\mu} + \mathbf{b}_{\alpha}^{\mu} ,$$

where μ describes the coordinate, \mathfrak{R}_j^{μ} is a Bravais lattice vector and $\mathbf{b}_{\alpha}^{\mu}$ an atom position in the unit cell. This choice is possible, as no other restriction, that conflicts with

this choice, has been made before. We recall, that adding a Bravais vector $\mathbf{B} \in L$ to a Bravais lattice point $\mathfrak{R}_{\vec{j}}$, one obtains another lattice point:

$$\mathfrak{R}_{\vec{j}} + \mathbf{B} = \mathfrak{R}_{\vec{k}} .$$

Since we have (implicitly) chosen an origin for the Bravais lattice, lattice points and Bravais vectors can be identified. Thus, there is an \vec{j}' , such that $\mathbf{B} = \mathfrak{R}_{\vec{j}'}$, which leads to the following notation:

$$\boxed{\mathfrak{R}_{\vec{j}} + \mathfrak{R}_{\vec{j}'} = \mathfrak{R}_{\vec{k}} \equiv \mathfrak{R}_{\vec{j}+\vec{j}'} .} \quad (3.3)$$

Definition 3.2.1.

We define the translation operator $T_{\vec{j}}$ for a Bravais vector $\mathfrak{R}_{\vec{j}}$ by

$$T_{\vec{j}}f(\mathbf{r}) = f(\mathbf{r} + \mathfrak{R}_{\vec{j}}) .$$

With 3.3, we observe that the notation $T_{\vec{j}+\vec{\ell}}$ is well defined, and calculate:

$$\begin{aligned} T_{\vec{j}}T_{\vec{\ell}}f(\mathbf{r}) &= f(\mathbf{r} + \mathfrak{R}_{\vec{j}} + \mathfrak{R}_{\vec{\ell}}) = f(\mathbf{r} + \mathfrak{R}_{\vec{j}+\vec{\ell}}) = T_{\vec{j}+\vec{\ell}}f(\mathbf{r}) , \\ \Rightarrow \quad &\boxed{T_{\vec{j}} \circ T_{\vec{\ell}} = T_{\vec{j}+\vec{\ell}}} . \end{aligned}$$

3.2.2 Bloch's theorem

Let A be an operator that commutes with any translation operator $T_{\vec{j}}$, i.e. $[A, T_{\vec{j}}] = 0$. Then there are common eigen states, i.e.

$$A\psi(\mathbf{r}) = a\psi(\mathbf{r}) \quad \text{and} \quad T_{\vec{j}}\psi(\mathbf{r}) = c(\mathfrak{R}_{\vec{j}})\psi(\mathbf{r}) .$$

For the eigen values $c(\mathfrak{R}_{\vec{j}})$, the following holds:

$$\begin{aligned} c(\mathfrak{R}_{\vec{j}})c(\mathfrak{R}_{\vec{\ell}})\psi(\mathbf{r}) &= T_{\vec{j}}T_{\vec{\ell}}\psi(\mathbf{r}) = T_{\vec{j}+\vec{\ell}}\psi(\mathbf{r}) = c(\mathfrak{R}_{\vec{j}+\vec{\ell}})\psi(\mathbf{r}) , \\ \Rightarrow \quad &c(\mathfrak{R}_{\vec{j}})c(\mathfrak{R}_{\vec{\ell}}) = c(\mathfrak{R}_{\vec{j}+\vec{\ell}}) . \end{aligned}$$

As usual in quantum mechanics, we normalize the states as follows:

$$\begin{aligned} 1 &= \int dr^n |\psi(\mathbf{r})|^2 = \int dr^n |\psi(\mathbf{r} + \mathfrak{R}_{\vec{j}})|^2 = |c(\mathfrak{R}_{\vec{j}})|^2 \int dr^n |\psi(\mathbf{r})|^2 \\ &= |c(\mathfrak{R}_{\vec{j}})|^2 \end{aligned}$$

This leads to the conclusion

$$c(\mathfrak{R}_{\vec{j}}) = e^{i\langle \mathbf{k}, \mathfrak{R}_{\vec{j}} \rangle} \quad \text{with} \quad \mathbf{k} \in \mathbb{R}^n .$$

Because $e^{i\langle \mathbf{G}, \mathfrak{R}_{\vec{j}} \rangle} = 1$ for all reciprocal lattice vectors $\mathbf{G} \in \tilde{L}$, one can reduce \mathbf{k} to a primitive cell in the reciprocal lattice, e.g. the unit cell or the Brillouin zone. Because of this, \mathbf{k} becomes interpreted as momentum.

It should be noted, that \mathbf{k} is not uniquely determined. It is rather a parameter, that distinguishes the ψ further, like a quantum number. Putting all together, one obtains the **Bloch condition**

$$\psi_{\mathbf{k}}(\mathbf{r} + \mathfrak{R}_{\vec{j}}) = e^{i\langle \mathbf{k}, \mathfrak{R}_{\vec{j}} \rangle} \psi_{\mathbf{k}}(\mathbf{r}) . \quad (3.4)$$

The $\psi_{\mathbf{k}}$ can be written in terms of periodic functions $u_{\mathbf{k}}$, i.e. $u_{\mathbf{k}}(\mathbf{r}) = u_{\mathbf{k}}(\mathbf{r} + \mathfrak{R}_{\vec{j}})$ for all $\mathfrak{R}_{\vec{j}}$, called **Bloch functions**:

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\langle \mathbf{k}, \mathbf{r} \rangle} u_{\mathbf{k}}(\mathbf{r}) . \quad (3.5)$$

3.2.3 Periodic boundary conditions

To use translational invariance, we had to set $N_a \rightarrow \infty$. However, real solids have finite size. To keep both, the finite nature of real solids as well as the computational convenience of Bravais lattices, the finite lattice is repeated periodically, with respective boundary conditions. The construction of a finite crystal from the infinite model is

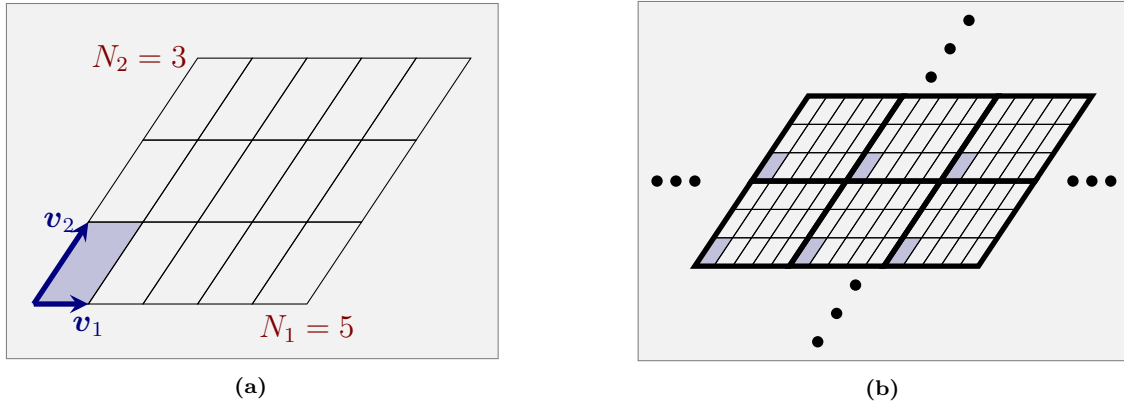


Figure 3.1: Example of finite lattice 3.1a and its periodic continuation 3.1b.

straightforward. Consider the Bravais lattice $L = \{\sum_{j=1}^n \mathbf{v}_j m_j \mid m_j \in \mathbb{Z}\}$ that describes the solid. Then repeat the unit cell N_j times in the direction \mathbf{v}_j . One obtains a finite lattice \bar{L} . This lattice has $\prod_{j=1}^n N_j$ unit cells. If there are d atoms per unit cell, the number of ions N_a is given by

$$N_a = d \cdot \prod_{j=1}^n N_j$$

One recovers the Bravais lattice by repeating the finite lattice. For functions of lattice points $f(\mathbf{R}_{\vec{m}})$, this leads to **periodic boundary conditions**:

$$f(\mathbf{R}_{\vec{m}} + N_j \mathbf{v}_j) = f(\mathbf{R}_{\vec{m}}) \quad \forall j = 1, \dots, n .$$

Since $N_j \mathbf{v}_j \in L$, the periodic boundary conditions have the following consequences for Bloch functions $\psi_{\mathbf{k}}$:

$$\begin{aligned} \psi_{\mathbf{k}}(\mathbf{r} + N_j \mathbf{v}_j) &= e^{i\langle \mathbf{k}, N_j \mathbf{v}_j \rangle} \psi_{\mathbf{k}}(\mathbf{r}) \stackrel{!}{=} \psi_{\mathbf{k}}(\mathbf{r}) , \\ \Rightarrow e^{i\langle \mathbf{k}, N_j \mathbf{v}_j \rangle} &= 1 \quad \Rightarrow \quad N_j \langle \mathbf{k}, \mathbf{v}_j \rangle \in 2\pi\mathbb{Z} . \end{aligned}$$

This has the following important implications:

- The momentum \vec{k} has become discrete.
- The momentum can be written as $\mathbf{k} = \sum_j k_j \mathbf{g}_j$ with reciprocal primitive vectors $\mathbf{g}_j \in \tilde{L}$. As we have seen, the momenta can be reduced to a primitive cell. We chose the unit cell and the Brillouin cell here, because the restriction for the k_j becomes easy in these cases:

$$\text{uni cell: } k_j \in [0, 1) \quad \text{B.z.: } k_j \in [-\frac{1}{2}, \frac{1}{2}) \quad \forall j .$$

Combining these observations by plugging in, we find:

$$2\pi\mathbb{Z} \ni N_j \langle \mathbf{k}, \mathbf{v}_j \rangle = N_j \sum_{\ell} k_{\ell} \langle \mathbf{g}_{\ell}, \mathbf{v}_j \rangle = 2\pi N_j k_j$$

$$\Rightarrow \quad k_j = \frac{m}{N_j} \quad \text{for } m < N_j \in \mathbb{N}_0 .$$

For a finite lattice of lengths N_1, \dots, N_j , the possible momenta are given by

$$\boxed{\mathbf{k}_{\vec{m}} = \sum_{j=1}^n \frac{m_j}{N_j} \mathbf{g}_j} \quad (3.6)$$

$$\text{unit cell: } m_j = 0, 1, \dots, N_j - 1 \quad \text{B.z.: } m_j = -\frac{N_j}{2}, -\frac{N_j}{2} + 1, \dots, \frac{N_j}{2} - 1 .$$

3.2.4 Diagonalization of the dynamical matrix

To obtain (3.2), we have used, that the dynamical matrix is diagonalizable. However, existence of a diagonalization does not give an algorithm. In fact, without further considerations, diagonalizing an $n \cdot N_a \times n \cdot N_a$ matrix is impossible for contemporary computers. Yet, translational invariance can be used to reduce the problem to the diagonalization of an $d \cdot n \times d \cdot n$ -matrix, where d is the number of atoms in the unit cell.

Applying the notation of subsection 3.2.1, the dynamical matrix reads:

$$D_{\vec{j}, \alpha; \vec{\ell}, \beta}^{\mu\nu} = \frac{1}{\sqrt{M_{\alpha} M_{\beta}}} \frac{\partial^2}{\partial R_{\vec{j}, \alpha}^{\mu} \partial R_{\vec{\ell}, \beta}^{\nu}} V(\vec{\mathbf{R}}) \Big|_{\vec{\mathbf{R}} = \vec{\mathfrak{R}}} .$$

There are two things to observe here. The first is, that the masses only depend on the atom index, because of the periodicity of the lattice. The second is, that the dynamical matrix coefficient $D_{\vec{j}, \alpha; \vec{\ell}, \beta}^{\mu\nu}$ depends on the indices $\vec{j} \in \mathbb{Z}^n$ and $\vec{\ell} \in \mathbb{Z}^n$, which correspond to Bravais vectors $\mathfrak{R}_{\vec{j}}$ and $\mathfrak{R}_{\vec{\ell}}$:

$$\boxed{D_{\vec{j}, \alpha; \vec{\ell}, \beta}^{\mu\nu} = D_{\alpha; \beta}^{\mu\nu}(\vec{j}, \vec{\ell})}$$

Remark 3.2.2.

Because of the correspondence $\vec{j} \leftrightarrow \mathfrak{R}_{\vec{j}}$, it is common practice to write:

$$D_{\alpha;\beta}^{\mu\nu}(\vec{j}, \vec{\ell}) = D_{\alpha;\beta}^{\mu\nu}(\mathfrak{R}_{\vec{j}}, \mathfrak{R}_{\vec{\ell}}) .$$

This is similar to the vector space isomorphism $V \rightarrow \mathbb{R}^n$ defined by $\sum_i a_i \mathbf{v}_i \rightarrow (a_1, \dots, a_n)$. In fact, it is a modul isomorphism $I: L \rightarrow \mathbb{Z}^n$. Formally we actually have

$$D_{\alpha;\beta}^{\mu\nu}(\mathfrak{R}_{\vec{j}}, \mathfrak{R}_{\vec{\ell}}) = (D_{\alpha;\beta}^{\mu\nu} \circ I)(\mathfrak{R}_{\vec{j}}, \mathfrak{R}_{\vec{\ell}}) .$$

This notation should be used with care, as the position dependence of the dynamical matrix is

$$D_{\vec{j}, \alpha; \vec{\ell}, \beta}^{\mu\nu}(\vec{\mathfrak{R}})$$

The translational invariance of the Bravais lattice means, that translating every point by an Bravais vector, the lattice still looks the same. For the dynamical matrix, this means:

$$D_{\alpha;\beta}^{\mu\nu}(\vec{j}, \vec{\ell}) = D_{\alpha;\beta}^{\mu\nu}(\vec{j} + \vec{a}, \vec{\ell} + \vec{a}) \quad \forall \vec{a} \in \mathbb{Z}^n .$$

Choosing $\vec{a} = -\vec{\ell}$, we see that $D_{\alpha;\beta}^{\mu\nu}(\vec{j}, \vec{\ell})$ depends only on the difference:

$$D_{\alpha;\beta}^{\mu\nu}(\vec{j}, \vec{\ell}) = D_{\alpha;\beta}^{\mu\nu}(\vec{j} - \vec{\ell}, 0) \equiv D_{\alpha;\beta}^{\mu\nu}(\vec{j} - \vec{\ell}) .$$

We obtain the following $n \cdot d \times n \cdot d$ matrix, for $\vec{m} = \vec{j} - \vec{\ell}$:

$$D(\vec{m}) = \begin{pmatrix} D_{1;1}^{11}(\vec{m}) & D_{1;1}^{12}(\vec{m}) & \dots & D_{1;n}^{1d}(\vec{m}) \\ D_{1;1}^{21}(\vec{m}) & D_{1;1}^{22}(\vec{m}) & \dots & D_{2;n}^{1;d}(\vec{m}) \\ \vdots & & \ddots & \vdots \\ D_{n;1}^{d1}(\vec{m}) & D_{n;1}^{d2}(\vec{m}) & \dots & D_{n;n}^{dd}(\vec{m}) \end{pmatrix}$$

To diagonalize the matrix $(D_{\vec{j}, \alpha; \vec{\ell}, \beta}^{\mu\nu})$, we have to solve the following eigen value equation:

$$\lambda a_{\vec{j}, \alpha}^{\mu} = \sum_{\vec{\ell}, \beta, \nu} D_{\alpha;\beta}^{\mu\nu}(\vec{j} - \vec{\ell}) a_{\vec{\ell}, \beta}^{\nu} .$$

Following [Pru13] we make the ansatz (using the correspondence of remark 3.2.2):

$$a_{\vec{j}, \alpha}^{\mu} = \psi^{\mu}(\mathbf{b}_{\alpha} + \mathfrak{R}_{\vec{j}}) = T_{\vec{j}} \psi^{\mu}(\mathbf{b}_{\alpha}) .$$

Since the dynamical matrix is a derived quantity from the potential V , it commutes with the translation operators $T_{\vec{j}}$, because of $T_{\vec{j}} V(\vec{\mathbf{R}}) = V(\vec{\mathbf{R}} + \mathfrak{R}_{\vec{j}}) = V(\vec{\mathbf{R}})$. The, eigen states of $D(\vec{m})$ thus can be chosen to be also eigen states of $T_{\vec{j}}$. By Bloch's theorem we obtain:

$$a_{\vec{j}, \alpha}^{\mu} = e^{i\langle \mathbf{k}, \mathfrak{R}_{\vec{j}} \rangle} \psi_{\mathbf{k}}^{\mu}(\mathbf{b}_{\alpha}) \equiv e^{i\langle \mathbf{k}, \mathfrak{R}_{\vec{j}} \rangle} \psi_{\alpha, \mathbf{k}}^{\mu} .$$

Plugging in, into the eigen value equation, writing the eigen values in the form $\lambda = \omega^2$ and using the quantum number \mathbf{k} , it follows that:

$$\Leftrightarrow \omega^2 \psi_{\alpha, \mathbf{k}}^{\mu} = \sum_{\beta, \nu} \psi_{\beta, \mathbf{k}}^{\nu} \sum_{\vec{\ell}} D_{\alpha;\beta}^{\mu\nu}(\vec{j} - \vec{\ell}) e^{-i\langle \mathbf{k}, \mathfrak{R}_{\vec{j}} - \mathfrak{R}_{\vec{\ell}} \rangle}$$

$$= \sum_{\beta, \nu} \psi_{\beta, \mathbf{k}}^{\nu} \sum_{\vec{m}} D_{\alpha; \beta}^{\mu\nu}(\vec{m}) e^{-i\langle \mathbf{k}, \mathfrak{R}_{\vec{m}} \rangle} . \quad (3.7)$$

In the second line, we have used, that there is an $\vec{m} \in \mathbb{Z}^n$, such that $\mathfrak{R}_{\vec{j}} - \mathfrak{R}_{\vec{\ell}} = \mathfrak{R}_{\vec{m}}$. Also, because the sum over $\vec{\ell}$ covers all indices \mathbb{Z}^n , we could also sum over \vec{m} . The term

$$\sum_{\vec{m}} D_{\alpha; \beta}^{\mu\nu}(\vec{m}) e^{-i\langle \mathbf{k}, \mathfrak{R}_{\vec{m}} \rangle}$$

is a discrete inverse Fourier series:

$$\widetilde{D}_{\alpha; \beta}^{\mu\nu}(\mathbf{k}) = \sum_{\vec{m}} D_{\alpha; \beta}^{\mu\nu}(\mathfrak{R}_{\vec{m}}) e^{-i\langle \mathbf{k}, \mathfrak{R}_{\vec{m}} \rangle} .$$

This further motivates the interpretation of \mathbf{k} as momentum. The components $\widetilde{D}_{\alpha; \beta}^{\mu\nu}(\mathbf{k})$ define a matrix $\widetilde{D}(\mathbf{k})$. Plugging in, into (3.7):

$$\omega^2(\mathbf{k}) \psi_{\alpha, \mathbf{k}}^{\mu} = \sum_{\beta, \nu} \widetilde{D}_{\alpha; \beta}^{\mu\nu}(\mathbf{k}) \psi_{\beta, \mathbf{k}}^{\nu}$$

This is again an eigen value equation. However, it is now an eigen value equation of a $n \cdot d \times n \cdot d$ matrix for each \mathbf{k} .

3.2.5 Ion Hamiltonian

Using the notation and results of this section, the ion Hamiltonian (3.2) becomes:

$$H = \sum_{\mathbf{k}} \sum_{j=1}^{n \cdot d} \hbar \omega_j(\mathbf{k}) (b_{j, \mathbf{k}}^{\dagger} b_{j, \mathbf{k}} + \frac{1}{2}) . \quad (3.8)$$

The \mathbf{k} are momenta, wave vectors to be precise, that are given by (3.6), if one chooses the unit cell as confinement. Here, the index j does not label the Bravais index, but the dimension and the atom in the unit cell.

3.3 Phonons

Usually in second quantization, the creation and annihilation operators create/annihilate particles. Here one adopts that picture, imagining quasi particles:

Definition 3.3.1.

The quasi particles created/annihilated by $b_{j, \mathbf{k}}^{\dagger}/b_{j, \mathbf{k}}$ are called **phonons**.

Owing to the commutation relations, the phonons are bosons. For each atom and each dimension, there is one phonon, labeled by the index j . Phonons, as created by $b_{j, \mathbf{k}}^{\dagger}$, also carry a momentum \mathbf{k} . However, one should be careful with this imagery, as phonons are no real particles. There is no conservation of particles for example. Hence the word quasi particle. It is best to think of phonons as vibration excitations of the atoms in the solid.

3.3.1 Phonon branches

The Hamiltonian (3.8) shows that for each phonon, there is a **dispersion relation** $\omega_j(\mathbf{k})$, i.e. a relation between the wave vector \mathbf{k} and the angular frequency ω .

Definition 3.3.2.

A dispersion relation $\omega_j(\mathbf{k})$ for a phonon is called **phonon branch**. A phonon branch that satisfies

$$\omega_j(\mathbf{k}) \rightarrow 0 \quad \mathbf{k} \rightarrow 0$$

is called **acoustic branch**. A phonon branch that satisfies

$$\omega_j(\mathbf{k}) \rightarrow \omega_j \neq 0 \quad \mathbf{k} \rightarrow 0$$

is called **optical branch**.

It can be shown², that for each dimension, there is an acoustic branch. So there are in

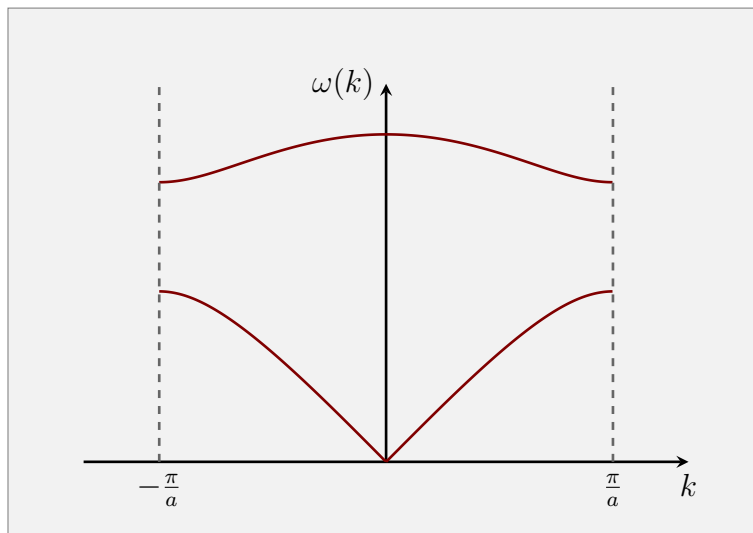


Figure 3.2: Example for phonon branches in a two-atomic linear chain with momenta restricted to the Brillouin zone. See [Dem10, section 12.1] for further information.

total n acoustic branches and $n \cdot d - n$ optical branches. Furthermore, it can be shown, that for acoustic branches, it holds that

$$\omega_{\text{ac}}(\mathbf{k}) \sim \|\mathbf{k}\| \quad \text{for } \|\mathbf{k}\| \rightarrow 0 .$$

3.3.2 Thermodynamics of phonons

To describe macroscopic, thermodynamical effects caused by lattice vibrations, such as the specific heat for example, the phonon description has to be brought in the language

²See for example [Czy16, pp. 52ff.]. The basic idea is, that all atoms are displaced in the same direction, but with possible different phases for different unit cells. Developing the classical equations of motions for small \mathbf{k} , one finds $\omega^2(\mathbf{k}) \rightarrow 0$ for $\mathbf{k} \rightarrow 0$ in these cases. But there are only n linearly independent directions in n dimensions.

of statistical mechanics. From second quantization we know, that the eigen states of (3.8) are the occupation number states

$$|\{n_{j,\mathbf{k}}\}\rangle := |n_{1,\mathbf{k}_1}, n_{1,\mathbf{k}_2}, \dots, n_{2,\mathbf{k}_1}, \dots\rangle$$

with the eigen values

$$E_{\{n_{j,\mathbf{k}}\}} = \sum_{\mathbf{k}} \sum_{j=1}^{n-d} \hbar\omega_j(\mathbf{k})(n_{j,\mathbf{k}} + \frac{1}{2}) .$$

The set $\{n_{j,\mathbf{k}}\}$ completely describes the quantum system and thus is a micro state in the sense of statistical mechanics. Thus, the description of phonons becomes a problem of statistical mechanics.

Internal energy

The internal energy is the expectation value of the $E_{\{n_{j,\mathbf{k}}\}}$, i.e. $U = \langle E_{\{n_{j,\mathbf{k}}\}} \rangle$. Phonons are non-interacting bosons. A result from statistical mechanics is the so called **Bose-Einstein statistics**:

$$\langle n_{j,\mathbf{k}} \rangle = \frac{1}{\exp\left(\frac{\hbar\omega_j(\mathbf{k})}{k_B T}\right) - 1} ,$$

where k_B is the Boltzmann constant and T the temperature. For the internal energy

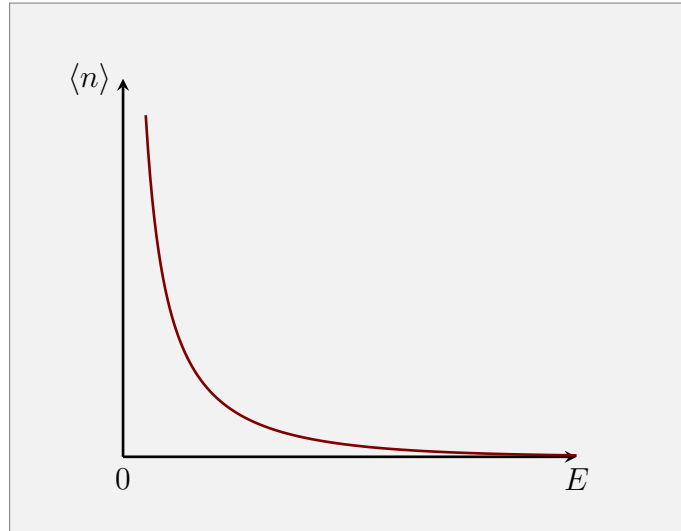


Figure 3.3: Energy dependence of the particle number expectation value for the Bose-Einstein statistics.

we calculate:

$$\begin{aligned} U &= \langle E_{\{n_{j,\mathbf{k}}\}} \rangle = \sum_{\mathbf{k}} \sum_{j=1}^{n-d} \hbar\omega_j(\mathbf{k})(\langle n_{j,\mathbf{k}} \rangle + \frac{1}{2}) \\ &= \frac{1}{2} \sum_{\mathbf{k}} \sum_{j=1}^{n-d} \hbar\omega_j(\mathbf{k}) + \sum_{\mathbf{k}} \sum_{j=1}^{n-d} \frac{\hbar\omega_j(\mathbf{k})}{\exp\left(\frac{\hbar\omega_j(\mathbf{k})}{k_B T}\right) - 1} \\ &\equiv E_0 + \sum_{\mathbf{k}} \sum_{j=1}^{n-d} \frac{\hbar\omega_j(\mathbf{k})}{\exp\left(\frac{\hbar\omega_j(\mathbf{k})}{k_B T}\right) - 1} \end{aligned}$$

Remark 3.3.3.

Here the quasi particle nature of phonons becomes noticeable. For acoustic phonons, the expectation value of zero momentum phonons becomes infinite, i.e. the Bose-Einstein function has a singularity. The particle number is no good description for phonons, at least for zero momentum acoustical phonons. Yet in case of the internal energy, these problems do not appear. Using l'hospital's theorem, one can see that

$$\frac{x}{e^x - 1} \rightarrow 1 \quad \text{for } x \searrow 0 .$$

High temperature limit

In the high temperature limit, it holds that $\hbar\omega_j(\mathbf{k}) \ll k_B T$, for all (j, \mathbf{k}) . This allows to consider only the first order contribution around 0 in $\frac{\hbar\omega_j(\mathbf{k})}{k_B T}$ of the exponential function $\exp\left(\frac{\hbar\omega_j(\mathbf{k})}{k_B T}\right)$:

$$\exp\left(\frac{\hbar\omega_j(\mathbf{k})}{k_B T}\right) \approx 1 + \frac{\hbar\omega_j(\mathbf{k})}{k_B T} .$$

For the internal energy it follows that:

$$\begin{aligned} U &\approx E_0 + \sum_{\mathbf{k}} \sum_{j=1}^{n-d} \frac{\hbar\omega_j(\mathbf{k})}{1 + \frac{\hbar\omega_j(\mathbf{k})}{k_B T} - 1} = E_0 + k_B T \sum_{\mathbf{k}} \sum_{j=1}^{n-d} 1 \\ &= E_0 + k_B T \cdot \prod_{j=1}^n N_j \cdot n \cdot d . \end{aligned}$$

The product $N \equiv \prod_{j=1}^n N_j$ is the number of unit cells. In the limiting case $T \rightarrow \infty$, the internal energy approaches asymptotically:

$$U \simeq N n d k_B T .$$

This equation is called **Dulong-Petit law**.

Low temperature limit

Because of the acoustic branches, the assumption $k_B T \ll \hbar\omega_j(\mathbf{k})$ is wrong, as $\omega(\mathbf{k}) \rightarrow 0$ for $\mathbf{k} \rightarrow 0$. For that reason, we have to treat the optical and acoustic phonons separately.

Assume, that $k_B T \ll \hbar\omega_{j_0}(\mathbf{k})$, where j_0 labels the optical phonons. Then

$$\frac{\hbar\omega_j(\mathbf{k})}{k_B T} \gg 1 \quad \Rightarrow \quad \exp\left(\frac{\hbar\omega_j(\mathbf{k})}{k_B T}\right) - 1 \approx \exp\left(\frac{\hbar\omega_j(\mathbf{k})}{k_B T}\right) .$$

This shows, that the contribution of the optical phonons can be neglected, because:

$$\sum_{\mathbf{k}, j_0} \frac{\hbar\omega_{j_0}(\mathbf{k})}{\exp\left(\frac{\hbar\omega_{j_0}(\mathbf{k})}{k_B T}\right) - 1} \approx \sum_{\mathbf{k}, j_0} \hbar\omega_{j_0}(\mathbf{k}) \exp\left(-\frac{\hbar\omega_{j_0}(\mathbf{k})}{k_B T}\right) ,$$

which tends to zero for $T \rightarrow 0$.

For the analysis of the acoustic phonons, we assume, that the temperature is low enough, such that only low momentum phonons are thermally excited. This allows to use the dispersion relation $\omega_{j_a}(\mathbf{k}) = c_{j_a} \|\mathbf{k}\|$:

$$U - E_0 = \sum_{\mathbf{k}, j_a} \frac{\hbar c_{j_a} \|\mathbf{k}\|}{\exp\left(\frac{\hbar c_{j_a} \|\mathbf{k}\|}{k_B T}\right) - 1}.$$

To evaluate the sum, it is approximated by integration, a standard approach in statistical physics.

Remark 3.3.4 (Integral approximation).

For the integral approximation we recall the definition of the Riemann integral:

$$\int_a^b f(x) dx \underset{\Delta x \rightarrow 0}{\approx} \Delta x \sum_{k=0}^{\frac{b-a}{\Delta x}} f(x_k).$$

In case of a one dimensional lattice, this leads to the approximation:

$$\sum_{k \in \text{pc}} f(k) \approx \frac{1}{\Delta k} \int_{\text{pc}} f(k) dk,$$

where pc is the primitive cell to which the momenta k are confined. From (3.6), we find, that $\Delta k = \frac{2\pi}{Na}$, where $\frac{2\pi}{a} = \|g\|$. With the lattice length $L = Na$, this leads to the one dimensional approximation

$$\sum_{k \in \text{pc}} f(k) \approx \frac{L}{2\pi} \int_{\text{pc}} f(k) dk.$$

In n dimensions, with $V = NV_{\text{pc}}$ it follows that:

$$\sum_{\mathbf{k} \in \text{pc}} f(\mathbf{k}) \approx \frac{V}{(2\pi)^n} \int_{\text{pc}} f(\mathbf{k}) dk^n.$$

With the integral approximation we obtain:

$$\begin{aligned} U - E_0 &\approx \frac{V}{(2\pi)^n} \sum_{j_a} \int \frac{\hbar c_{j_a} \|\mathbf{k}\|}{\exp\left(\frac{\hbar c_{j_a} \|\mathbf{k}\|}{k_B T}\right) - 1} dk^n \\ &= \frac{V}{(2\pi)^n} \sum_{j_a} \left(\frac{k_B T}{\hbar c_{j_a}}\right)^n k_B T \int \frac{\|\mathbf{x}\|}{e^{\|\mathbf{x}\|} - 1} dx^n \\ &\approx \frac{V}{(2\pi)^n} \sum_{j_a} \left(\frac{k_B T}{\hbar c_{j_a}}\right)^3 k_B T \int_{\mathbb{R}^n} \frac{\|\mathbf{x}\|}{e^{\|\mathbf{x}\|} - 1} dx^n. \end{aligned}$$

In the second line, we made the substitution $x_i = \frac{\hbar c_{j_a} k_i}{k_B T}$. The additional $k_B T$ arises, because the numerator $\hbar c_{j_a} \|\mathbf{k}\|$ lacks a $\frac{1}{k_B T}$. For the integration boundaries, this leads to $[-\frac{\pi}{a_i}, \frac{\pi}{a_i}] \rightarrow [-\frac{\pi \hbar c_{j_a}}{a_i k_B T}, \frac{\pi \hbar c_{j_a}}{a_i k_B T}]$ for the unit cell. The approximation in the third line uses, that we are interested in the limiting case $T \rightarrow 0$, such that the boundaries can

be set to $\pm\infty$ as approximation. In three dimensions, the integral can be evaluated in spherical coordinates $dx^3 = r^2 \sin(\theta) dr d\theta d\phi$:

$$\int_{\mathbb{R}^3} \frac{\|\mathbf{x}\|}{e^{\|\mathbf{x}\|} - 1} dx^3 = 4\pi \int_0^\infty \frac{r^3}{e^r - 1} dr = 4\pi \frac{\pi^4}{15} .$$

There are 3 acoustic branches. Replacing the c_{j_a} by the averaged speed of sound c_s , the sum becomes $\sum_{j_a} \frac{1}{c_s^3} = 3 \frac{1}{c_s^3}$. Plugging all in, we obtain:

$$U - E_0 = \frac{\pi^2}{10} \frac{V}{(\hbar c_s)^3} (k_B T)^4 .$$

Specific heat

The **specific heat** per unit volume is defined by

$$C_V = \frac{1}{V} \frac{\partial U}{\partial T} .$$

For the limiting cases, this means:

$$\begin{array}{ll} \text{low temperature:} & \frac{2}{5} \pi^2 k_B \left(\frac{k_B T}{\hbar c_s} \right)^3 \\ \text{high temperatures} & 3N d k_B \end{array} .$$

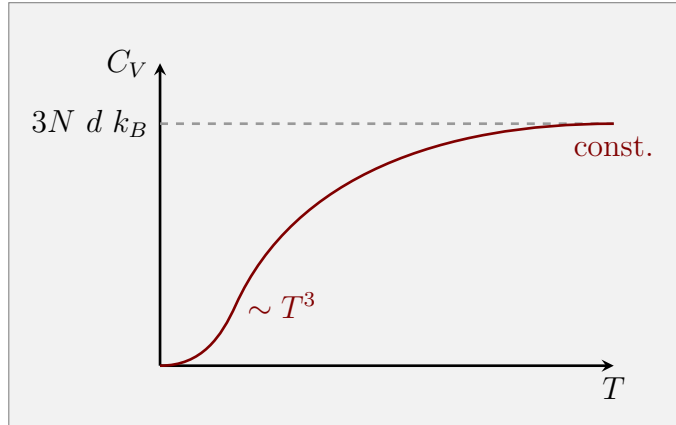


Figure 3.4: Specific heat

Debye model

The Debye model tries to model the internal energy between the limiting cases. It uses the following two assumptions:

1. Instead of the Brillouin zone, choose a sphere (called **Debye sphere**) around $\mathbf{k} = 0$. The radius k_D of the sphere is chosen, such that there are still $N = \prod_{j=1}^n N_j$ discrete momenta contained:

$$N = \sum_{\mathbf{k}, \|\mathbf{k}\| < k_D} 1 \approx \frac{V}{(2\pi)^3} \int_{\mathbf{k}, \|\mathbf{k}\| < k_D} dk^3 = \frac{V}{6\pi^2} k_D^3 ,$$

$$\Rightarrow k_D = \sqrt[3]{6\pi^2 \frac{N}{V}} .$$

2. The dispersion relation is assumed to be $\omega(\mathbf{k}) = c_s \|\mathbf{k}\|$ for all phonons on the whole Debye sphere. Here c_s is again the averaged speed of sound.

The result of the Debye model for the internal energy, repeating the steps of the low temperature limit, is:

$$U - E_0 = 3d \frac{V}{2\pi^2} \left(\frac{k_B T}{\hbar c_s} \right)^3 k_B T \int_0^{\frac{\Theta_D}{T}} \frac{r^3}{e^r - 1} dr ,$$

where $\Theta_D = \frac{\hbar \omega_D}{k_B}$ is the **Debye temperature** and $\omega_D = c_s k_D$ the **Debye frequency**. Furthermore, the first assumption also yields $V = \frac{6\pi^2 N}{k_D^3}$, such that (using that $d \cdot N = N_a$ is the number of atoms):

$$U - E_0 = 9N_a \left(\frac{T}{\Theta_D} \right)^3 k_B T \int_0^{\frac{\Theta_D}{T}} \frac{r^3}{e^r - 1} dr .$$

Phonon state density

Definition 3.3.5.

The **phonon state density** $\eta(\omega) d\omega$ is defined by

$$\eta(\omega) := \frac{1}{N} \sum_{j=1}^{n \cdot d} \sum_{\mathbf{k}} \delta(\omega - \omega_j(\mathbf{k}))$$

The integration $\int_{\omega_1}^{\omega_2} \eta(\omega) d\omega$ of the phonon density yields the number of possible phonon states per unit cell, with $\omega_j(\mathbf{k}) \in [\omega_1, \omega_2]$. Hence, the phonon state density has the following normalization:

$$\int_0^{\infty} \eta(\omega) d\omega = n \cdot d .$$

Let $e(x) = \frac{x}{\exp(\frac{x}{k_B T}) - 1}$, to keep the notation short.

Lemma 3.3.6.

The *internal energy* can be calculated as follows:

$$U - E_0 = N \int_{\mathbb{R}} \eta(\omega) e(\hbar\omega) d\omega .$$

Proof 3.3.7.

$$\begin{aligned} N \int_{\mathbb{R}} \eta(\omega) e(\hbar\omega) d\omega &= \sum_{j=1}^{n \cdot d} \sum_{\mathbf{k}} \int_{\mathbb{R}} \delta(\omega - \omega_j(\mathbf{k})) e(\hbar\omega) d\omega \\ &= \sum_{j=1}^{n \cdot d} \sum_{\mathbf{k}} e(\hbar\omega_j(\mathbf{k})) = \sum_{j=1}^{n \cdot d} \sum_{\mathbf{k}} \frac{\hbar\omega_j(\mathbf{k})}{\exp\left(\frac{\hbar\omega_j(\mathbf{k})}{k_B T}\right) - 1} \end{aligned}$$

$$= U - E_0 .$$

□

In the integration approximation, the phonon state density function reads:

$$\eta(\omega) \approx \frac{1}{N} \frac{V}{(2\pi)^n} \sum_j \int_{\text{pc}} \delta(\omega - \omega_j(\mathbf{k})) dk^n .$$

For the Debye model, in 3 dimensions, we have $\omega_j(\mathbf{k}) = c_s \|\mathbf{k}\|$, $\sum_j 1 = 3d$ and $V = \frac{6\pi^2 N}{k_D^3}$:

$$\begin{aligned} n(\omega) &\approx \frac{3d}{4\pi k_D^3} \int_{B_{k_D}(0)} \delta(\omega - c_s \|k\|) dk^3 \\ &= \frac{3d}{k_D^3} \int_0^{k_D} r^2 \delta(\omega - c_s r) dr \\ &= \frac{3d}{k_D^3} \int_0^{k_D c_s} \left(\frac{x}{c_s}\right)^2 \delta(\omega - x) \frac{dx}{c_s} \\ &= \frac{3d}{k_D^3} \frac{\omega^2}{c_s^3} = 3d \cdot \omega^2 \frac{\hbar^3}{(k_B \Theta_D)^3} . \end{aligned}$$

Non-interacting electrons in solids

For the Born-Oppenheimer approximation we assumed, that we can solve the Schrödinger equation for the electrons, if the ion positions are fixed. However, the problem that made the Born-Oppenheimer approximation necessary, also applies to the electron Schrödinger equation. The huge number of electrons makes a direct solution impossible. To approximate the problem, we assume non-interacting/weakly interacting electrons first.

4.1 Electrons in a periodic potential

Recall the Hamiltonian H_0 from section 2.2:

$$H_0 = T_e + V_{ee} + V_{aa} + V_{ea} .$$

The ion-ion contribution is just an additive constant for fixed ion positions, such that we can include it in V_{ea} or set to zero. Non-interacting electrons means, that $V_{ee} \equiv 0$. This allows to treat the electrons individually, reducing the problem to a single particle problem. Thus, we consider the following stationary Schrödinger equation:

$$H\psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad \text{with} \quad H = T + V .$$

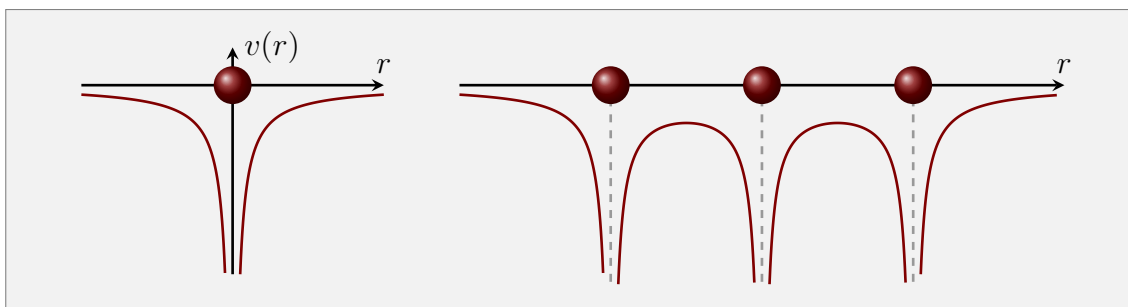


Figure 4.1: Coulomb potential in one dimension.

For the ion positions $\mathbf{R}_{\vec{j},\alpha}$ we assume that $\mathbf{R}_{\vec{j},\alpha} = \mathbf{R}_{\vec{j}} + \mathbf{b}_\alpha$, where $\mathbf{R}_{\vec{j}}$ is a Bravais lattice point and \mathbf{b}_α the atom position. Let $v_\alpha(\mathbf{r})$ denote the potential of the α -th ion at the origin. Then, $v_\alpha(\mathbf{r} - \mathbf{R}_{\vec{j}} - \mathbf{b}_\alpha)$ is the potential for the α -th ion at position $\mathbf{R}_{\vec{j},\alpha}$. Thus:

$$V(\mathbf{r}) = \sum_{\vec{j}} \sum_{\alpha} v_\alpha(\mathbf{r} - \mathbf{R}_{\vec{j}} - \mathbf{b}_\alpha) =: \sum_{\vec{j}} \tilde{v}(\mathbf{r} - \mathbf{R}_{\vec{j}}) .$$

Because of the periodicity of the lattice (periodic boundary conditions), the potential V is periodic w.r.t. Bravais vectors $\mathbf{R}_{\vec{\ell}}$:

$$V(\mathbf{r} + \mathbf{R}_{\vec{\ell}}) = \sum_{\vec{j}} \tilde{v}(\mathbf{r} - \mathbf{R}_{\vec{j}} + \mathbf{R}_{\vec{\ell}}) = \sum_{\vec{m}} \tilde{v}(\mathbf{r} - \mathbf{R}_{\vec{m}}) = V(\mathbf{r}) .$$

4.1.1 Application of Bloch's theorem

The potential is invariant under Bravais vector translations. Hence the Hamiltonian and translation operators commute $[T_{\vec{j}}, H] = 0$. This allows to use Bloch's theorem (3.5) as ansatz for the Schrödinger equation:

$$H\psi_{\mathbf{k}}(\mathbf{r}) = E_{\mathbf{k}}\psi_{\mathbf{k}}(\mathbf{r}) \quad \text{with} \quad \psi_{\mathbf{k}}(\mathbf{r}) = e^{i\langle \mathbf{k}, \mathbf{r} \rangle} u_{\mathbf{k}}(\mathbf{r}) .$$

Recall, that the Bloch functions $u_{\mathbf{k}}$ are periodic, i.e. $T_{\vec{j}}u = u$. Though the wave function $\psi_{\mathbf{k}}$ is not periodic (unless $\mathbf{k} = 0$), the electron density $|\psi_{\mathbf{k}}|^2$ is. Again, the momenta are given by (3.6).

A direct calculation shows that:

$$\Delta e^{i\langle \mathbf{k}, \mathbf{r} \rangle} u_{\mathbf{k}}(\mathbf{r}) = e^{i\langle \mathbf{k}, \mathbf{r} \rangle} \left(-\|\mathbf{k}\|^2 + 2i \sum_{m=1}^3 k_m \partial_m + \Delta \right) u_{\mathbf{k}}(\mathbf{r}) .$$

This leads to the equation

$$h(\mathbf{k})u_{\mathbf{k}}(\mathbf{r}) = E_{\mathbf{k}}u_{\mathbf{k}}(\mathbf{r}) ,$$

with the effective Hamiltonian

$$h(\mathbf{k}) = \frac{\hbar^2}{2m} \left(\|\mathbf{k}\|^2 - 2i \sum_{m=1}^3 k_m \partial_m - \Delta \right) + V(\mathbf{r}) .$$

This equation has to be solved for every \mathbf{k} individually. We are only interested in bound states, for which the energy is quantized. As usual, we use the quantum number n , to label the energy levels:

$$h(\mathbf{k})u_{n,\mathbf{k}}(\mathbf{r}) = E_{n,\mathbf{k}}u_{n,\mathbf{k}}(\mathbf{r}) . \quad (4.1)$$

Because of the periodicity of V and $u_{\mathbf{k}}$, it is enough to solve this equation for \mathbf{r} confined to a unit cell.

Remark 4.1.1.

Because of the periodicity of $u_{n,\mathbf{k}}$, the wave functions $\psi_{n,\mathbf{k}}$ are no $L^2(\mathbf{R}^3)$ functions, i.e. have no normalization. However, since the solid is finite anyway, this is no serious problem. Still, the periodic boundary conditions create an infinitely large lattice. But, because of the restriction to the unit cell, there is still a normalization possible:

$$\int_{V_{\text{uc}}} \overline{u_{n',\mathbf{k}'}}(\mathbf{r}) u_{n,\mathbf{k}}(\mathbf{r}) d\mathbf{r}^3 = \delta_{n'n} \delta_{\mathbf{k}'\mathbf{k}} .$$

Let N be the number of unit cells, then the wave functions are normalized by

$$\int_V \overline{\psi_{n\mathbf{k}}}(\mathbf{r}) \psi_{n\mathbf{k}}(\mathbf{r}) d\mathbf{r}^3 = N \delta_{n'n} \delta_{\mathbf{k}'\mathbf{k}} .$$

Definition 4.1.2.

The energy dependence $E_{n,\mathbf{k}}$ of the momentum \mathbf{k} is called **electronic band**. For that reason, the index n is called **band index**, as it labels the electronic bands.

The effective Hamiltonian $h(\mathbf{k})$ has no upper bound, such that unlike phonos, there are

infinitely many electronic bands. The set of electronic bands is called **band structure** of the solid. It can happen, that $E_{n,\mathbf{k}} = E_{n',\mathbf{k}}$ for $n \neq n'$. This is called **band degeneracy**.

4.1.2 Fourier transformation of the Schrödinger equation

The goal of this subsection is to diagonalize $H\psi(\mathbf{r}) = E\psi(\mathbf{r})$. Having section 1.2 in mind, we find:

$$V(\mathbf{r}) = \sum_{\mathbf{G} \in \tilde{L}} V_{\mathbf{G}} e^{i\langle \mathbf{G}, \mathbf{r} \rangle} \quad \text{and} \quad \psi(\mathbf{r}) = \sum_{\mathbf{q} \in S} \psi_{\mathbf{q}} e^{i\langle \mathbf{q}, \mathbf{r} \rangle} .$$

The wave functions are not lattice periodic, but only satisfy the periodic boundary conditions from subsection 3.2.3. For the set S , the \mathbf{q} are summed over, it follows that

$$S = \left\{ \mathbf{q} \left| \mathbf{q} = \sum_{j=1}^n \frac{m_j}{N_j} \mathbf{g}_j, \quad m_j \in \mathbb{Z} \right. \right\} ,$$

where the \mathbf{g}_j are the reciprocal primitive vectors. It follows that:

$$\begin{aligned} H\psi(\mathbf{r}) &= \dots = \sum_{\mathbf{q}} \left(\frac{\hbar^2 \|\mathbf{q}\|^2}{2m} + \sum_{\mathbf{G} \in \tilde{L}} V_{\mathbf{G}} e^{i\langle \mathbf{G}, \mathbf{r} \rangle} \right) \psi_{\mathbf{q}} e^{i\langle \mathbf{q}, \mathbf{r} \rangle} \\ &= E_n \sum_{\mathbf{q}} \psi_{\mathbf{q}} e^{i\langle \mathbf{q}, \mathbf{r} \rangle} . \end{aligned}$$

With the substitution $\mathbf{q}' = \mathbf{q} - \mathbf{G}$:

$$\sum_{\mathbf{q}, \mathbf{G}} V_{\mathbf{G}} \psi_{\mathbf{q}} e^{i\langle \mathbf{G} + \mathbf{q}, \mathbf{r} \rangle} = \sum_{\mathbf{q}', \mathbf{G}} V_{\mathbf{G}} \psi_{\mathbf{q}' - \mathbf{G}} e^{i\langle \mathbf{q}', \mathbf{r} \rangle} = \sum_{\mathbf{q}, \mathbf{G}} V_{\mathbf{G}} \psi_{\mathbf{q} - \mathbf{G}} e^{i\langle \mathbf{q}, \mathbf{r} \rangle} ,$$

where we relabeled \mathbf{q}' in the last step. Hence:

$$\sum_{\mathbf{q}} \left(\left(\frac{\hbar^2 \|\mathbf{q}\|^2}{2m} - E_n \right) \psi_{\mathbf{q} - \mathbf{G}} + \sum_{\mathbf{G} \in \tilde{L}} V_{\mathbf{G}} \psi_{\mathbf{q} - \mathbf{G}} \right) e^{i\langle \mathbf{q}, \mathbf{r} \rangle} = 0 .$$

The $e^{i\langle \mathbf{q}, \mathbf{r} \rangle}$ are linearly independent and non-zero, such that the brackets must be zero for each \mathbf{q} individually. Furthermore, it is possible, that the energy E depends on the momentum \mathbf{q} . Thus, we obtain the following equation

$$\left(\frac{\hbar^2 \|\mathbf{q}\|^2}{2m} - E_{\mathbf{q}} \right) \psi_{\mathbf{q} - \mathbf{G}} + \sum_{\mathbf{G} \in \tilde{L}} V_{\mathbf{G}} \psi_{\mathbf{q} - \mathbf{G}} = 0 .$$

Any \mathbf{q} can be written as

$$\mathbf{q} = \mathbf{k} + \mathbf{K} , \quad \text{with} \quad \mathbf{K} \in \tilde{L} \quad \text{and} \quad \mathbf{k} = \sum_{j=1}^n \frac{m_j}{N_j} \mathbf{g}_j ,$$

$$\text{where} \quad m_j = -\frac{N_j}{2}, -\frac{N_j}{2} + 1, \dots, \frac{N_j}{2} - 1 .$$

Hence, $\mathbf{k} = \mathbf{q} - \mathbf{K}$ is again restricted to the Brillouin zone. Substituting G by $\mathbf{G} + \mathbf{K}$ in the sum leads to

$$\left(\frac{\hbar^2 \|\mathbf{k} - \mathbf{K}\|^2}{2m} - E_{\mathbf{K}}(\mathbf{k}) \right) \psi_{\mathbf{k}-\mathbf{K}} + \sum_{\mathbf{G} \in \tilde{\mathcal{L}}} V_{\mathbf{G}-\mathbf{K}} \psi_{\mathbf{k}-\mathbf{G}} = 0. \quad (4.2)$$

For the energies, we have used the notation $E_{\mathbf{K}}(\mathbf{k}) := E_{\mathbf{k}-\mathbf{K}}$. Here, \mathbf{K} is the band index.

Remark 4.1.3.

We recall that (here we choose the Brillouin zone)

$$V_{\mathbf{G}} = \frac{1}{V_{\text{Bz}}} \int_{\text{Bz}} V(\mathbf{r}) e^{-i\langle \mathbf{G}, \mathbf{r} \rangle} d\mathbf{r}^3.$$

For real potentials, this leads to

$$\begin{aligned} \overline{V}_{\mathbf{G}} &= \frac{1}{V_{\text{Bz}}} \int_{\text{Bz}} \overline{V(\mathbf{r}) e^{-i\langle \mathbf{G}, \mathbf{r} \rangle}} d\mathbf{r}^3 = \frac{1}{V_{\text{Bz}}} \int_{\text{Bz}} V(\mathbf{r}) e^{i\langle \mathbf{G}, \mathbf{r} \rangle} d\mathbf{r}^3 \\ &= V_{-\mathbf{G}}. \end{aligned}$$

4.2 Nearly free electrons

In this section, we focus on the valance electrons. The bound electrons shield the valance electrons partially from the positively charged nuclei, i.e. decrease the potential V . To solve (4.2), the potential is treated as perturbation, with zeroth order $V = 0$:

$$\left(\frac{\hbar^2 \|\mathbf{k} - \mathbf{K}\|^2}{2m} - E_{\mathbf{K}}(\mathbf{k}) \right) \psi_{\mathbf{k}-\mathbf{G}} \Rightarrow E_{\mathbf{K}}(\mathbf{k}) = \frac{\hbar^2 \|\mathbf{k} - \mathbf{K}\|^2}{2m}$$

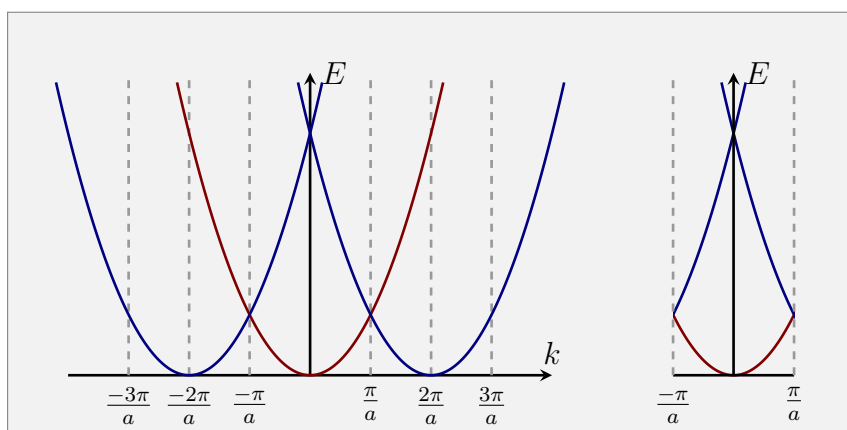


Figure 4.2: Electronic bands in one dimension and their restriction to the Brillouin zone.

In one dimension, the electronic bands are thus:

$$E_n(k) = \frac{\hbar^2}{2m} \left(k - \frac{2\pi}{a} n \right)^2, \quad n \in \mathbb{Z}.$$

For the band labeling we identify n and $-n$.

With $E_{\mathbf{K}}(\mathbf{k}) = \frac{\hbar^2 \|\mathbf{k} - \mathbf{K}\|^2}{2m}$, the equation Fourier transformed Schrödinger equation can be written

$$(E_{\mathbf{K}}(\mathbf{k}) - \mathcal{E})\psi_{\mathbf{k}-\mathbf{K}} = -V_0\psi_{\mathbf{k}-\mathbf{K}} - \sum_{\tilde{\mathbf{L}} \ni \mathbf{G} \neq \mathbf{K}} V_{\mathbf{G}-\mathbf{K}}\psi_{\mathbf{k}-\mathbf{G}},$$

where $V_0 = V_{\mathbf{K}-\mathbf{K}}$ and \mathcal{E} is the eigen energy with potential. Fixing a new gauge we can set $V_0 = 0$, i.e. setting $V' = V - V_0$:

$$\Rightarrow (E_{\mathbf{K}}(\mathbf{k}) - \mathcal{E})\psi_{\mathbf{k}-\mathbf{K}} = - \sum_{\tilde{\mathbf{L}} \ni \mathbf{G} \neq \mathbf{K}} V_{\mathbf{G}-\mathbf{K}}\psi_{\mathbf{k}-\mathbf{G}}. \quad (4.3)$$

The perturbative treatment of (4.3) looks as follows:

- Solve (4.3) for $\psi_{\mathbf{K}}(\mathbf{k})$:

$$\psi_{\mathbf{k}-\mathbf{K}} = \frac{1}{\mathcal{E} - E_{\mathbf{K}}(\mathbf{k})} \sum_{\mathbf{G} \neq \mathbf{K}} V_{\mathbf{G}-\mathbf{K}}\psi_{\mathbf{k}-\mathbf{G}}.$$

and adjust the indices for the next step:

$$\Rightarrow \psi_{\mathbf{k}-\mathbf{G}} = \frac{1}{\mathcal{E} - E_{\mathbf{G}}(\mathbf{k})} \sum_{\tilde{\mathbf{G}} \neq \mathbf{G}} V_{\tilde{\mathbf{G}}-\mathbf{G}}\psi_{\mathbf{k}-\tilde{\mathbf{G}}}. \quad (4.4)$$

- In the second step, we rewrite (4.3) into

$$\mathcal{E} = E_{\mathbf{K}}(\mathbf{K}) + \frac{1}{\psi_{\mathbf{k}-\mathbf{K}}} \sum_{\mathbf{G} \neq \mathbf{K}} V_{\mathbf{G}-\mathbf{K}}\psi_{\mathbf{k}-\mathbf{G}}$$

and plug (4.4) in:

$$\mathcal{E} = E_{\mathbf{K}}(\mathbf{K}) + \frac{1}{\psi_{\mathbf{k}-\mathbf{K}}} \sum_{\mathbf{G} \neq \mathbf{K}} \frac{V_{\mathbf{G}-\mathbf{K}}}{\mathcal{E} - E_{\mathbf{G}}(\mathbf{k})} \sum_{\tilde{\mathbf{G}} \neq \mathbf{G}} V_{\tilde{\mathbf{G}}-\mathbf{G}}\psi_{\mathbf{k}-\tilde{\mathbf{G}}} \quad (4.5)$$

- Consider the last sum of (4.5), and split it

$$\sum_{\tilde{\mathbf{G}} \neq \mathbf{G}} V_{\tilde{\mathbf{G}}-\mathbf{G}}\psi_{\mathbf{k}-\tilde{\mathbf{G}}} = V_{\mathbf{K}-\mathbf{G}} + \psi_{\mathbf{k}-\mathbf{K}} \sum_{\tilde{\mathbf{G}} \neq \mathbf{G}, \mathbf{K}} V_{\tilde{\mathbf{G}}-\mathbf{G}}\psi_{\mathbf{k}-\tilde{\mathbf{G}}}$$

Hence (4.5) reads:

$$\begin{aligned} \mathcal{E} &= E_{\mathbf{K}}(\mathbf{K}) + \sum_{\mathbf{G} \neq \mathbf{K}} \frac{V_{\mathbf{G}-\mathbf{K}}V_{\mathbf{K}-\mathbf{G}}}{\mathcal{E} - E_{\mathbf{G}}(\mathbf{k})} \\ &\quad + \frac{1}{\psi_{\mathbf{k}-\mathbf{K}}} \sum_{\mathbf{G} \neq \mathbf{K}} \sum_{\tilde{\mathbf{G}} \neq \mathbf{G}, \mathbf{K}} \frac{V_{\mathbf{G}-\mathbf{K}}V_{\tilde{\mathbf{G}}-\mathbf{G}}}{\mathcal{E} - E_{\mathbf{G}}(\mathbf{k})} \psi_{\mathbf{k}-\tilde{\mathbf{G}}}. \end{aligned} \quad (4.6)$$

- So far, all steps have been exact. Repeating the same procedure, i.e. plugging in (4.4), show that the the double sum term in (4.6) is of order $\mathcal{O}(V^3)$. The first approximation is (using remark 4.1.3):

$$\mathcal{E} = E_{\mathbf{K}}(\mathbf{K}) + \sum_{\mathbf{G} \neq \mathbf{K}} \frac{|V_{\mathbf{G}-\mathbf{K}}|^2}{\mathcal{E} - E_{\mathbf{G}}(\mathbf{k})} + \mathcal{O}(V^3).$$

- For the next approximation, we recall that band degeneracies can occur. We intend to set \mathcal{E} to $E_{\mathbf{K}}(\mathbf{K})$ on the right hand side since we assume V^2 to be small enough. Let \mathbf{K}_d be the point of band degeneracy for \mathbf{k} . Splitting the sum we can thus approximate:

$$\mathcal{E} = E_{\mathbf{K}}(\mathbf{K}) + \frac{|V_{\mathbf{K}_d-\mathbf{K}}|^2}{\mathcal{E} - E_{\mathbf{K}_d}(\mathbf{k})} + \sum_{\mathbf{G} \neq \mathbf{K}, \mathbf{K}_d} \frac{|V_{\mathbf{G}-\mathbf{K}}|^2}{E_{\mathbf{K}}(\mathbf{k}) - E_{\mathbf{G}}(\mathbf{k})} + \mathcal{O}(V^3).$$

- In the preceding approximation we used that $\mathcal{E} = E_{\mathbf{K}_d}(\mathbf{k}) \approx E_{\mathbf{K}}(\mathbf{K})$. Since we have filtered out \mathbf{K} and \mathbf{K}_d in the sum, the term $E_{\mathbf{K}}(\mathbf{k}) - E_{\mathbf{G}}(\mathbf{k})$ is large. Hence, the sum term can also be neglected, leading to the approximation for \mathcal{E} :

$$\boxed{\mathcal{E} = E_{\mathbf{K}}(\mathbf{K}) + \frac{|V_{\mathbf{K}_d-\mathbf{K}}|^2}{\mathcal{E} - E_{\mathbf{K}_d}(\mathbf{k})}} \quad (4.7)$$

Recalling, that we have started with an equation (4.3) for the \mathbf{K} -th electronic band, we write $\mathcal{E} = \mathcal{E}_{\mathbf{K}}(\mathbf{k})$. Solving (4.7) for $\mathcal{E} = \mathcal{E}_{\mathbf{K}}(\mathbf{k})$ leads to

$$\mathcal{E}_{\mathbf{K}}^{\pm}(\mathbf{k}) = \frac{1}{2} \left(E_{\mathbf{K}}(\mathbf{k}) + E_{\mathbf{K}_d}(\mathbf{k}) \pm \sqrt{(E_{\mathbf{K}}(\mathbf{k}) - E_{\mathbf{K}_d}(\mathbf{k}))^2 + 4|V_{\mathbf{K}_d-\mathbf{K}}|^2} \right).$$

There is an important case to discuss here. In the point \mathbf{k} of band degeneracy it holds

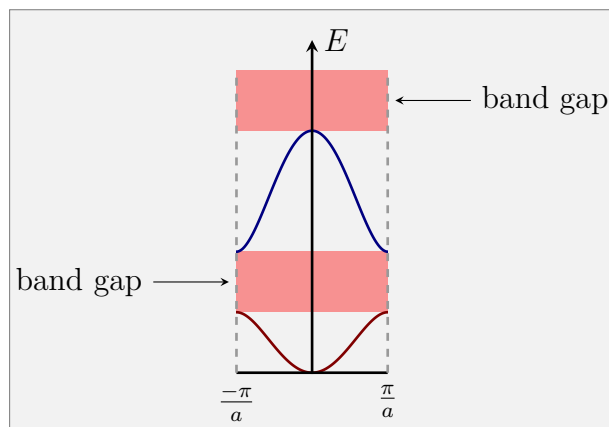


Figure 4.3: One dimensional case (see fig. 4.2) with band gaps.

that $E_{\mathbf{K}_d}(\mathbf{k}) = E_{\mathbf{K}}(\mathbf{k})$. Then we find

$$\boxed{\mathcal{E}_{\mathbf{K}}^{\pm}(\mathbf{k}) = E_{\mathbf{K}} \pm 2|V_{\mathbf{K}_d-\mathbf{K}}|}.$$

This means, that the band degeneracy vanishes and the bands \mathbf{K} and \mathbf{K}_d become separated by an energy gap of $2|V_{\mathbf{K}_d-\mathbf{K}}|$. This separation of bands is called **band gap**. Also note, that the bands are labeled by $\{\mathbf{K}, \pm\}$ now.

4.3 The effective mass

We return to (4.1), where we can write the effective Hamiltonian as

$$\begin{aligned} h(\mathbf{k}) &= \frac{\hbar^2}{2m} \|\mathbf{k}\|^2 - \frac{i\hbar^2}{m} \mathbf{k} \cdot \vec{\nabla} - \frac{\hbar^2}{2m} \Delta + V(\mathbf{r}) \\ &= \left(-\frac{\hbar^2}{2m} \Delta + V(\mathbf{r}) \right) + \left(\frac{\hbar^2}{2m} \|\mathbf{k}\|^2 + \frac{\hbar}{m} \mathbf{k} \cdot \mathbf{p} \right) \\ &\equiv h(0) + h_1(\mathbf{k}) . \end{aligned}$$

Assuming to have found a solution for the case $\mathbf{k} = 0$, i.e.

$$h(0)u_{n,0}(\mathbf{r}) = E_{n,0}u_{n,0}(\mathbf{r}) ,$$

we can consider the full equation for small \mathbf{k} as perturbation problem in $\mathbf{k} \cdot \mathbf{p}$ with perturbation Hamiltonian $h_1(\mathbf{k})$. In most cases, it holds that $\{E_{n,0}\}$ is non-degenerate, which we assume here to be true in general.

4.3.1 The Rayleigh-Schrödinger approximation

The perturbation problem can be solved with the Rayleigh-Schrödinger approximation. Here, we sketch the treatment of [Sch13, section 11.1]

Consider the Hamiltonian $H = H_0 + \lambda h$, where h is the perturbation and λ the development coefficient. Assume that for H_0 , the eigen states $|n_0\rangle$ and eigen energies $E_0(n)$ are known:

$$H_0|n_0\rangle = E_0(n)|n_0\rangle ,$$

and that the $\ker n_0$ are a Hilbert basis. Also, we use that the eigen energies are non-degenerate.

The goal is to find the solution for

$$H|n\rangle = E(n)|n\rangle .$$

We expand these solutions in orders of the development coefficient:

$$E(n) = E_0(n) + \lambda E_1(n) + \lambda^2 E_2(n) + \dots \quad \text{and} \quad |n\rangle = |n_0\rangle + \lambda |n_1\rangle + \lambda^2 |n_2\rangle + \dots .$$

With the normalization $\langle n_0, n_j \rangle = \delta_{0j}$. Plugging in, leads to:

$$\begin{aligned} H|n\rangle &= (H_0 + \lambda h) \sum_j \lambda^j |n_j\rangle = \sum_j \lambda^j E_j(n) \sum_\ell \lambda^\ell |n_\ell\rangle . \\ \Rightarrow \sum_j \lambda^j H_0 |n_j\rangle + \lambda^{j+1} h |n_j\rangle &= \sum_{j+\ell} \lambda^{j+\ell} E_j(n) |n_\ell\rangle . \end{aligned}$$

Comparing the coefficients shows that:

$$\begin{aligned} \text{zeroth order:} \quad & H_0 |n_0\rangle = E_0(n) |n_0\rangle \\ \text{first order:} \quad & H_0 |n_1\rangle + h |n_0\rangle = E_0(n) |n_1\rangle + E_1(n) |n_0\rangle \\ \text{second order:} \quad & H_0 |n_2\rangle + h |n_1\rangle = E_0(n) |n_2\rangle + E_1(n) |n_1\rangle + E_2(n) |n_0\rangle \end{aligned}$$

Multiplying $\langle n_0|$ from the left with the first order, we obtain

$$E_1(n) = \langle n_0|h|n_0\rangle .$$

For the second order contribution $E_2(n)$, we use, that $\{|n_0\rangle\}$ is a Hilbert basis, such that

$$|n_1\rangle = \sum_{m_0 \neq n_0} c(m_0)|m_0\rangle .$$

The case $m_0 = n_0$ could be excluded, because $\langle n_0|n_1\rangle = 0$. Plugging this development in into the first order, and multiplying from the left with $\langle \tilde{m}_0| \neq \langle n_0|$ it follows that:

$$c(\tilde{m}) (E_0(n) - E_0(\tilde{m})) = \langle \tilde{m}_0|h|n_0\rangle .$$

Because of the assumed non-degeneracy, we can divide by $(E_0(n) - E_0(\tilde{m}))$ to find the development coefficient:

$$\begin{aligned} c(\tilde{m}) &= \frac{\langle \tilde{m}_0|h|n_0\rangle}{(E_0(n) - E_0(\tilde{m}))} \\ \Rightarrow |n_1\rangle &= \sum_{m_0 \neq n_0} \frac{\langle m_0|h|n_0\rangle}{(E_0(n) - E_0(m_0))} |m_0\rangle . \end{aligned}$$

Plugging this in, into the second order, and multiplying from the left with $\langle n_0|$, we find:

$$E_2(n) = E_0(n) + \langle n_0|h|n_1\rangle = \sum_{m_0 \neq n_0} \frac{\langle m_0|h|n_0\rangle \langle n_0|h|m_0\rangle}{(E_0(n) - E_0(m_0))} = \sum_{m_0 \neq n_0} \frac{|\langle m_0|h|n_0\rangle|^2}{(E_0(n) - E_0(m_0))} .$$

The development coefficient was a tool to develop the perturbation, and is usually contained in the perturbation Hamiltonian $\mathfrak{h} = \lambda h$. With this, the energy, developed to second order reads:

$$E(n) = E_0(n) + \langle n_0|\mathfrak{h}|n_0\rangle + \sum_{m_0 \neq n_0} \frac{|\langle m_0|\mathfrak{h}|n_0\rangle|^2}{(E_0(n) - E_0(m_0))} . \quad (4.8)$$

4.3.2 Application of the Rayleigh-Schrödinger approximation

To make the application of (4.8) more easy at a notational level, we define $u_{n,\mathbf{k}}(\mathbf{r}) = \langle \mathbf{r}|n, \mathbf{k}\rangle$. Then

$$h(0)|n, 0\rangle = E_{n,0}|n, 0\rangle .$$

For the energies in second order approximation, (4.8) results in:

$$E_{n,\mathbf{k}} = E_{n,0} + \langle n, 0|h_1(\mathbf{k})|n, 0\rangle + \sum_{m \neq n} \frac{|\langle m, 0|h_1(\mathbf{k})|n, 0\rangle|^2}{(E_{n,0} - E_{m,0})} . \quad (4.9)$$

For the first order contribution, we observe, that $V(\mathbf{r}) = V(-\mathbf{r})$ because of the periodicity and since $0 = \mathbf{R}_0$ is a lattice point. Then, the operator S , defined by $Sf(\mathbf{r}) = f(-\mathbf{r})$ and $h(0)$ commute, such that the $u_{n,0}(\mathbf{r})$ can be chosen to be symmetric: $u_{n,0}(\mathbf{r}) = u_{n,0}(-\mathbf{r})$.

Corollary 4.3.1.

It holds that

$$\langle n, 0 | \mathbf{k} \cdot \mathbf{p} | n, 0 \rangle = 0 .$$

Proof 4.3.2.

In position representation, it holds that

$$\langle n, 0 | \mathbf{k} \cdot \mathbf{p} | n, 0 \rangle = \frac{-i\hbar}{2m} \int_{-\infty}^{\infty} dr_1 \dots \int_{-\infty}^{\infty} dr_3 \overline{u_{n,0}}(\mathbf{r}) \sum_j k_j \partial_j u_{n,0}(\mathbf{r}) .$$

Since $u_{n,0}(\mathbf{r})$ is symmetric, $\partial_j u_{n,0}(\mathbf{r})$ is antisymmetric in r_j . Then $\overline{u_{n,0}}(\mathbf{r}) k_j \partial_j u_{n,0}(\mathbf{r})$ is also antisymmetric, and

$$\int_{-\infty}^{\infty} dr_j \overline{u_{n,0}}(\mathbf{r}) k_j \partial_j u_{n,0}(\mathbf{r}) = 0 .$$

This term can be achieved for all j , by changing the order of the integration for each summand of $\sum_j \overline{u_{n,0}}(\mathbf{r}) k_j \partial_j u_{n,0}(\mathbf{r})$. \square

Thus, the first order contribution is

$$\langle n, 0 | h_1(\mathbf{k}) | n, 0 \rangle = \frac{\hbar^2}{2m} \|\mathbf{k}\|^2 .$$

For the second order contribution we calculate:

$$\langle m, 0 | h_1(\mathbf{k}) | n, 0 \rangle = \underbrace{\langle m, 0 | \frac{\hbar^2}{2m} \|\mathbf{k}\|^2 | n, 0 \rangle}_{=\delta_{mn}} + \langle m, 0 | \frac{\hbar}{m} \mathbf{k} \cdot \mathbf{p} | n, 0 \rangle .$$

Hence (4.9) can be written as

$$E_{n,\mathbf{k}} = E_{n,0} + \frac{\hbar^2}{2m} \|\mathbf{k}\|^2 + \frac{\hbar}{m} \sum_{\alpha,\alpha'} k_\alpha k_{\alpha'} \sum_{m \neq n} \frac{\langle n, 0 | p_\alpha | m, 0 \rangle \langle m, 0 | p_{\alpha'} | n, 0 \rangle}{(E_{n,0} - E_{m,0})} \quad (4.10)$$

This is an approximation for $E_{n,\mathbf{k}}$ valid for small \mathbf{k} . This shows, that around $\mathbf{k} = 0$ the bands (i.e. $E_{n,\mathbf{k}}$) are quadratic in the k_α .

4.3.3 The effective mass tensor

Definition 4.3.3.

The **effective mass** $\frac{1}{m^*(n)}$ is a (0, 2) tensor, defined by

$$\left(\frac{1}{m^*(n)}\right)_{\alpha\beta} = \frac{1}{\hbar} \frac{\partial^2}{\partial k_\alpha \partial k_\beta} E_{n,\mathbf{k}} .$$

In case of free electrons the energies are $E_{n,\mathbf{k}} = \frac{\hbar^2 \|\mathbf{k}\|^2}{2m}$, such that

$$\left(\frac{1}{m^*(n)}\right)_{\alpha\beta} = \frac{1}{m} \delta_{\alpha\beta} .$$

So for free electrons, the effective mass and the mass are the same. For the second order approximated energies, $E_{n,\mathbf{k}}$, the effective mass reads:

$$\left(\frac{1}{m^*(n)}\right)_{\alpha\beta} = \frac{1}{m} \delta_{\alpha\beta} + \frac{2}{m^2} \sum_{m \neq n} \frac{\langle n, 0 | p_\alpha | m, 0 \rangle \langle m, 0 | p_\beta | n, 0 \rangle}{(E_{n,0} - E_{m,0})} .$$

This allows to write (4.10) in terms of the effective mass:

$$E_{n,\mathbf{k}} = E_{n,0} + \frac{\hbar^2}{2} \sum_{\alpha\beta} \left(\frac{1}{m^*(n)}\right)_{\alpha\beta} k_\alpha k_\beta . \quad (4.11)$$

4.4 The tight binding model

Not all electrons are valance electrons. In fact, most electrons are strongly bound by the Coulomb potential, i.e. are tightly bound to the nuclei. The tight binding model follows this idea, and assumes the potentials to be strong. This is the opposing limiting case, compared to the nearly free electron model.

We recall the construction of the lattice potential from section 4.1. The potential of an atom group, that belongs to a Bravais lattice position $\mathbf{R}_{\vec{j}}$ is $\tilde{v}(\mathbf{r} - \mathbf{R}_{\vec{j}})$. For the tight binding model, assume, that the problem for a single position has been solved, i.e.

$$H_a \varphi_n(\mathbf{r}) = E_n \varphi_n(\mathbf{r}) \quad \text{with} \quad H = \frac{-\hbar^2}{2m} \Delta + \tilde{v}(\mathbf{r}) .$$

Remark 4.4.1.

Although H_a is the Hamiltonian for an atom group, we still call it **single atom Hamiltonian**. In the same way, the φ_n are called single atom solutions, etc.

This solution is applicable for arbitrary Bravais lattice positions $\mathbf{R} \in L$ by

$$H_{a,\mathbf{R}} \varphi_n(\mathbf{r} - \mathbf{R}) = E_n \varphi_n(\mathbf{r} - \mathbf{R}) \quad \text{with} \quad H_a = \frac{-\hbar^2}{2m} \Delta + \tilde{v}(\mathbf{r} - \mathbf{R}) .$$

The full Hamiltonian from section 4.1 can be written as

$$H = \left(\frac{-\hbar^2}{2m} \Delta + \tilde{v}(\mathbf{r} - \mathbf{R}) \right) + \left(\sum_{L \ni \mathbf{R}' \neq \mathbf{R}} \tilde{v}(\mathbf{r} - \mathbf{R}') \right) \equiv H_{a,\mathbf{R}} + U_{\mathbf{R}(\mathbf{r})} . \quad (4.12)$$

It follows that

$$H\varphi_n(\mathbf{r} - \mathbf{R}') = H_{a,\mathbf{R}'}\varphi_n(\mathbf{r} - \mathbf{R}') + U_{\mathbf{R}'}(\mathbf{r})\varphi_n(\mathbf{r} - \mathbf{R}') . \quad (4.13)$$

4.4.1 Terminology and idea behind the tight binding model

By construction, the solutions $\varphi_n(\mathbf{r} - \mathbf{R})$ of the single atom Hamiltonian $H_{a,\mathbf{R}}$ describe particles, that are bound to the atom(s) at position \vec{R} . One says, the solutions $\varphi_n(\mathbf{r} - \mathbf{R})$ are **localized** at \vec{R} .

The interpretation of (4.12) is the following. The full Hamiltonian is the single atom Hamiltonian $H_{a,\mathbf{R}}$, with potential contributions from the remaining atoms $U_{\mathbf{R}}(\mathbf{r})$. The terms $\mathfrak{D}(\mathbf{r}) := R_{\mathbf{R}}(\mathbf{r})\varphi_n(\mathbf{r} - \mathbf{R})$ (see (4.13)) are called **overlaps** of the states with the potentials of the other atoms.

Remark 4.4.2.

The name overlap can be motivated from the probability amplitude. If $\mathfrak{D}(\mathbf{r})$ is non-zero at \mathbf{r} , this means that $|\mathfrak{D}(\mathbf{r})| \neq 0$, which in turn means, that neither $|\varphi_n(\mathbf{r} - \mathbf{R})|$ nor $|U_{\mathbf{R}}(\mathbf{r})|$ are zero at \mathbf{r} . This is the case for all \mathbf{r} , where the hyper surfaces $\mathbf{r} \rightarrow |U_{\mathbf{R}}(\mathbf{r})|$ and $\mathbf{r} \rightarrow |\varphi_n(\mathbf{r} - \mathbf{R})|$ overlap.

If the overlap tends to zero, the localized solutions become solutions of the full Hamiltonian.

The idea of the tight binding model is to assume, that the overlap is small, i.e. confined to neighboring atoms. Then, an approximative solution of the full Hamiltonian can be constructed from localized states. The smaller the overlap, the better the approximation is.

4.4.2 No overlap

Define a function that satisfies the Bloch condition by

$$\psi_{n,\mathbf{k}}(\mathbf{r}) := \frac{1}{\sqrt{N}} \sum_{\mathbf{R} \in L} e^{i\langle \mathbf{k}, \mathbf{R} \rangle} \varphi_n(\mathbf{r} - \mathbf{R}) .$$

Indeed, it holds that:

$$\begin{aligned} \sqrt{N}\psi_{n,\mathbf{k}}(\mathbf{r} + \mathbf{R}') &= \sum_{\mathbf{R} \in L} e^{i\langle \mathbf{k}, \mathbf{R} \rangle} \varphi_n(\mathbf{r} - \mathbf{R} + \mathbf{R}') \\ &= \sum_{\mathbf{R}' = \mathbf{R} - \mathbf{R}'} e^{i\langle \mathbf{k}, \mathbf{R}' \rangle} e^{i\langle \mathbf{k}, \mathbf{R}'' \rangle} \varphi_n(\mathbf{r} - \mathbf{R}'') \\ &= e^{i\langle \mathbf{k}, \mathbf{R}' \rangle} \sqrt{N}\psi_{n,\mathbf{k}}(\mathbf{r}) . \end{aligned}$$

A Bloch function is then given by (see (3.5)):

$$v_{n,\mathbf{k}}(\mathbf{r}) = e^{-i\langle \mathbf{k}, \mathbf{r} \rangle} \psi_{n,\mathbf{k}}(\mathbf{r}) .$$

If the overlap is zero, i.e. $\mathfrak{D}_{\mathbf{R}}(\mathbf{r}) = 0$, then because of (4.13) and since E_n does not depend on \mathbf{R} , the functions $\psi_{n,\mathbf{k}}(\mathbf{r})$ are eigen states of the full Hamiltonian:

$$H\psi_{n,\mathbf{k}}(\mathbf{r}) = E_n\psi_{n,\mathbf{k}}(\mathbf{r}) .$$

The consequence are dispersionless bands, i.e. E_n does not depend on \mathbf{k} .

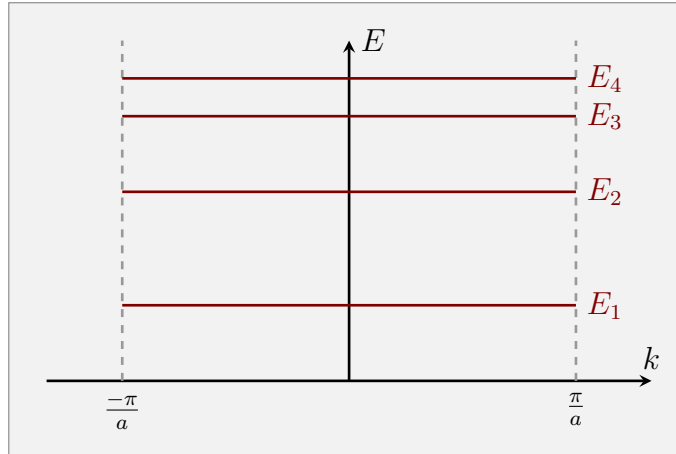


Figure 4.4: Dispersionless bands in one dimension.

4.4.3 General approximation for the overlap case

If the overlap is non-zero, the single atom solutions will no longer work exactly. To approximate the energies $E_{n,\mathbf{k}}$, we want to use the Ritz variation method. Having the idea of localization in the tight binding in mind, we generalize the definition of the $\psi_{n,\mathbf{k}}$ to

Definition 4.4.3.

The **trail functions** for the tight binding model are defined by

$$\psi_{n,\mathbf{k}}(\mathbf{r}) := \frac{1}{\sqrt{N}} \sum_{\mathbf{R} \in L} e^{i\langle \mathbf{k}, \mathbf{R} \rangle} \phi_n(\mathbf{r} - \mathbf{R}) ,$$

where the functions $\phi_{n,\mathbf{R}}(\mathbf{r}) \equiv \phi_n(\mathbf{r} - \mathbf{R})$ are localized at \mathbf{R}

In the general case, the trail functions are not orthonormal:

Lemma 4.4.4.

For the $\psi_{n,\mathbf{k}}$ it holds that:

$$\langle \psi_{n,\mathbf{k}} | \psi_{m,\ell} \rangle = \delta_{nm} + \sum_{\mathbf{K} \neq 0 \in L} e^{-i\langle \mathbf{k}, \mathbf{K} \rangle} \alpha_{nm}(\mathbf{K}) ,$$

where $\alpha_{nm}(\mathbf{K}) = \int_{\mathbb{R}^3} d\mathbf{r}'^3 \overline{\phi_n(\mathbf{r}' - \mathbf{K})} \phi_m(\mathbf{r}')$.

Proof 4.4.5.

$$\begin{aligned} \langle \psi_{n,\mathbf{k}} | \psi_{m,\ell} \rangle &= \frac{1}{N} \sum_{\mathbf{R}, \mathbf{G} \in L} e^{-i\langle \mathbf{k}, \mathbf{R} \rangle} e^{i\langle \mathbf{k}, \mathbf{G} \rangle} \int_{\mathbb{R}^3} d\mathbf{r}^3 \overline{\phi_n(\mathbf{r} - \mathbf{R})} \phi_m(\mathbf{r} - \mathbf{G}) \\ &\stackrel{r' = r - \mathbf{G}}{=} \frac{1}{N} \sum_{\mathbf{R}, \mathbf{G} \in L} e^{i\langle \mathbf{k}, \mathbf{G} - \mathbf{R} \rangle} \int_{\mathbb{R}^3} d\mathbf{r}'^3 \overline{\phi_n(\mathbf{r}' + \mathbf{G} - \mathbf{R})} \phi_m(\mathbf{r}') \end{aligned}$$

$$\begin{aligned}
& \stackrel{=}{=} \frac{1}{N} \sum_{\mathbf{R} \in L} \sum_{\mathbf{K} \in L} \underbrace{e^{-i\langle \mathbf{k}, \mathbf{K} \rangle} \int_{\mathbb{R}^3} dr'^3 \overline{\phi_n(\mathbf{r}' - \mathbf{K})} \phi_m(\mathbf{r}')}_{\text{does not depend on } \mathbf{R} \text{ and } \sum_{\mathbf{R}} 1 = N} \\
& = \sum_{\mathbf{K} \in L} e^{-i\langle \mathbf{k}, \mathbf{K} \rangle} \int_{\mathbb{R}^3} dr'^3 \overline{\phi_n(\mathbf{r}' - \mathbf{K})} \phi_m(\mathbf{r}') \\
& \equiv \sum_{\mathbf{K} \in L} e^{-i\langle \mathbf{k}, \mathbf{K} \rangle} \alpha_{nm}(\mathbf{K}) \\
& = \alpha_{nm}(0) + \sum_{\mathbf{K} \neq 0 \in L} e^{-i\langle \mathbf{k}, \mathbf{K} \rangle} \alpha_{nm}(\mathbf{K}) \\
& = \delta_{nm} + \sum_{\mathbf{K} \neq 0 \in L} e^{-i\langle \mathbf{k}, \mathbf{K} \rangle} \alpha_{nm}(\mathbf{K}) .
\end{aligned}$$

In the last line we used, that $\{\phi_n(\mathbf{r})\}$ are orthonormal, i.e. $\alpha_{nm}(0) = \delta_{nm}$. \square

For the Ritz variation method, we need to calculate:

$$E_{n,\mathbf{k}} = \frac{\langle \psi_{n,\mathbf{k}} | H | \psi_{n,\mathbf{k}} \rangle}{\langle \psi_{n,\mathbf{k}} | \psi_{n,\mathbf{k}} \rangle} .$$

The term $\langle \psi_{n,\mathbf{k}} | \psi_{n,\mathbf{k}} \rangle$ has already been calculated in lemma 4.4.4. We also have to calculate $\langle \psi_{n,\mathbf{k}} | H | \psi_{n,\mathbf{k}} \rangle$. With (4.13) and the definition of $U_{\mathbf{R}}(\mathbf{r}) = \sum_{L \ni \mathbf{R}' \neq \mathbf{R}} \tilde{v}(\mathbf{r} - \mathbf{R})$ one obtains

$$\begin{aligned}
\langle \psi_{n,\mathbf{k}} | H | \psi_{n,\mathbf{k}} \rangle & = \sum_{\mathbf{R}} \langle \psi_{n,\mathbf{k}} | H_{a,\mathbf{R}} | e^{i\langle \mathbf{k}, \mathbf{R} \rangle} \phi_{n,\mathbf{R}} \rangle \\
& + \frac{1}{N} \sum_{\mathbf{R}_1, \mathbf{R}_2} \sum_{\mathbf{R}_3 \neq \mathbf{R}_2} e^{i\langle \mathbf{k}, \mathbf{R}_2 - \mathbf{R}_1 \rangle} \int_{\mathbb{R}^3} dr^3 \overline{\phi_n(\mathbf{r} - \mathbf{R}_1)} \phi_n(\mathbf{r} - \mathbf{R}_2) \tilde{v}(\mathbf{r} - \mathbf{R}_3) . \tag{4.14}
\end{aligned}$$

The functions $\overline{\phi_n(\mathbf{r} - \mathbf{R}_1)}$, $\phi_n(\mathbf{r} - \mathbf{R}_2)$ and $\tilde{v}(\mathbf{r} - \mathbf{R}_3)$ are localized at \mathbf{R}_1 , \mathbf{R}_2 and \mathbf{R}_3 respectively. If $\mathbf{R}_1 \neq \mathbf{R}_2 \neq \mathbf{R}_3$, i.e. the \mathbf{R}_j are pair wise disjoint, for all \mathbf{r} , there is a product of two functions that have small overlap, such that this case will be neglected. If however two of the \mathbf{R}_j 's are equal, then there is a region, where the overlap of the corresponding functions is large. These cases can not be neglected. The case $\mathbf{R}_1 = \mathbf{R}_2 = \mathbf{R}_3$ is forbidden. In the following we consider the remaining two possible cases:

$$\mathbf{R}_1 = \mathbf{R}_2 \neq \mathbf{R}_3$$

Fixing $\mathbf{R}_2 = \mathbf{R}_1$, the double sum $\sum_{\mathbf{R}_1, \mathbf{R}_2}$ becomes a single sum $\sum_{\mathbf{R}_1}$. Also, we will use the substitution $\mathbf{r}' = \mathbf{r} - \mathbf{R}_1$, such that

$$U_{\mathbf{R}_1}(\mathbf{r}' + \mathbf{R}_1) = \sum_{\mathbf{R}_3 \neq \mathbf{R}_1} \tilde{v}(\mathbf{r}' + \mathbf{R}_1 - \mathbf{R}_3) = \sum_{\mathbf{K} \in L} \tilde{v}(\mathbf{r}' - \mathbf{K}) = U_0(\mathbf{r}') ,$$

since for $\mathbf{R}_3 \neq \mathbf{R}_1$ the difference $\mathbf{R}_1 - \mathbf{R}_3$ never is zero. Hence the second line of (4.14) becomes:

$$\frac{1}{N} \sum_{\mathbf{R}_1} \sum_{\mathbf{R}_3 \neq \mathbf{R}_1} e^{i\langle \mathbf{k}, 0 \rangle} \int_{\mathbb{R}^3} dr^3 \overline{\phi_n(\mathbf{r} - \mathbf{R}_1)} \phi_n(\mathbf{r} - \mathbf{R}_1) \tilde{v}(\mathbf{r} - \mathbf{R}_3)$$

$$\begin{aligned}
&= \frac{1}{N} \sum_{\mathbf{R}_1} e^{i\langle \mathbf{k}, 0 \rangle} \int_{\mathbb{R}^3} dr^3 \overline{\phi_n(\mathbf{r} - \mathbf{R}_1)} \phi_n(\mathbf{r} - \mathbf{R}_1) U_{\mathbf{R}_1}(\mathbf{r}) \\
&= \frac{1}{N} \sum_{\mathbf{R}_1} \int_{\mathbb{R}^3} dr'^3 \overline{\phi_n(\mathbf{r}')}\phi_n(\mathbf{r}') \underbrace{U_{\mathbf{R}_1}(\mathbf{r}' + \mathbf{R}_1)}_{=U_0(\mathbf{r}')} \\
&= \frac{1}{N} \int_{\mathbb{R}^3} dr'^3 \overline{\phi_n(\mathbf{r}')}\phi_n(\mathbf{r}') U_0(\mathbf{r}') \sum_{\mathbf{R}_1} 1 \\
&= \int_{\mathbb{R}^3} dr'^3 \overline{\phi_n(\mathbf{r}')}\phi_n(\mathbf{r}') U_0(\mathbf{r}') \\
&\equiv \beta_n .
\end{aligned}$$

This is just a constant.

$$\mathbf{R}_1 = \mathbf{R}_3 \neq \mathbf{R}_2$$

In this case, we will substitute $\mathbf{R} = \mathbf{R}_1 - \mathbf{R}_2$. Then, the sum $\sum_{\mathbf{R}_2 \neq \mathbf{R}_1}$ becomes $\sum_{\mathbf{R} \neq 0}$. For the second line of (4.14) we find:

$$\begin{aligned}
&\frac{1}{N} \sum_{\mathbf{R}_1} \sum_{\mathbf{R}_2 \neq \mathbf{R}_1} e^{i\langle \mathbf{k}, \mathbf{R}_2 - \mathbf{R}_1 \rangle} \int_{\mathbb{R}^3} dr^3 \overline{\phi_n(\mathbf{r} - \mathbf{R}_1)} \phi_n(\mathbf{r} - \mathbf{R}_2) \tilde{v}(\mathbf{r} - \mathbf{R}_1) \\
&\stackrel{r' = r - \mathbf{R}_2}{=} \frac{1}{N} \sum_{\mathbf{R}_1} \sum_{\mathbf{R}_2 \neq \mathbf{R}_1} e^{i\langle \mathbf{k}, \mathbf{R}_2 - \mathbf{R}_1 \rangle} \\
&\quad \int_{\mathbb{R}^3} dr'^3 \overline{\phi_n(\mathbf{r}' - \mathbf{R}_1 + \mathbf{R}_2)} \phi_n(\mathbf{r}') \tilde{v}(\mathbf{r} - \mathbf{R}_1 + \mathbf{R}_2) \\
&= \stackrel{\mathbf{R} = \mathbf{R}_1 - \mathbf{R}_2}{=} \frac{1}{N} \sum_{\mathbf{R} \neq 0} e^{-i\langle \mathbf{k}, \mathbf{R} \rangle} \int_{\mathbb{R}^3} dr'^3 \overline{\phi_n(\mathbf{r}' - \mathbf{R})} \phi_n(\mathbf{r}') \tilde{v}(\mathbf{r} - \mathbf{R}) \sum_{\mathbf{R}_1} 1 \\
&= \sum_{\mathbf{R} \neq 0} e^{-i\langle \mathbf{k}, \mathbf{R} \rangle} \int_{\mathbb{R}^3} dr'^3 \overline{\phi_n(\mathbf{r}' - \mathbf{R})} \phi_n(\mathbf{r}') \tilde{v}(\mathbf{r}' - \mathbf{R}) \\
&\equiv \sum_{\mathbf{R} \neq 0} e^{-i\langle \mathbf{k}, \mathbf{R} \rangle} \lambda_n(\mathbf{R}) .
\end{aligned}$$

Collecting the results

Collecting the results, we obtain:

$$\boxed{E_{n,\mathbf{k}} = \mathcal{E}_n + \frac{\beta_n + \sum_{\mathbf{R} \neq 0} e^{-i\langle \mathbf{k}, \mathbf{R} \rangle} \lambda_n(\mathbf{R})}{1 + \sum_{\mathbf{R} \neq 0 \in L} e^{-i\langle \mathbf{k}, \mathbf{R} \rangle} \alpha_{nn}(\mathbf{R})}} . \quad (4.15)$$

Remark 4.4.6.

Here we defined

$$\mathcal{E}_n = \frac{1}{\langle \psi_{n,\mathbf{k}} | \psi_{n,\mathbf{k}} \rangle} \sum_{\mathbf{R}} \langle \psi_{n,\mathbf{k}} | H_{a,\mathbf{R}} | e^{i\langle \mathbf{k}, \mathbf{R} \rangle} \phi_{n,\mathbf{R}} \rangle .$$

Although at first not known, if the $\phi_{n,\mathbf{R}}$ diagonalize the single atom Hamiltonian i.e. $\langle \phi_{n,\mathbf{R}} | H_{a,\mathbf{R}} | \phi_{n',\mathbf{R}'} \rangle = c_n \delta_{nn'} \delta_{\mathbf{R}\mathbf{R}'}$, one finds

$$\mathcal{E}_n = c_n .$$

If one uses $\phi_{n,\mathbf{R}} = \varphi_{n,\mathbf{R}}$ one even obtains $\mathcal{E}_n = E_n$. Hence the notation.

We can expand the fraction of (4.15) by $1 - \sum_{\mathbf{R} \neq 0 \in L} e^{-i\langle \mathbf{k}, \mathbf{R} \rangle} \alpha_{nn}(\mathbf{R})$:

$$E_{n,\mathbf{k}} = \mathcal{E}_n + \frac{\beta_n + \sum_{\mathbf{R} \neq 0} e^{-i\langle \mathbf{k}, \mathbf{R} \rangle} (\lambda_n(\mathbf{R}) - \beta_n \alpha_{nn}(\mathbf{R})) + \sum_{\mathbf{R}, \mathbf{R}' \neq 0} A(\mathbf{R}, \mathbf{R}')}{1 - \left(\sum_{\mathbf{R} \neq 0 \in L} e^{-i\langle \mathbf{k}, \mathbf{R} \rangle} \alpha_{nn}(\mathbf{R}) \right)^2},$$

where the double sum term is $A(\mathbf{R}, \mathbf{R}') = e^{-i\langle \mathbf{k}, \mathbf{R} + \mathbf{R}' \rangle} \lambda_n(\mathbf{R}) \alpha_{nn}(\mathbf{R}')$. Now we assume, that $\sum_{\mathbf{R} \neq 0 \in L} e^{-i\langle \mathbf{k}, \mathbf{R} \rangle} \alpha_{nn}(\mathbf{R})$ and $\sum_{\mathbf{R} \neq 0} e^{-i\langle \mathbf{k}, \mathbf{R} \rangle}$ are small, such that the square terms

$$A(\mathbf{R}, \mathbf{R}') \quad \text{and} \quad \left(\sum_{\mathbf{R} \neq 0 \in L} e^{-i\langle \mathbf{k}, \mathbf{R} \rangle} \alpha_{nn}(\mathbf{R}) \right)^2$$

can be neglected. We define $\gamma_n(\mathbf{R}) = \lambda_n(\mathbf{R}) - \beta_n \alpha_{nn}(\mathbf{R})$ and obtain:

$$E_{n,\mathbf{k}} = \mathcal{E}_n + \sum_{\mathbf{R} \neq 0} e^{-i\langle \mathbf{k}, \mathbf{R} \rangle} \gamma_n(\mathbf{R}). \quad (4.16)$$

The Ritz variation approximation is now, to choose the ϕ_n to be depending on some parameters, and to minimize $E_{n,\mathbf{k}}$ with respect to these parameters.

4.4.4 Wannier functions

A first approximation for the trial functions would be to return to the initial definition of the $\psi_{n\mathbf{k}}$, i.e. to use $\phi_{n,\mathbf{k}}(\mathbf{r} - \mathbf{R}) = \varphi_{n,\mathbf{k}}(\mathbf{r} - \mathbf{R})$. However, for the $\varphi_{n,\mathbf{k}}(\mathbf{r} - \mathbf{R})$

$$\int_{\mathbb{R}^3} dr^3 \overline{\varphi_n(\mathbf{r} - \mathbf{R})} \varphi_{n'}(\mathbf{r} - \mathbf{R}') \neq \delta_{nn'} \delta_{\mathbf{R}\mathbf{R}'}.$$

Hence $\alpha_{nn'}(\mathbf{R}) \neq 0$ which is the reason why the $\psi_{n,\mathbf{k}}$ are not orthonormal. For that reason we formally define:

Definition 4.4.7.

The **Wannier functions** are functions $w_{n,\mathbf{R}}(\mathbf{r}) \equiv w_n(\mathbf{r} - \mathbf{R})$, that are localized at \mathbf{R} and that satisfy

$$\langle w_{n,\mathbf{R}} | w_{n',\mathbf{R}'} \rangle = \delta_{nn'} \delta_{\mathbf{R}\mathbf{R}'}.$$

Choosing the Wannier functions, i.e. $\phi_n(\mathbf{r} - \mathbf{R}) = w_n(\mathbf{r} - \mathbf{R})$, the $\psi_{n,\mathbf{k}}$ become orthonormal in the excitation:

$$\langle \psi_{n,\mathbf{k}} | \psi_{m,\ell} \rangle = \delta_{nm}.$$

Also, for the Wannier functions, $\alpha_{nn}(\mathbf{R}) = 0$ by definition, such that one obtains (4.16) from (4.15) without further assumptions.

However, the Wannier functions are not known (except for numerical approximations and in toy models). If one new the orthonormal solutions of the full Hamiltonian,

$H\Psi_{n,\mathbf{k}}(\mathbf{r}) = E_{n,\mathbf{k}}\Psi_{n,\mathbf{k}}(\mathbf{r})$ (yet, in this case one would not need the Wannier functions in the first place), they could be constructed by

$$w_n(\mathbf{r} - \mathbf{R}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k} \in \text{B.Z.}} e^{-i\langle \mathbf{k}, \mathbf{R} \rangle} \Psi_{n,\mathbf{k}}(\mathbf{r}) .$$

A direct calculation shows, that the orthonormality condition is satisfied. The localization property depends on the $\Psi_{n,\mathbf{k}}$.

4.4.5 Example: d -dimensional cubic lattice with nearest neighbor interaction

The lattice vectors of a cubic Bravais lattice are $\mathbf{R}_{\vec{m}} = a \cdot \vec{m}$. Nearest neighbor interaction means, that we assume, that the overlaps at $\mathbf{R} = 0$ are only meaningful for the nearest neighbors. For $\gamma_n(\mathbf{R})$ from (4.16) this assumption translated to

$$\gamma_n(\mathbf{R}) = \begin{cases} -t_n & , m_j = 1 \text{ for only one } j \\ 0 & , \text{else} \end{cases}$$

The minus sign can be motivated as follows. We assume $\alpha_{nn}(\mathbf{R})$ to be small (or zero in the case of Wannier functions). Then $\gamma_n(\mathbf{R}) \approx \lambda_n(\mathbf{R})$, which describes an attracting potential. Plugging in into (4.16) and using trigonometric identities yields:

$$\begin{aligned} E_{n,\mathbf{k}} &= c_n - t_n \sum_{j=1}^d \left(e^{-ik_j a} + e^{ik_j a} \right) \\ &= c_n - t_n \sum_{j=1}^d \cos(k_j a) . \end{aligned}$$

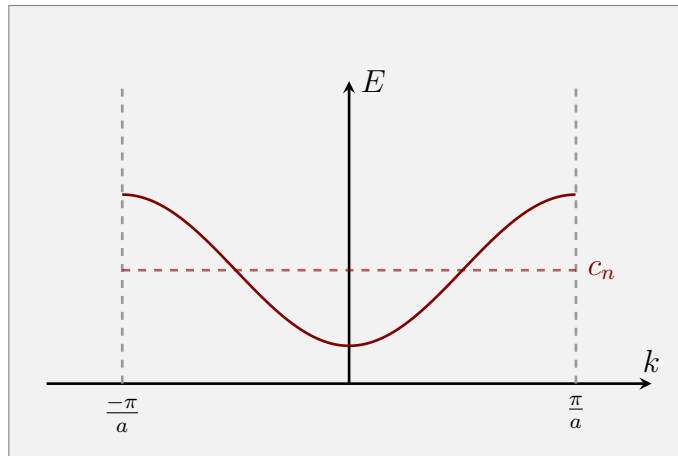


Figure 4.5: Dispersion relation of a one dimensional cubic lattice with nearest neighbor interaction.

4.5 Density of electron states and Fermi surface

As with phonons, we can define an electron state density.

Definition 4.5.1.

The **electron state density** $\rho(E) dE$ is defined by

$$\rho(E) = \frac{1}{N} \sum_{n,\sigma} \sum_{\mathbf{k} \in \text{B.z.}} \delta(E - E_n(\mathbf{k}, \sigma)) ,$$

where n is the band index as usual and σ is the spin index.

If the energies do not depend on the spin (e.g. free electrons), the sum over the spins results in a factor 2.

Contrary to phonons, electrons are fermion. Because of the Pauli exclusion principle, the electron can not all be in the lowest energy state. Each state (n, \mathbf{k}, σ) can only be occupied by one electron. The ground state is thus obtained by successively filling the lowest energy states. The energy, occupied by the last electron added to obtain the ground state, i.e. the largest occupied energy state, is called Fermi energy E_F . With the electron state density, the Fermi energy can be formally defined as follows:

Definition 4.5.2.

The **Fermi energy** E_F is the upper boundary, such that

$$\int_0^{E_F} \rho(E) dE = N_e ,$$

where N_e is the number of electrons.

A ready example for the electron state density is the free electron model in the continuum approximation ($\sum_{\mathbf{k}} \rightarrow \frac{V}{(2\pi)^3} \int d^3k$). Since the energy bands are degenerate at the boundary of the Brillouin zone, the sum of the bands becomes:

$$\sum_{\mathbf{K} \in \tilde{L}} E_{\mathbf{K}}(\mathbf{k}) = \sum_{\mathbf{K} \in \tilde{L}} \frac{\hbar^2 \|\mathbf{k} - \mathbf{K}\|^2}{2m} = E(\mathbf{k}) = \frac{\hbar^2 \|\mathbf{k}\|^2}{2m} ,$$

where \mathbf{k} is no longer restricted to the Brillouin zone on the right hand side. Then, the electron state density can be calculated by integration,¹ since it does not depend on the spin:

$$\begin{aligned} \rho(E) &= \frac{1}{N} \sum_{\mathbf{K}, \sigma} \sum_{\mathbf{k} \in \text{B.z.}} \delta \left(E - \frac{\hbar^2 \|\mathbf{k} - \mathbf{K}\|^2}{2m} \right) \\ &= \frac{2}{N} \sum_{\mathbf{k} \in \tilde{L}} \delta \left(E - \frac{\hbar^2 \|\mathbf{k}\|^2}{2m} \right) \approx \frac{2V}{(2\pi)^3 N} \int_{\mathbb{R}^3} dk^3 \delta \left(E - \frac{\hbar^2 \|\mathbf{k} - \mathbf{K}\|^2}{2m} \right) \\ &= \frac{8\pi V}{(2\pi)^3 N} \int_0^\infty dr r^2 \delta \left(E - \frac{\hbar^2 r^2}{2m} \right) = \frac{8\pi V}{(2\pi)^3 N} \int_0^\infty dr r^2 \frac{\delta(r - \frac{1}{\hbar} \sqrt{2mE})}{\frac{\hbar}{m} \sqrt{2mE}} \\ &= \frac{Vm}{N\pi^2 \hbar^3} \sqrt{2mE} . \end{aligned}$$

¹For a delta function it holds that $\delta(f(x)) = \sum_j \frac{\delta(x - x_j)}{|f'(x_j)|}$, where the $\{x_j\}$ are the zeros of $f(x)$.

For the Fermi energy it follows that

$$N_e = \int_0^{E_F} dE \frac{Vm}{N\pi^2\hbar^3} \sqrt{2mE} = \frac{V}{3N\pi^2\hbar^3} (2mE_F)^{\frac{3}{2}},$$

$$\Rightarrow \boxed{E_F = \frac{\hbar^2}{2m} \left(\frac{3N\pi^2 N_e}{V} \right)^{\frac{2}{3}}}.$$

Definition 4.5.3.

The **Fermi surface** is $S(E_F)$, where $S(E)$ is defined by

$$S(E) = \{\mathbf{k} \in \text{B.z.} \mid E_n(\mathbf{k}) = E\}.$$

The Fermi surface separates the momentum space into the part that is occupied in the ground state ($E_n(\mathbf{k}) \leq E_F$) and the part that is unoccupied ($E > E_F$). For the free electrons, it is a sphere with radius $k_F = \frac{\sqrt{2mE_F}}{\hbar}$.

4.6 Thermodynamics of electrons in a solid

As in subsection 3.3.2, the occupation numbers $\{n_{\ell, \mathbf{k}, \sigma}\}$ describe the micro state. Here $n_{\ell, \mathbf{k}, \sigma}$ is the number of electrons in the $(\ell, \mathbf{k}, \sigma)$ state with energy $E_\ell(\mathbf{k}, \sigma)$. Since electrons are fermions, the $n_{\ell, \mathbf{k}, \sigma}$ are either 1 or 0. Another result from statistical mechanics is the so called **Fermi-Dirac statistics**:

$$\langle n_{\ell, \mathbf{k}, \sigma} \rangle \equiv f(E_\ell(\mathbf{k}, \sigma)) = \frac{1}{\exp\left(\frac{E_\ell(\mathbf{k}, \sigma) - \mu}{k_B T}\right) + 1}.$$

The total number of electrons is given by

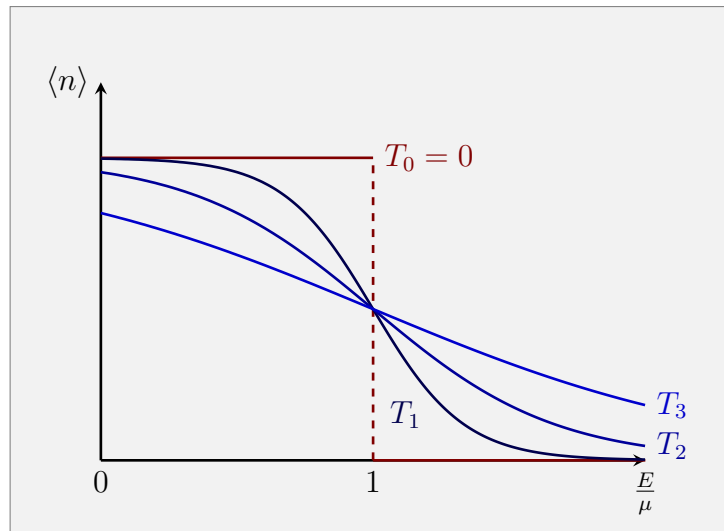


Figure 4.6: Energy dependence of the particle number expectation value for the Fermi-Dirac statistics for different temperatures $T_0 = 0 < T_1 < T_2 < T_3$.

$$N_e = \sum_{\ell, \mathbf{k}, \sigma} n_{\ell, \mathbf{k}, \sigma}.$$

We chose the **chemical potential** μ , such that the expectation value of the electron number is the electron number:

$$N_e \stackrel{!}{=} \langle N_e \rangle = \sum_{\ell, \mathbf{k}, \sigma} \langle n_{\ell, \mathbf{k}, \sigma} \rangle = \sum_{\ell, \mathbf{k}, \sigma} f(E_{\ell}(\mathbf{k}, \sigma)) = \int_{\mathbb{R}} \rho(E) f(E) dE .$$

The last step can be proven in exactly the same way² as lemma 3.3.6. This also yields:

$$U = \langle E \rangle = \sum_{\ell, \mathbf{k}, \sigma} \langle n_{\ell, \mathbf{k}, \sigma} \rangle E_{\ell}(\mathbf{k}, \sigma) = \int_{\mathbb{R}} E \rho(E) f(E) dE .$$

4.6.1 Sommerfeld expansion

The Sommerfeld expansion is a method to evaluate integrals of the type $\int_{-\infty}^{\infty} h(E) f(E) dE$, where $f(E)$ is the Fermi-Dirac statistics. However, for the method to work, $h(E)$ has to be smooth and fast decaying for $E \rightarrow -\infty$.

Remark 4.6.1 (Warning of application).

In the case of interest, $h(E) = E\rho(E)$ is not even a function, but the kernel representation of a singular distribution. To solve this, one can smear out the delta functions of $\rho(E)$ to Gaussian functions, obtaining $\tilde{\rho}(E)$. Then $E\tilde{\rho}(E)$ is smooth, yet fast decaying. This can be fixed by multiplying with an appropriate cutoff function, such that one obtains $h(E) = Ec(E)\tilde{\rho}(E)$. This discussion is usually omitted in the physical literature and one identifies $\rho(E) = c(E)\tilde{\rho}(E)$. In the following subsections, we will do the same, but keeping this regularization in mind.

With this assumption fast decay for $E \rightarrow -\infty$ it follows the antiderivative of $h(E)$ also has fast decay, up to a constant. So let $H(E) + c$ be an antiderivative of $h(E)$, such that $H(E)$ has fast decay for $E \rightarrow -\infty$:

$$H(E) = \int_{-\infty}^E h(A) dA .$$

Using partial integration, we find:

$$\begin{aligned} \int_{-\infty}^{\infty} h(E) f(E) dE &= [H(E) f(E)]_{-\infty}^{\infty} + [c f(E)]_{-\infty}^{\infty} \\ &\quad - \int_{-\infty}^{\infty} c f'(E) dE - \int_{-\infty}^{\infty} H(E) f'(E) dE \\ &= [H(E) f(E)]_{-\infty}^{\infty} + [c f(E)]_{-\infty}^{\infty} \\ &\quad - [c f(E)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} H(E) f'(E) dE \\ &= [H(E) f(E)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} H(E) f'(E) dE \\ &= \int_{-\infty}^{\infty} H(E) (-f'(E)) dE . \end{aligned}$$

²Note that in the lemma $e(\hbar\omega) = \hbar\omega g(\hbar\omega)$, if $g(\hbar\omega)$ is the Bose-Einstein statistics.

In the last line, $[H(E)f(E)]_{-\infty}^{\infty}$ vanishes, because $H(E) \rightarrow 0$ for $E \rightarrow -\infty$ and $f(E) \rightarrow 0$ for $E \rightarrow \infty$. The derivative of the Fermi-Dirac statistics is

$$-f'(E) = \frac{1}{k_B T} \frac{1}{\left(\exp\left(\frac{E_\ell(\mathbf{k}, \sigma) - \mu}{k_B T}\right) + 1\right) \left(\exp\left(-\frac{E_\ell(\mathbf{k}, \sigma) - \mu}{k_B T}\right) + 1\right)}$$

This is a function that is symmetric around $E = \mu$ and decays for $E \rightarrow \pm\infty$. In

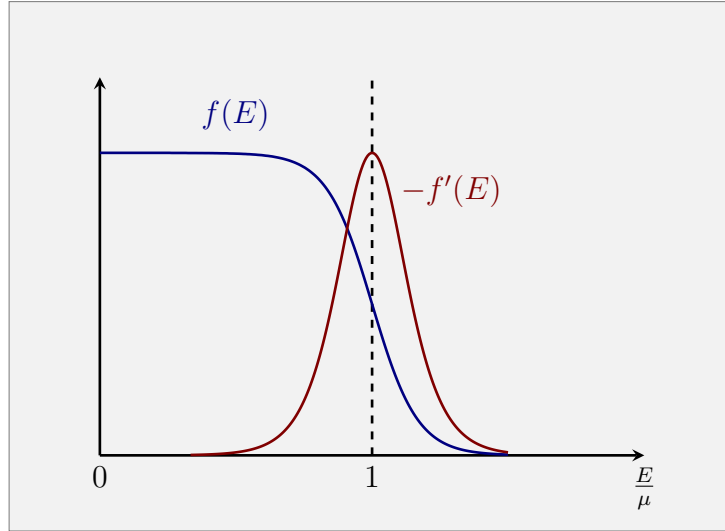


Figure 4.7: Fermi-Dirac statistics $f(E)$ and its negative derivative $-f'(E)$.

fact, the lower the temperature, the faster the decay, until it becomes a delta peak (see figure 4.7). This makes sense, as $f(E)_{T=0}$ is just a Heaviside function. Since the main contribution, where $H(E)f'(E)$ is non zero is in a small interval $[\mu - \varepsilon, \mu + \varepsilon]$ for sufficiently small temperatures, the function $H(E)$ can be developed as Taylor series:

$$\begin{aligned} H(E) &= H(\mu) + \sum_{n=1}^{\infty} \frac{H^{(n)}(\mu)}{n!} (E - \mu)^n \\ &= \int_{-\infty}^{\mu} h(E) dE + \sum_{n=1}^{\infty} \frac{h^{(n-1)}(\mu)}{n!} (E - \mu)^n . \end{aligned}$$

For the second line, we used the definition of $H(E)$ as constant less antiderivative of $h(E)$. For the Fermi-Dirac statistics it holds that $F(E) \rightarrow 1$ for $E \rightarrow -\infty$ and $f(E) \rightarrow 0$ for $E \rightarrow \infty$, such that

$$\int_{-\infty}^{\infty} -f'(E) dE = -[f(E)]_{-\infty}^{\infty} = -[0 - 1] = 1 .$$

We are also going to use the substitution $E = k_B T(x + \mu)$, such that $dE = k_B T dx$ and

$$-f'(E) = -f(k_B T(x + \mu)) = \frac{1}{k_B T} \frac{1}{(e^x + 1)(e^{-x} + 1)} .$$

With these results at hand, we can continue to calculate $\int_{-\infty}^{\infty} h(E)f(E) dE$:

$$\begin{aligned} \int_{-\infty}^{\infty} h(E)f(E) dE &= \int_{-\infty}^{\mu} h(E) dE + \sum_{n=1}^{\infty} \frac{h^{(n-1)}(\mu)}{n!} \int_{-\infty}^{\infty} (E - \mu)^n (-f'(E)) dE \\ &= \int_{-\infty}^{\mu} h(E) dE + \sum_{n=1}^{\infty} h^{(n-1)}(\mu) \int_{-\infty}^{\infty} \frac{(k_B T)^n x^n}{n!} \frac{1}{(e^x + 1)(e^{-x} + 1)} dx \end{aligned}$$

The function $\frac{x^n}{(e^x+1)(e^{-x}+1)}$ is odd, for odd n , such that the second integral vanishes whenever n is odd. As short notation, we define

$$a_m = \int_{-\infty}^{\infty} \frac{x^{2m}}{(2m)!} \frac{1}{(e^x + 1)(e^{-x} + 1)} dx .$$

Then we arrive at:

$$\boxed{\int_{-\infty}^{\infty} h(E)f(E) dE = \int_{-\infty}^{\mu} h(E) dE + \sum_{m=1}^{\infty} h^{(2m-1)}(\mu)(k_B T)^{2m} a_m .} \quad (4.17)$$

With further analysis (e.g. contour integration), one can show, that a_m can be expressed in term of the Riemann zeta function. However, we will only use the first order approximation (i.e. neglect terms of order $\mathcal{O}(T^4)$), such only a_1 is needed. It can be shown that $a_1 = \frac{\pi^2}{6}$.

$$\int_{-\infty}^{\infty} h(E)f(E) dE = \int_{-\infty}^{\mu} h(E) dE + \frac{\pi^2}{6} h^{(1)}(\mu)(k_B T)^2 + \mathcal{O}(T^4) . \quad (4.18)$$

4.6.2 The chemical potential

We return to the thermodynamics of electrons and calculate $\langle N_e \rangle$ with (4.18):

$$\langle N_e \rangle = \int_{-\infty}^{\infty} \rho(E)f(E) dE \approx \int_{-\infty}^{\mu} \rho(E) dE + \frac{\pi^2}{6} \rho'(\mu)(k_B T)^2$$

Looking at figure 4.6, we can assume that $\mu = \mu(T) \approx E_F$ and thus

$$\rho'(\mu) \approx \rho'(E_F) .$$

For $T = 0$ this has to be exact, as all states below $E = \mu = E_F$ are occupied. This assumption is compatible with the Sommerfeld expansion, which required small temperatures T in the first place. This allows to approximate:

$$\int_{E_F}^{\mu} \rho(E) dE \approx \rho(E_F)(\mu - E_F) .$$

Using the definition of the Fermi energy, 4.5.2, we get

$$\int_{-\infty}^{\infty} \rho(E)f(E) dE = \int_{-\infty}^{E_F} \rho(E) dE + \int_{E_F}^{\mu} \rho(E) dE \approx N_e + \rho(E_F)(\mu - E_F) .$$

$$\Rightarrow \quad \langle N_e \rangle = N_e + \rho(E_F)(E_F - \mu) + \frac{\pi}{6} \rho'(E_F)(k_B T)^2 .$$

The chemical potential is chosen, such that $\langle N_e \rangle = N_e$:

$$\Rightarrow \quad \boxed{\mu = E_F - \frac{\pi^2}{2} \frac{\rho'(E_F)}{\rho(E_F)} (k_B T)^2} . \quad (4.19)$$

4.6.3 The specific heat

To calculate the specific heat, we need to calculate the internal energy first. With the same assumptions as in the last subsection it follows³:

$$U = U_0 + E_F \rho(E_F)(\mu - E_F) + \frac{\pi}{6} (\rho(E_F) + E_F \rho'(E_F)) (k_B T)^2 .$$

Here $U_0 = \int_{-\infty}^{E_F} E \rho(E) dE$. Plugging (4.19) in we find:

$$\boxed{U = U_0 + \frac{\pi^2}{6} \rho(E_F) (k_B T)^2} .$$

The specific heat then reads:

$$\boxed{C_v(T) = \frac{\pi^2}{3V} \rho(E_F) k_B^2 T} .$$

³Here $h'(E_F) = \frac{d}{dE}(E\rho(E))|_{E_F} = \rho(E_F) + E_F \rho'(E_F)$

Outlook: Electron-electron interaction

In the last chapter, we treated the electrons as non-interacting and obtained qualitative results. However, because of their charge, electrons have a strong interaction. In this chapter we give an outlook on how to include the electron-electron interaction.

5.1 The electron Hamiltonian in second quantization

The non-interaction of the last chapter allowed us, to consider single particle problems. If we are to include interactions between electrons, this is no longer valid, and we need to pass to second quantization.

We assume to have found the solutions $\{\Psi_\alpha\}$ with quantum numbers $\{\alpha\}$ for the single electron Hamiltonian

$$H = \frac{-\hbar^2}{2m}\Delta + \sum_{\mathbf{R} \in L} \tilde{v}(\mathbf{r} - \mathbf{R})$$

from section 4.1. In fact, it is enough to choose a Hilbert basis $\{\Phi_\alpha\}$ of the single particle Hilbert space, such that the Wannier functions also work. At least, if their linear span is dense in the Hilbert space.

Remark 5.1.1.

To recover the wave function representation, from which we have started in chapter 2, one uses the **Slater determinant**. Let $\Psi^{(j)}(\mathbf{r}_j)$ denote the wave function of the j -th particle. Then the N particle wave function Ψ is

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \det \begin{pmatrix} \Psi^{(1)}(\mathbf{r}_1) & \dots & \Psi^{(1)}(\mathbf{r}_N) \\ \vdots & \ddots & \vdots \\ \Psi^{(N)}(\mathbf{r}_1) & \dots & \Psi^{(N)}(\mathbf{r}_N) \end{pmatrix}.$$

In fact, the fermionic Fock space is $F = \bigoplus_{j=0}^{\infty} \wedge^j(\mathcal{H})$, where \mathcal{H} is the single particle Hamiltonian. Then the Slater determinant is just a representation of the wedge product in $L^2(\mathbb{R}^n)$.

To keep the notation short, we write $|\alpha\rangle = \Phi_\alpha$. Then, the occupation number representation reads:

$$|n_{\alpha_1}, n_{\alpha_2}, \dots\rangle = |\alpha_1\rangle^{n_{\alpha_1}} \wedge |\alpha_2\rangle^{n_{\alpha_2}} \wedge \dots,$$

where $|\alpha_j\rangle^{n_{\alpha_j}} = |\alpha_j\rangle \wedge \dots \wedge |\alpha_j\rangle$ with n_{α_j} -times $|\alpha_j\rangle$ and $|\alpha_j\rangle^0 = 1$. However, since we are dealing with fermions, $n_{\alpha_j} \in \{0, 1\}$ anyway.

If the many electron Hamiltonian is given by

$$\mathbf{H} = \sum_{j=1}^{N_e} H_j + U,$$

where H_j is the single electron Hamiltonian, acting on the j -th electron and $U = V_{ee}$ the electron-electron interaction, the second quantized Hamiltonian reads:

$$\mathbf{H} = \sum_{\alpha,\beta} H_{\alpha\beta} c_{\alpha}^{\dagger} c_{\beta} + \frac{1}{2} \sum_{\alpha,\beta,\gamma,\delta} U_{\alpha\beta\gamma\delta} c_{\alpha}^{\dagger} c_{\beta}^{\dagger} c_{\delta} c_{\gamma}.$$

One has to be careful with the ordering of the indices for fermions, because of the anticommutativity. Indeed, the order of the indices of U does not match the order of the creation/annihilation operators. The coefficients are given by

$$H_{\alpha\beta} = \langle \alpha | H | \beta \rangle \quad \text{and} \quad U_{\alpha\beta\gamma\delta} = \langle \alpha | \langle \beta | U | \gamma \rangle | \delta \rangle.$$

5.2 Single band Hubbard model

With $U = V_{ee} = \sum_{j < k} \frac{e^2}{\|\mathbf{r}_j - \mathbf{r}_k\|}$ for the electron electron, the matrix elements of the second quantized Hamiltonian are

$$H_{\alpha\beta} = \int_{\mathbb{R}^3} \overline{\Phi_{\alpha}(\mathbf{r})} \left(\frac{-\hbar^2}{2m} \Delta + \sum_{\mathbf{R} \in L} \tilde{v}(\mathbf{r} - \mathbf{R}) \right) \Phi_{\beta}(\mathbf{r}) \, d\mathbf{r}^3,$$

$$U_{\alpha\beta\gamma\delta} = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \overline{\Phi_{\alpha}(\mathbf{r}') \Phi_{\beta}(\mathbf{r})} \frac{e^2}{\|\mathbf{r} - \mathbf{r}'\|} \Phi_{\gamma}(\mathbf{r}) \Phi_{\delta}(\mathbf{r}') \, d\mathbf{r}^3 \, d\mathbf{r}'^3.$$

If one chooses for the Φ_{α} the Bloch states, which are solutions of the single electron Hamiltonian H , the matrix elements $H_{\alpha\beta}$ are

$$H_{n,\mathbf{k},\sigma;n',\mathbf{k}',\sigma'} = E_n(\mathbf{k}, \sigma) \delta_{n,n'} \delta_{\mathbf{k},\mathbf{k}'} \delta_{\sigma,\sigma'}.$$

However, the interaction part is more difficult/impossible to calculate exactly. Hence, one needs approximations by choosing appropriate models.

One such model is the Hubbard model. For the single electron Hilbert space, we choose the Wannier functions $w_{n,\mathbf{R}}$ as basis and the tight binding approximation. This however means, that the matrix element $H_{\alpha\beta}$ is no longer diagonal (except for the spin). The second quantized Hamiltonian reads:

$$\mathbf{H} = \sum_{n,\mathbf{R};n',\mathbf{R}';\sigma} H_{n,\mathbf{R};n',\mathbf{R}'} c_{n,\mathbf{R},\sigma}^{\dagger} c_{n',\mathbf{R}',\sigma} + \frac{1}{2} \sum_{\substack{n_1,\mathbf{k}_1,\sigma_1;\dots; \\ n_4,\mathbf{k}_4,\sigma_4}} U_{n_1,\mathbf{k}_1,\sigma_1;\dots;n_1,\mathbf{k}_1,\sigma_1} c_{n_1,\mathbf{k}_1,\sigma_1}^{\dagger} c_{n_2,\mathbf{k}_2,\sigma_2}^{\dagger} c_{n_4,\mathbf{k}_4,\sigma_4} c_{n_3,\mathbf{k}_3,\sigma_3}.$$

So far, this is an exact expression of the second quantized Hamiltonian. For the **single band Hubbard model**, the approximations are:

- Consider only the valance band contributions, i.e. drop the band index sums.
- For the one body part, consider only next neighbor contributions of magnitude t , in agreement with the tight binding model. This means, that $\mathbf{R}' = \mathbf{R} + \Delta\mathbf{R}$, where $\Delta\mathbf{R}$ is a Bravais vector to the next neighbors:

$$t \sum_{\mathbf{R}, \Delta\mathbf{R}; \sigma} c_{\mathbf{R}, \sigma}^\dagger c_{\mathbf{R} + \Delta\mathbf{R}, \sigma}$$

- The main contribution of interaction is $\mathbf{R} = \mathbf{R}'$, i.e. the electrons are as close as possible. Because of the Pauli principle, electrons can only be at the same lattice position (which is a quantum number in the Wannier basis), if they have differing spins. With a parameter U for the interaction strength, the interaction part reads:

$$\frac{U}{2} \sum_{\mathbf{R}} c_{\mathbf{R}, \uparrow}^\dagger c_{\mathbf{R}, \uparrow} c_{\mathbf{R}, \downarrow}^\dagger c_{\mathbf{R}, \downarrow} .$$

With these assumptions, we arrive at the single band Hubbard Hamiltonian

$$\mathbf{H} = t \sum_{\mathbf{R}, \Delta\mathbf{R}; \sigma} c_{\mathbf{R}, \sigma}^\dagger c_{\mathbf{R} + \Delta\mathbf{R}, \sigma} + \frac{U}{2} \sum_{\mathbf{R}} c_{\mathbf{R}, \uparrow}^\dagger c_{\mathbf{R}, \uparrow} c_{\mathbf{R}, \downarrow}^\dagger c_{\mathbf{R}, \downarrow} .$$

5.3 Outlook

The single band Hubbard model is an important model, but just one of many. Also, there are other approximation methods, numerical and analytical ones. An important example would be the Hartree–Fock method. Furthermore, the interaction between phonons and electrons has to be investigated. . .

Index

A

- acoustic branch 22
- atomic position vectors 6
- atomic units 11

B

- band degeneracy 31
- band gap 34
- band index 30
- band structure 31
- Bloch condition 18
- Bloch function 18
- Born-Oppenheimer approximation .. 13
- Bose-Einstein statistics 23
- Bravais lattice 1
- Brillouin zone 9

C

- chemical potential 47
- coordination number 5

D

- Debye frequency 27
- Debye sphere 26
- Debye temperature 27
- Dulong-Petit law 24
- dynamical matrix 15

E

- effective mass 38
- electron state density 45
- electronic band 30

F

- Fermi energy 45
- Fermi surface 46
- Fermi-Dirac statistics 46

H

- harmonic approximation 14

L

- lattice vectors 1
- localized states 39

N

- nearest neighbor 4

O

- optical branch 22
- overlap of states and potentials 39

P

- periodic boundary condition 18
- phonon 21
- phonon branch 22
- phonon state density 27
- primitive cell 3
- primitive vectors 1

R

- reciprocal lattice 7

S

- single atom Hamiltonian 38
- single band Hubbard model 52
- Slater determinant 51
- solid Hamiltonian 11
- specific heat 26

T

- trial functions tight binding 40
- translation vector 2

U

- unit cell 3

W

- Wannier functions 43
- Wigner-Seitz cell 4

Bibliography

- [Czy16] G. Czycholl. *Theoretische Festkörperphysik Band 1*. Vol. 4. Springer-Verlag Berlin Heidelberg, 2016.
- [Dem10] W. Demtröder. *Experimentalphysik 3. Atome, Moleküle und Festkörper*. Vol. 4. Springer-Verlag Berlin Heidelberg, 2010.
- [Pru13] T. Pruschke. *Introduction to solid state theory*. Ed. by T. Pruschke. 2013. URL: http://www.tifr.res.in/~kbmaiti/Course/Note_Thomas.pdf.
- [Sch13] F. Schwabl. *Quantenmechanik (QM I): Eine Einführung*. 7th ed. Springer-Lehrbuch. Springer Berlin Heidelberg, 2013.