ΦSX

Advanced Topics of Theoretical Physics I

The electronic structure of matter

Peter E. Blöchl

Caution! This is an unfinished draft version
Errors are unavoidable!

Institute of Theoretical Physics; Clausthal University of Technology;
D-38678 Clausthal Zellerfeld; Germany;
`http://www.pt.tu-clausthal.de/atp/`

don't panic!

[1]

---

[1] To the title page: What is the meaning of ΦSX? Firstly, it sounds like "Physics". Secondly, the symbols stand for the three main pillars of theoretical physics: "X" is the symbol for the coordinate of a particle and represents Classical Mechanics. "Φ" is the symbol for the wave function and represents Quantum Mechanics and "S" is the symbol for the Entropy and represents Statistical Physics.

Foreword and Outlook

This book is work in progress. This book is not completed and not proofread. You can use this book currently as a basis for following through the material presented in the course. However, you should never blindly copy formulas an apply them. The chances that there are still errors is too large.

This book provides an introduction into the quantum mechanics of the interacting electron gas. It aims at students at the graduate level, that already have a good understanding on one-particle quantum mechanics. My aim is to include all proofs in a comprehensive and hopefully water-tight manner. I am also avoiding special systems of units. I consider this important in view of a large variety of notations, differing definitions etc. While the proofs have the advantage to train the student to perform the typical operations, it clutters the course with material far from applications. Therefore, I have placed many of the more detailed derivations into appendices. On the graduate level, the student should be able to follow the proofs without problem. It is strongly recommended to follow through the derivations.

A recommended reading are the Lecture notes[1] by Wolf-Dieter Schöne, which has been used in the preparation of this text. Another book that is very recommendable is the one by Fetter and Walecka[2]. Very useful has been also the book by Nolting[3]. While I did not follow the book of Mattuck[4]. A recommendable book for the reader, who wants to dig a little deeper into the density-functional theory is the Book by Eschrig[5]. Another excellent book on many particle physics, which however takes a quite different route from this text is the book by Negele and Orland[6].

The goal of the lecture is that the student

- understands the Born-Oppenheimer approximation and its main limitations,
- is able to construct the band structure of a non-interacting problem using the tight-binding method and is able to interpret a band structure,
- has a sound understanding of the connection between wave functions and the operator formalism of second quantization, i.e. creation and annihilation operators,
- understand the foundation of density functional theory, the dominant tool for first-principles electronic structure calculations,
- understands the role of Green's functions in many-particle physics and its relation to one-particle physics, and that he
- understands the main theorems underlying the perturbation expansion fo Green's functions, including Feynman diagrams.

I have tried to provide detailed and complete derivations of everything presented in this course. Some of the more extensive derivations are placed into an appendix. I strongly encourage the student to understand and rationalize these derivations. Only this will make the student familiar with the limitations of the theory and also with its flexibility. For a theoretical physicist this is of central importance, because he needs to be able to extend the existing theory, if necessary.

In the online version, there are many hyperreferences that allow to jump to the relevant information in the book. To avoid the additional cost for color printouts just for the references they are not visible. They still work, when you click on them. Here is a list of links set:

- List of contents
- references to equation numbers, figures, tables, sections, appendices, etc.
- citations will take you to the corresponding position in the list of references

Contents

Part I

Introduction to solid state physics

Topics:

1. Separate nuclear and electron dynamics
2. Spin-orbitals ans many-particle wave functions
3. non-interacting electrons: band structure
4. Hartree-Fock approximation
5. Density functional theory
6. Phonons
7. Transport theory: Boltzmann equation
    - Open systems
    - Decoherence
8. Magnetism

Chapter 1

The Standard model of solid-state physics

Solid state physics has, in contrast to elementary particle physics, the big advantage that the fundamental particles and their interactions are completely known. The particles are electrons and nuclei[1] and their interaction is described by electromagnetism. Gravitation is such a weak force at small distances that it is completely dominated by the electromagnetic interaction. The strong force, which is active inside the nucleus, has such a short range that it does not affect the relative motion of electrons and nuclei.

Some argue that solid-state physics is not at the very frontier of research, because its particles and interactions are fully known. However, a second look shows the contrary: While the constituents are simple, the complexity develops via the interaction of many of these simple particles. The simple constituents act together to form an entire zoo of phenomena of fascinating complexity.

The dynamics is described by the Schrödinger equation:

$$i\hbar\partial_t|\Phi\rangle = \hat{H}|\Phi\rangle \tag{1.1}$$

where the wave function depends on the coordinates of all electrons and nuclei in the system. In addition it depends on the spin degrees of freedom of the electrons. Thus, a wave function for N electrons and M nuclei has the form

$$\Phi_{\sigma_1,\dots,\sigma_N}(\vec{r}_1,\dots,\vec{r}_N,\vec{R}_1,\dots,\vec{R}_M)$$

where we denote the electronic coordinates by a lower-case $\vec{r}$ and the nuclear positions by an uppercase $\vec{R}$. The symbols σ_j are the spin indices, which may assume values $\sigma_j \in \{-\frac{1}{2},+\frac{1}{2}\}$. We will often use the notation $\sigma_j \in \{\downarrow,\uparrow\}$. The spin quantum number in z-direction is $s_z = \hbar\sigma$.

Often we will also combine position and spin index of the electrons in the form $\vec{x} = (\sigma,\vec{r})$, so that

$$\Phi(\vec{x}_1,\dots,\vec{x}_N,\vec{R}_1,\dots,\vec{R}_M) \stackrel{\text{def}}{=} \Phi_{\sigma_1,\dots,\sigma_N}(\vec{r}_1,\dots,\vec{r}_N,\vec{R}_1,\dots,\vec{R}_M)$$

For the integrations and summations we use the short-hand notation

$$\int d^4x \stackrel{\text{def}}{=} \sum_{\sigma\in\uparrow,\downarrow}\int d^3r \tag{1.2}$$

---

[1] The cautious reader may object that the nuclei are objects that are far from being fully understood. However, the only properties that are relevant for us are their charge, their mass, and their size.

Most important is the antisymmetry of the wave function with respect to interchange of two electrons, which is the cause for the **Pauli principle**

PAULI PRINCIPLE

$$\Phi(\ldots,\vec{x}_i,\ldots,\vec{x}_j,\ldots,\vec{R}_1,\ldots,\vec{R}_M)=-\Phi(\ldots,\vec{x}_j,\ldots,\vec{x}_i,\ldots,\vec{R}_1,\ldots,\vec{R}_M) \tag{1.3}$$

An electronic many-particle wave function for Fermions is antisymmetric under particle exchange. As a consequence, no two electrons with the same spin can occupy the same point in space or the same one-particle orbital.

The Standard Model

Now we arrive at the most important equation, which will be the basis of all that will be said in this book. This equation specifies the many-particle Hamiltonian of our standard model of solid-state physics.

HAMILTONIAN OF THE STANDARD MODEL

The Hamiltonian for a many particle system, such as a molecule or a solid, is

$$\begin{aligned}\hat{H}=&\underbrace{\sum_{j=1}^{M}\frac{-\hbar^2}{2M_j}\vec{\nabla}^2_{\vec{R}_j}}_{E_{kin,nuc}}+\underbrace{\sum_{i=1}^{N}\frac{-\hbar^2}{2m_e}\vec{\nabla}^2_{\vec{r}_i}}_{E_{kin,e}}\\&+\underbrace{\frac{1}{2}\sum_{i\neq j}^{M}\frac{e^2Z_iZ_j}{4\pi\epsilon_0|\vec{R}_i-\vec{R}_j|}}_{E_{C,nuc-nuc}}-\underbrace{\sum_{i=1}^{N}\sum_{j=1}^{M}\frac{e^2Z_j}{4\pi\epsilon_0|\vec{r}_i-\vec{R}_j|}}_{E_{C,e-nuc}}\\&+\underbrace{\frac{1}{2}\sum_{i\neq j}^{N}\frac{e^2}{4\pi\epsilon_0|\vec{r}_i-\vec{r}_j|}}_{E_{C,e-e}}\end{aligned} \tag{1.4}$$

The electrons are characterized by their charge $q=-e$, their mass m_e and their position $\vec{r}_j$. The nuclei are characterized by their masses M_j, their atomic numbers Z_j and their positions $\vec{R}_j$. The first term, denoted by $E_{kin,nuc}$, describes the kinetic energy of the nuclei. The second term, denoted by $E_{kin,e}$, describes the kinetic energy of the electrons. The third term, denoted by $E_{C,nuc-nuc}$, describes the electrostatic repulsion between the nuclei. The fourth term, denoted by $E_{C,e-nuc}$, describes the electrostatic attraction between electrons and nuclei. The last term, denoted by $E_{C,e-e}$, describes the electrostatic repulsion between the electrons.

Deficiencies of the standard model

No model is perfect. Therefore, let me discuss what I consider its main deficiencies. The following effects are missing:

- **Relativistic effects** are important for the heavier elements, because the deep Coulomb potential of the nuclei results in relativistic velocities for the electrons near the nucleus. Relativistic effects are important for magnetic effects such as the **magnetic anisotropy**.
- **Magnetic fields** have only a small effect on solids. The dominant magnetic properties of

materials do not originate from magnetic interactions, but result from exchange effects of the spin-distribution. Important are, however, **stray fields**, which tend to force the magnetization at a surface to be in-plane. Magnetic effects are also important to simulate the results of NMR or Mössbauer experiments, which are sensitive to the magnetic interaction of nuclei and electrons.

- **Size, shape and spin of the nuclei** play a role in certain nuclear experiments which measure isomer shifts, electric-field gradients and magnetic hyperfine parameters.
- **Photons**. The use of electrostatics is justified, because the electrons and nuclei are moving sufficiently slow that the electrostatic field responds instantaneously. Excitations by photons can be treated explicitly.

These effects are not necessarily ignored, but can be incorporated into the standard model, when required.

Dimensional Bottleneck

It is immediately clear that the standard model, as simple as it can be stated, is impossible to solve as such. To make this argument clear let us estimate the amount of storage needed to store the wave function of a simple molecule such as N_2. Each coordinate may be discretized into 100 grid points. For 2 nuclei and 14 electrons we need $100^{(3*16)}$ grid points. In addition, we need to consider the 2^{14} sets of spin indices. Thus, we need to store about $2^{14} \cdot 100^{3*16} \approx 10^{100}$ numbers. One typical hard disk can hold about 100 GByte=10^{14} Byte, which corresponds to 10^{13} complex numbers. A complex number occupies 16 byte. This implies that we would need 10^{87} hard discs to store a single wave function. A hard-disc occupies a volume of 5 cm×5cm×0.5 cm=1.25×10^{-5} m^3. The volume occupied by the hard discs with our wave function would be 10^{82} m^3 corresponding to a sphere with a radius of 10^{27} m or 10^{11} light years! I hope that this convinced you that there is a problem...

The problem described here is called the **dimensional bottle-neck**.

In order to make progress we need to simplify the problem. Thus, we have to make approximations and we will develop simpler models that can be understood in detail. This is what solid state physics is about.

In the following we will perform a set of common approximations, and will describe their limitations.

Chapter 2

Born-Oppenheimer approximation and beyond

The **Born-Oppenheimer approximation**[7, 8] simplifies the many-particle problem considerably by separating the electronic and nuclear coordinates. This method is similar, but not identical to the method of separation of variables. We refer here to the formalism as described in appendix VIII of the Book of Born and Huang[8].

2.1 Separation of electronic and nuclear degrees of freedom

Our ultimate goal is to determine the wave function $\Phi(\vec{x}_1, \dots, \vec{x}_N, \vec{R}_1, \dots, \vec{R}_M, t)$, which describes the electronic degrees of freedom $\vec{x}_1, \dots, \vec{x}_N$ and the atomic positions $\vec{R}_1, \dots, \vec{R}_M$ on the basis of the time-dependent Schrödinger equation Eq. 1.1

$$i\hbar\partial_t|\Phi(t)\rangle \overset{\text{Eq. 1.1}}{=} \hat{H}|\Phi(t)\rangle \tag{2.1}$$

The Hamiltonian $\hat{H}$ is the one given in Eq. 1.4.

In the Born-Oppenheimer approximation, one first determines the electronic eigenstates for a fixed set of atomic positions. This solution forms the starting point for the description of the dynamics of the nuclei.

We will formulate an ansatz for the time-dependent many-particle wave functions. For this ansatz we need the **Born-Oppenheimer wave functions**, which are the electronic wave function for a frozen set of nuclear positions.

Born-Oppenheimer wave functions and Born-Oppenheimer surfaces

Firstly, we construct the Born-Oppenheimer Hamiltonian $\hat{H}^{BO}$ by removing all terms that depend on the momenta of the nuclei from the many-particle Hamiltonian given in Eq. 1.4.

$$\begin{aligned}\hat{H}^{BO}(\vec{R}_1, \dots, \vec{R}_M) = &\underbrace{\sum_{i=1}^{N} \frac{-\hbar^2}{2m_e}\vec{\nabla}^2_{\vec{r}_i}}_{E_{kin,e}} + \underbrace{\frac{1}{2}\sum_{i\neq j}^{M} \frac{e^2 Z_i Z_j}{4\pi\epsilon_0|\vec{R}_i - \vec{R}_j|}}_{E_{C,nuc-nuc}} - \underbrace{\sum_{i=1}^{N}\sum_{j=1}^{M} \frac{e^2 Z_j}{4\pi\epsilon_0|\vec{r}_i - \vec{R}_j|}}_{E_{C,e-nuc}} \\ &+ \underbrace{\frac{1}{2}\sum_{i\neq j}^{N} \frac{e^2}{4\pi\epsilon_0|\vec{r}_i - \vec{r}_j|}}_{E_{C,e-e}}\end{aligned} \tag{2.2}$$

The Born-Oppenheimer Hamiltonian acts in the Hilbert space of electronic wave functions, and depends parametrically on the nuclear positions. In other words, the Born-Oppenheimer Hamiltonian does not contain any gradients acting on the nuclear positions.

In order to simplify the equations, I combine all the electronic coordinates and spin indices into a vector $\vec{x} = (\vec{x}_1, \ldots, \vec{x}_N)$ and the nuclear positions into a $3M$ dimensional vector $\vec{R} = (\vec{R}_1, \ldots, \vec{R}_M)$. Thus, the many-particle wave function acting on electrons and nuclei will be written as $\Phi(\vec{x}, \vec{R}, t) = \langle \vec{x}, \vec{R} | \Phi(t) \rangle$. I combine the nuclear masses into a diagonal matrix $\boldsymbol{M}$ with dimension $3M$.

In addition, I introduce another notation, where the wave function is a regular function of the atomic positions and a quantum state $|\Phi(\vec{R}, t)\rangle$ in the Hilbert space for the electrons, so that $\Phi(\vec{x}, \vec{R}, t) = \langle \vec{x} | \Phi(\vec{R}, t) \rangle$.

The full Hamiltonian of Eq. 1.4

$$\hat{H} = \frac{1}{2} \left(\frac{\hbar}{i} \vec{\nabla}_R \right) \boldsymbol{M}^{-1} \left(\frac{\hbar}{i} \vec{\nabla}_R \right) \hat{1} + \hat{H}^{BO}(\vec{R}) \tag{2.3}$$

is obtained by adding the nuclear kinetic energy to the Born-Oppenheimer Hamiltonian $\hat{H}^{BO}$. The unit operator $\hat{1}$ acts on the electronic Hilbert space.

The Born-Oppenheimer wave functions $|\Psi_n^{BO}(\vec{R})\rangle$ and the Born-Oppenheimer surfaces $E_n^{BO}(\vec{R})$ are obtained from the Born-Oppenheimer equation

$$\left[\hat{H}^{BO}(\vec{R}) - E_n^{BO}(\vec{R}) \right] \left| \Psi_n^{BO}(\vec{R}) \right\rangle = 0 \,. \tag{2.4}$$

Eq. 2.4 corresponds to a Schrödinger equation for the electrons feeling the static potential from the nuclei, which are fixed in space. The **Born-Oppenheimer surfaces** $E_n^{BO}(\vec{R})$ are simply the position-dependent energy eigenvalues of the Born-Oppenheimer Hamiltonian. For each atomic configuration we obtain a ground-state Born-Oppenheimer surface and infinitely many excited-state Born-Oppenheimer surfaces, which are labeled by the quantum number n. Similarly, the **Born-Oppenheimer wave functions** $\Psi_n^{BO}(\vec{x}_1, \ldots, \vec{x}_N, \vec{R}_1, \ldots, \vec{R}_M) = \langle \vec{x}_1, \ldots, \vec{x}_N | \Psi_n^{BO}(\vec{R}_1, \ldots, \vec{R}_M) \rangle$ depend parametrically on the nuclear positions $\vec{R}$.

The Born-Oppenheimer wave functions are chosen[1] to be orthogonal for each set of atomic positions, i.e.

$$\left\langle \Psi_m^{BO}(\vec{R}) \middle| \Psi_n^{BO}(\vec{R}) \right\rangle = \int d^{4N}x \; \Psi_m^{BO*}(\vec{x}, \vec{R}) \Psi_n^{BO}(\vec{x}, \vec{R}) = \delta_{m,n} \tag{2.5}$$

Ansatz for the electronic-nuclear wave function

Now we are ready to write down the Born-Huang ansatz for the many-particle wave function of a system of electrons and nuclei:

BORN-HUANG EXPANSION

$$\Phi(\vec{x}, \vec{R}, t) = \sum_n \left\langle \vec{x} \middle| \Psi_n^{BO}(\vec{R}) \right\rangle \phi_n(\vec{R}, t) \tag{2.6}$$

The multi-valued nuclear wave function obeys the normalization condition

$$\sum_n \int d^{3M}R \; \phi_n^*(\vec{R}, t) \phi_n(\vec{R}, t) = 1 \tag{2.7}$$

The many-particle wave function $\Phi(\vec{x}, \vec{R}, t)$ is expressed as a sum of products of the Born-Oppenheimer wave functions $|\Psi_n^{BO}(\vec{R})\rangle$ and time-dependent nuclear wave functions $\phi_n(\vec{R}, t)$. It is better to

---

[1] They are orthogonal whenever the energies differ, and each set of degenerate states can be orthonormalized.

think of Born-Oppenheimer wave functions and nuclear wave functions as multi-valued wave functions with a component for each ground or excited state labelled by n. While the Born-Oppenheimer wave functions depend on the positions of the nuclei, the nuclear wave functions $\phi_n(\vec{R}_1, \dots, \vec{R}_M, t)$ do no more refer to the electronic positions, except for the state index n.

Some caution is required to keep the symbols for the different types of wave functions apart

$\Phi(\vec{x}, \vec{R}, t)$	many-particle wave function of electrons and nuclei
$\Psi_n^{BO}(\vec{x}, \vec{R})$	Born-Oppenheimer wave function of electrons
$\phi_n(\vec{R}, t)$	component of the nuclear wave function

Physical meaning of the Born-Huang expansion: In order to convey some of the physical meaning of the quantities defined in the Born-Huang ansatz, let me anticipate a few observations that will be derived and stated more rigorously in the following sections.

- One Born-Oppenheimer surface can be considered as the potential energy surface governing the motion of the nuclei. The forces acting on the atoms are the negative gradients of the Born-Oppenheimer surface. The Born-Oppenheimer surface is the object that is parameterized in classical molecular dynamics simulations.
- The squared nuclear wave function describes the probability density $p_n(\vec{R}, t) \stackrel{\text{def}}{=} \phi_n^*(\vec{R}, t)\phi_n(\vec{R}, t)$ for the nuclei to be at positions $\vec{R}$ and for the electrons to be in the n-th excited state.
- If the electrons remain in the ground state, only the nuclear wave function $\phi_0(\vec{R}, t)$ with $n = 0$ differs from zero, and the atomic motion is entirely determined by the single valued potential energy $E_0^{BO}(\vec{R})$.
- An optical absorption event is usually sufficiently rapid, that the nuclei do not move by any appreciable distance during that process. This is the **Franck-Condon principle**. In that case we can estimate the absorption energy as the energy difference between two Born-Oppenheimer surfaces for the same nuclear coordinates.

The Born-Huang Expansion is not an approximation: It is important to realize that the Born-Huang expansion for the wave function is not an approximation. Every wave function can exactly be represented in this form. For any given vector R and time t, the wave function $\Phi(\vec{x}, \vec{R}, t)$ is a function of the electronic coordinates $\vec{x}$. This function can be expanded into any complete basisset $\{|\bar{\Psi}_n\rangle\}$ in the form

$$\Phi(\vec{x}, \vec{R}, t) = \sum_n \left\langle \vec{x} \middle| \bar{\Psi}_n \right\rangle \bar{c}_n(\vec{R}, t) \tag{2.8}$$

with complex coefficients $\bar{c}(\vec{R})$. Because the wave function changes with $\vec{R}$ and t, the coefficients depend on nuclear positions and time.

With the same right, one can choose a different complete basisset $\{|\tilde{\Psi}_n(\vec{R})\rangle\}$ for every set of nuclear coordinates, so that

$$\Phi(\vec{x}, \vec{R}, t) = \sum_n \left\langle \vec{x} \middle| \tilde{\Psi}_n(\vec{R}) \right\rangle \tilde{c}_n(\vec{R}, t) \tag{2.9}$$

The eigenstates of the Born-Oppenheimer approximation form such a complete basisset. Thus we can replace $|\tilde{\Psi}_n(\vec{R})\rangle$ by the Born-Oppenheimer wave functions. The corresponding coefficients are the nuclear wave functions

$$\Phi(\vec{x}, \vec{R}, t) = \sum_n \left\langle \vec{x} \middle| \Psi_n^{BO}(\vec{R}) \right\rangle \phi_n(\vec{R}, t) \tag{2.10}$$

which is nothing but the Born-Huang expansion Eq. 2.6.

Nuclear Schrödinger equation

When the Born-Huang expansion, Eq. 2.6, is inserted into the many-particle Schrödinger equation Eq. 1.1 with the full Hamiltonian Eq. 1.4, we obtain

$$\begin{aligned}
i\hbar\partial_t\left[\sum_n\left|\Psi_n^{BO}(\vec{R})\right\rangle\phi_n(\vec{R},t)\right] &\overset{Eq.\ 1.1}{=} \hat{H}\left[\sum_n\left|\Psi_n^{BO}(\vec{R})\right\rangle\phi_n(\vec{R},t)\right]\\
&\overset{Eq.\ 2.3}{=} \left[\frac{1}{2}\left(\frac{\hbar}{i}\vec{\nabla}_R\right)\boldsymbol{M}^{-1}\left(\frac{\hbar}{i}\vec{\nabla}_R\right)+\hat{H}^{BO}\right]\left[\sum_n\left|\Psi_n^{BO}(\vec{R})\right\rangle\phi_n(\vec{R},t)\right]\\
&\overset{Eq.\ 2.4}{=} \sum_n\left|\Psi_n^{BO}(\vec{R})\right\rangle\left[\frac{1}{2}\left(\frac{\hbar}{i}\vec{\nabla}_R\right)\boldsymbol{M}^{-1}\left(\frac{\hbar}{i}\vec{\nabla}_R\right)+E_n^{BO}(\vec{R})\right]\phi_n(\vec{R},t)\\
&+\quad\sum_n\left\{\left[\left(\frac{\hbar}{i}\vec{\nabla}_R\right)\left|\Psi_n^{BO}(\vec{R})\right\rangle\right]\boldsymbol{M}^{-1}\left(\frac{\hbar}{i}\vec{\nabla}_R\right)+\frac{1}{2}\left(\frac{\hbar}{i}\vec{\nabla}_R\right)\boldsymbol{M}^{-1}\left(\frac{\hbar}{i}\vec{\nabla}_R\right)\left|\Psi_n^{BO}(\vec{R})\right\rangle\right\}\phi_n(\vec{R},t)
\end{aligned}\tag{2.11}$$

Now we multiply Eq. 2.11 from the left with the bra of the Born-Oppenheimer wave function $\left\langle\Psi_m^{BO}(\vec{R})\right|$ and integrate over the electronic coordinates. Hereby, we exploit the orthonormality Eq. 2.5 of the Born-Oppenheimer wave functions.

As the result, we obtain a Schrödinger equation for the nuclear motion alone.

$$\begin{aligned}
i\hbar\partial_t\phi_m(\vec{R},t) &\overset{Eqs.\ 2.11,2.5}{=} \left[\frac{\hbar}{i}\vec{\nabla}_R\boldsymbol{M}^{-1}\frac{\hbar}{i}\vec{\nabla}_R+E_m^{BO}(\vec{R})\right]\phi_m(\vec{R},t)\\
&+\sum_n\underbrace{\left[\langle\Psi_m^{BO}|\frac{\hbar}{i}\vec{\nabla}_R|\Psi_n^{BO}\rangle\boldsymbol{M}^{-1}\frac{\hbar}{i}\vec{\nabla}_R+\frac{1}{2}\langle\Psi_m^{BO}|\frac{\hbar}{i}\vec{\nabla}_R\boldsymbol{M}^{-1}\frac{\hbar}{i}\vec{\nabla}_R|\Psi_n^{BO}\rangle\right]}_{\text{derivative couplings}}\phi_n(\vec{R},t)
\end{aligned}\tag{2.12}$$

Note, that the brackets $\langle\ldots\rangle$ are evaluated by integrating over the electronic degrees of freedom only. Thus, the matrix elements still depend explicitly on the nuclear coordinates. Note also, that the gradients $\vec{\nabla}_R$ in the parentheses act on the nuclear and not the electronic wave functions, which is easily overlooked. The gradients $\vec{\nabla}_R$ inside the matrix elements only act on the ket $|\Psi_n^{BO}\rangle$, but not on the functions outside the matrix element.

Up to now, we did not introduce any approximations to arrive at the nuclear Schrödinger equation Eq. 2.12. We are still on solid ground.

Derivative couplings

In the following, we will discuss the individual terms of Eq. 2.12. Before we continue, however, let us simplify the notation again: The general structure of Eq. 2.12 is

$$i\hbar\partial_t\phi_m(\vec{R},t)=\left[\frac{1}{2}\vec{P}\boldsymbol{M}^{-1}\vec{P}+E_m^{BO}(\vec{R})\right]\phi_m(\vec{R},t)+\sum_n\left[\vec{A}_{m,n}(\vec{R})\boldsymbol{M}^{-1}\hat{\vec{P}}+B_{m,n}(\vec{R})\right]\phi_n(\vec{R},t)\tag{2.13}$$

where $\vec{P}=\frac{\hbar}{i}\vec{\nabla}_R$ is the momentum vector of the nuclear coordinates and the first- and second-**derivative couplings** are defined by

$$\vec{A}_{m,n}(\vec{R}):=\left\langle\Psi_m^{BO}(\vec{R})\right|\frac{\hbar}{i}\vec{\nabla}_R\left|\Psi_n^{BO}(\vec{R})\right\rangle\tag{2.14}$$

$$B_{m,n}(\vec{R}):=\frac{1}{2}\left\langle\Psi_m^{BO}(\vec{R})\right|\left(\frac{\hbar}{i}\vec{\nabla}_R\right)\boldsymbol{M}^{-1}\left(\frac{\hbar}{i}\vec{\nabla}_R\right)\left|\Psi_n^{BO}(\vec{R})\right\rangle\tag{2.15}$$

Second-order couplings from first-order derivative couplings: The first and the second-derivative couplings are not independent of each other: The second-derivative coupling $B_{m,n}(\vec{R})$ can be expressed by the first-derivative couplings $\vec{A}_{m,n}(\vec{R})$. This is shown as follows (see [9]):

$$
\begin{aligned}
\frac{\hbar}{i}\vec{\nabla}_R\boldsymbol{M}^{-1}\vec{A}_{m,n}(\vec{R}) &\overset{\text{Eq. 2.14}}{=} \frac{\hbar}{i}\vec{\nabla}_R\boldsymbol{M}^{-1}\Big\langle\Psi^{BO}_m(\vec{R})\Big|\frac{\hbar}{i}\vec{\nabla}_R\Big|\Psi^{BO}_n(\vec{R})\Big\rangle\\
&= -\Big\langle\frac{\hbar}{i}\vec{\nabla}_R\Psi^{BO}_m(\vec{R})\Big|\boldsymbol{M}^{-1}\frac{\hbar}{i}\vec{\nabla}_R\Big|\Psi^{BO}_n(\vec{R})\Big\rangle+\underbrace{\Big\langle\Psi^{BO}_m(\vec{R})\Big|\Big(\frac{\hbar}{i}\vec{\nabla}_R\Big)\boldsymbol{M}^{-1}\Big(\frac{\hbar}{i}\vec{\nabla}_R\Big)\Big|\Psi^{BO}_n(\vec{R})\Big\rangle}_{2B_{m,n}(\vec{R}}\\
&\overset{\text{Eq. 2.15}}{=} -\Big\langle\frac{\hbar}{i}\vec{\nabla}_R\Psi^{BO}_m(\vec{R})\Big|\underbrace{\Big(\sum_k\Big|\Psi^{BO}_k(\vec{R})\Big\rangle\Big\langle\Psi^{BO}_k(\vec{R})\Big|\Big)}_{=\hat{1}}\boldsymbol{M}^{-1}\frac{\hbar}{i}\vec{\nabla}_R\Big|\Psi^{BO}_n(\vec{R})\Big\rangle+2B_{m,n}(\vec{R})\\
&= -\sum_k\Big\langle\frac{\hbar}{i}\vec{\nabla}_R\Psi^{BO}_m(\vec{R})\Big|\Psi^{BO}_k(\vec{R})\Big\rangle\boldsymbol{M}^{-1}\Big\langle\Psi^{BO}_k(\vec{R})\Big|\frac{\hbar}{i}\vec{\nabla}_R\Big|\Psi^{BO}_n(\vec{R})\Big\rangle+2B_{m,n}(\vec{R})\\
&= -\sum_k\Big\langle\Psi^{BO}_k(\vec{R})\Big|\frac{\hbar}{i}\vec{\nabla}_R\Big|\Psi^{BO}_m(\vec{R})\Big\rangle^*\boldsymbol{M}^{-1}\Big\langle\Psi^{BO}_k(\vec{R})\Big|\frac{\hbar}{i}\vec{\nabla}_R\Big|\Psi^{BO}_n(\vec{R})\Big\rangle+2B_{m,n}(\vec{R})\\
&= -\sum_k\vec{A}^*_{k,m}(\vec{R})\boldsymbol{M}^{-1}\vec{A}_{k,n}(\vec{R})+2B_{m,n}(\vec{R})
\end{aligned}
$$

Next, we exploit that the first-derivative couplings $\vec{A}_{m,n}$ are hermitian in the band indices, i.e. $\vec{A}^*_{k,m}(\vec{R})=\vec{A}_{m,k}(\vec{R})$, which can be shown as follows: We start from the orthonormality condition Eq. 2.5 of the Born-Oppenheimer wave functions, which holds for all $\vec{R}$. Thus the nucler gradient of the scalar products vanishes

$$
\begin{aligned}
0 &\overset{\text{Eq. 2.5}}{=} \vec{\nabla}_R\overbrace{\langle\Psi^{BO}_m|\Psi^{BO}_n\rangle}^{\delta_{m,n}}\\
&= \langle\vec{\nabla}_R\Psi^{BO}_m|\Psi^{BO}_n\rangle+\langle\Psi^{BO}_m|\vec{\nabla}_R|\Psi^{BO}_n\rangle\\
&= \langle\Psi^{BO}_n|\vec{\nabla}_R\Psi^{BO}_m\rangle^*+\langle\Psi^{BO}_m|\vec{\nabla}_R|\Psi^{BO}_n\rangle\\
\Rightarrow\quad 0 &= -\langle\Psi^{BO}_n|\frac{\hbar}{i}\vec{\nabla}_R\Psi^{BO}_m\rangle^*+\langle\Psi^{BO}_m|\frac{\hbar}{i}\vec{\nabla}_R|\Psi^{BO}_n\rangle\\
\Rightarrow\quad \vec{A}^*_{k,m}(\vec{R}) &= \Big\langle\Psi^{BO}_k\Big|\frac{\hbar}{i}\vec{\nabla}_R\Big|\Psi^{BO}_m\Big\rangle^*=\Big\langle\Psi^{BO}_m\Big|\frac{\hbar}{i}\vec{\nabla}_R\Big|\Psi^{BO}_k\Big\rangle=\vec{A}_{m,k}(\vec{R})
\end{aligned}
\tag{2.16}
$$

This allows us to rewrite Eq. 2.16 in the form

$$
2B_{m,n}(\vec{R})=\frac{\hbar}{i}\vec{\nabla}_R\boldsymbol{M}^{-1}\vec{A}_{m,n}(\vec{R})+\sum_k\vec{A}_{m,k}(\vec{R})\boldsymbol{M}^{-1}\vec{A}_{k,n}(\vec{R})
\tag{2.17}
$$

Thus, we only need to evaluate the first-derivative couplings, while the second-derivative couplings can be obtained from the former via Eq. 2.17.

Final form for the nuclear Schrödinger equation

Insertion of the second derivative couplings from Eq. 2.17 into the nuclear Schrödinger equation Eq. 2.13 into the final form of Eq. 2.20.

Let me sketch here the line of thought schematically. I am dropping the matrix structure of the equation for the sake of clarity. The terms in Eq. 2.13 related to momenta and derivative couplings have the form

$$
\begin{aligned}
\Big(\frac{1}{2}\vec{P}\boldsymbol{M}^{-1}\vec{P}+\vec{A}\boldsymbol{M}^{-1}\vec{P}+B\Big)|\phi\rangle \overset{\text{Eq. 2.17}}{=} &\ \frac{1}{2}\Big(\vec{P}\boldsymbol{M}^{-1}\vec{P}+2\vec{A}\boldsymbol{M}^{-1}\vec{P}+(\vec{P}\boldsymbol{M}^{-1}\vec{A})+\vec{A}\boldsymbol{M}^{-1}\vec{A}\Big)|\phi\rangle \\
= &\ \frac{1}{2}\Big(\vec{P}\boldsymbol{M}^{-1}\vec{P}+\vec{A}\boldsymbol{M}^{-1}\vec{P}+\vec{P}\boldsymbol{M}^{-1}\vec{A}+\vec{A}\boldsymbol{M}^{-1}\vec{A}\Big)|\phi\rangle \\
= &\ \frac{1}{2}\Big(\vec{P}+\vec{A}\Big)\boldsymbol{M}^{-1}\Big(\vec{P}+\vec{A}\Big)|\phi\rangle \qquad (2.18)
\end{aligned}
$$

We used the the chain rule for $\vec{P}$, which is a gradient operator.

$$
\vec{P}\boldsymbol{M}^{-1}\vec{A}|\phi\rangle=(\vec{P}\boldsymbol{M}^{-1}\vec{A})|\phi\rangle+\vec{A}\boldsymbol{M}^{-1}\vec{P}|\phi\rangle \qquad (2.19)
$$

In this manner one arrives at our final form Eq. 2.20 for the nuclear Schrödinger equation.

SCHRÖDINGER EQUATION FOR THE NUCLEAR WAVE FUNCTIONS IN TERMS OF BORN-OPPENHEIMER WAVE FUNCTIONS

$$
\begin{aligned}
i\hbar\partial_t\phi_m(\vec{R},t)=\sum_n\Bigg[\frac{1}{2}\sum_k\Big(\delta_{m,k}\frac{\hbar}{i}\vec{\nabla}_R+\vec{A}_{m,k}(\vec{R})\Big)\boldsymbol{M}^{-1}\Big(\delta_{k,n}\frac{\hbar}{i}\vec{\nabla}_R+\vec{A}_{k,n}(\vec{R})\Big) \\
+\delta_{m,n}E^{BO}_m(\vec{R})\Bigg]\phi_n(\vec{R},t) \qquad (2.20)
\end{aligned}
$$

or, combining the components of the nuclear wave functions in a vector-matrix notation – using a single underline to denote a vector and a double underline for a matrix –

$$
i\hbar\partial_t\underline{\phi}(\vec{R},t)=\left[\frac{\Big(\underline{\underline{1}}\hat{\vec{P}}+\underline{\underline{\vec{A}}}(\vec{R})\Big)^2}{2\boldsymbol{M}}+\underline{\underline{E}}^{BO}(\vec{R})\right]\underline{\phi}(\vec{R},t) \qquad (2.21)
$$

$\underline{\underline{E}}^{BO}$ denotes the diagonal matrix with the Born-Oppenheimer energies as diagonal elements and $\hat{\vec{P}}=\frac{\hbar}{i}\vec{\nabla}_R$ is the $3M$-dimensional momentum operator of the nuclei. The kinetic energy expression in Eq. 2.21 is mathematically not allowed, because I wrote a vector-matrix-vector expression as vector-squared-divided-by-matrix. I used it here to allude to the corresponding equation from electrodynamics so that it can be memorized more easily, and because the matrix $\boldsymbol{M}$ is a diagonal matrix so that the expression can be evaluated component by component.
The first-derivative couplings $\vec{A}_{m,k}(\vec{R})$ are defined in Eq. 2.14.

$$
\vec{A}_{m,n}(\vec{R})\overset{\text{Eq. 2.14}}{:=}\Big\langle\Psi^{BO}_m(\vec{R})\Big|\frac{\hbar}{i}\vec{\nabla}_R\Big|\Psi^{BO}_n(\vec{R})\Big\rangle=\Big\langle\Psi^{BO}_m(\vec{R})\Big|\hat{\vec{P}}\Big|\Psi^{BO}_n(\vec{R})\Big\rangle \qquad (2.22)
$$

As shown in Eq. 2.16, each vector component of the first-derivative couplings is a hermitian matrix in the indices m, n.

This form of the nuclear Schrödinger equation, i.e. Eq. 2.21, has been mentioned[2] already in 1969 by F. Smith[10].

We have not introduced any approximations to bring Eq. 2.12 into the new form given by Eq. 2.21. That is, we are still on solid grounds.

---

[2] see his Eq. 10. His $\underline{\vec{P}}(R)$ are the first-derivative couplings.

Relation to electrodynamics

Interesting is the formal similarity of the nuclear Hamilton function in Eq. 2.14 with that of a charged particle in a electromagnetic field, which has the form

$$H(\vec{p},\vec{r})=\frac{1}{2m}\Big(\vec{p}-q\vec{A}(\vec{r})\Big)^2+q\Phi(\vec{r})$$

This analogy shows that the diagonal terms (in the band indices) of the derivative couplings act like a magnetic field $\vec{B}$ expressed by the vector potential $\vec{A}$ as $\vec{B}=\vec{\nabla}\times\vec{A}$ and an electric potential Φ. We will discuss this analogy later in somewhat more detail in the appendix D.2.

2.2 Approximations

2.2.1 Born-Oppenheimer approximation

The Born-Oppenheimer approximation amounts to ignoring the derivative couplings in Eq. 2.20

BORN-OPPENHEIMER APPROXIMATION FOR THE NUCLEAR WAVE FUNCTION

$$i\hbar\partial_t\phi_n(\vec{R},t)=\left[\sum_{j=1}^{M}\frac{-\hbar^2}{2M_j}\vec{\nabla}^2_{R_j}+E^{BO}_n(\vec{R})\right]\phi_n(\vec{R},t) \tag{2.23}$$

This equation describes nuclei that move on a given total energy surface $E^{BO}_n(\vec{R})$, which is called the **Born-Oppenheimer surface**. The Born-Oppenheimer surface may be an excited-state surface or the ground-state surface. The wave function may also have contributions simultaneously on different total energy surfaces. However, within the Born-Oppenheimer approximation, the contributions on different surfaces do not influence each other.

In other words, the system is with probability $P_n=\int d^{3M}R\ \phi^*_n(\vec{R})\phi_n(\vec{R})$ on the excited state surface $E^{BO}_n(\vec{R})$. In the Born-Oppenheimer approximation, this probability does not change with time.

The neglect of the non-adiabatic effects is the essence of the **Born-Oppenheimer approximation**. In the absence of the non-adiabatic effects, we could start the system in a particular eigenstate of the Born-Oppenheimer Hamiltonian, and the system would always evolve on the same total energy surface $E^{BO}_n(\vec{R})$. Thus, if we start the system in the electronic ground state, it will remain exactly in the instantaneous electronic ground state, while the nuclei are moving. Band crossings are the exceptions: Here, the Born-Oppenheimer approximation does not give a unique answer.

The separation of nuclear and electronic degrees of freedom have already been in use before the original Born-Oppenheimer approximation[11]. Born and Oppenheimer [7] gave a justification for neglecting the non-adiabatic effects in many cases. As we will see the Born-Oppenheimer approximation breaks down when Born-Oppenheimer surfaces come very close or even cross.

2.2.2 Classical approximation

Classical approximation in the Born-Oppenheimer approximation

The most simple approach to the classical approximation is to take the Hamilton operator in the Born-Oppenheimer approximation from Eq. 2.23, constrain it to a particular total-energy surface specified by n, and to form the corresponding Hamilton function.

$$H_n(\vec{P},\vec{R})=\sum_j\frac{\vec{P}_j^2}{2M_j}+E^{BO}_n(\vec{R})$$

The Hamilton function defines the classical Hamilton equations of motion

$$\partial_t \vec{R}_j = \vec{\nabla}_{P_j} H_n(\vec{P}, \vec{R}, t) = \frac{1}{M_j}\vec{P}_j$$
$$\partial_t \vec{P}_j = -\vec{\nabla}_{R_j} H_n(\vec{P}, \vec{R}, t) = -\vec{\nabla}_{R_j} E_n^{BO}(\vec{R})$$

This in turn leads to the Newton's equations of motion for the nuclei

$$M_j \partial_t^2 \vec{R}_j = -\vec{\nabla}_{R_j} E_n^{BO}(\vec{R})$$

This is the approximation that is most widely used to study the dynamics of the atoms in a molecule or a solid. A simulation of classical atoms using some kind of parameterized Born-Oppenheimer surface is called **molecular dynamics simulation**. When the Born-Oppenheimer surface is obtained on the fly from a quantum mechanical electronic structure method, the method is called **ab-initio molecular dynamics**.

The Born-Oppenheimer surface $E_n^{BO}(\vec{R})$ acts just like a total-energy surface for the motion of the nuclei. Once $E_n^{BO}(\vec{R})$ is known, the electrons are taken completely out of the picture. Within the Born-Oppenheimer approximation, no transitions between the ground-state and the excited-state surface take place. This implies not only that a system in the electronic ground state remains in the ground state. It also implies that a system in the excited state will remain there for ever. Transitions between different sheets $E_n^{BO}(\vec{R})$ of the Born-Oppenheimer energy are only possible when non-adiabatic effects are included.

2.3 Non-adiabatic corrections

While the Born-Oppenheimer approximation is a widely used and tremendously successful approach, its limitations are important for relaxation processes, photochemistry, charge transfer processes and many more interesting processes. They shall be briefly discussed here.

We will see that the derivative couplings governing the non-adiabatic effects are singular where two Born-Oppenheimer surface meet. Unfortunately, these are exactly the points of physical interest, namely those where a system makes a non-radiative transition from one Born-Oppenheimer surface to another. Even worse, these singular points have non-local consequences, so-called topological effects.

Interesting is a close formal similarity between non-adiabatic effects and field theory[12]. Related topics are the **Aharonov-Bohm effect**[13] and **Yang-Mills fields**[14].

Under normal circumstances, the diagonal terms of the derivative couplings are of secondary importance. The diagonal term of the derivative couplings are discussed in the appendix D.2. They become, however, important in the presence of conical intersections and in the presence of a geometrical phase.

2.3.1 Non-diagonal derivative couplings

The off-diagonal elements of the first-derivative couplings defined in Eq. 2.22 can be brought into the form

$$\vec{A}_{m,n} = \frac{\left\langle \Psi_m^{BO} \left| \left(\frac{\hbar}{i}\vec{\nabla}_R \hat{H}^{BO}(\vec{R})\right) \right| \Psi_n^{BO} \right\rangle}{E_n^{BO}(\vec{R}) - E_m^{BO}(\vec{R})} \qquad \text{for } E_m^{BO} \neq E_n^{BO} \tag{2.24}$$

The gradient acts only on the Born-Oppenheimer Hamiltonian but not further to the state on the right. This expression makes it evident that non-adiabatic effects are important when two Born-Oppenheimer surfaces become degenerate, i.e. when $E_m^{BO} = E_n^{BO}$.

Derivation of Eq. D.2 The off-diagonal elements of $\vec{A}_{n,m}(\vec{R})$ of the first-derivative couplings are obtained as follows: We begin with the Schrödinger equation for the Born-Oppenheimer wave function

$$\Big(\hat{H}^{BO}(\vec{R})-E^{BO}_n(\vec{R})\Big)|\Psi^{BO}_n(\vec{R})\rangle \overset{\text{Eq. 2.4}}{=} 0$$

Because this equation is valid for all atomic positions, the gradient of the above expression vanishes.

$$\begin{aligned}0 &= \vec{\nabla}_R\Big(\hat{H}^{BO}(\vec{R})-E^{BO}_n(\vec{R})\Big)\Big|\Psi^{BO}_n(\vec{R})\Big\rangle\\ &= \Big(\vec{\nabla}_R,\hat{H}^{BO}(\vec{R})\Big)\Big|\Psi^{BO}_n(\vec{R})\Big\rangle-\Big|\Psi^{BO}_n(\vec{R})\Big\rangle\vec{\nabla}_R E^{BO}_n(\vec{R})+\Big(\hat{H}^{BO}(\vec{R})-E^{BO}_n(\vec{R})\Big)\vec{\nabla}_R\Big|\Psi^{BO}_n(\vec{R})\Big\rangle\end{aligned}$$

Now, we multiply from the left with the bra $\langle\Psi^{BO}_m|$ and form the scalar products in the electronic Hilbert space.[3]

$$\begin{aligned}0 &= \langle\Psi^{BO}_m|\Big(\vec{\nabla}_R\hat{H}^{BO}(\vec{R})\Big)|\Psi^{BO}_n\rangle-\langle\Psi^{BO}_m\,|\Psi^{BO}_n\rangle\vec{\nabla}_R E^{BO}_n(\vec{R})+\langle\Psi^{BO}_m|\Big(\hat{H}^{BO}(\vec{R})-E^{BO}_n(\vec{R})\Big)\vec{\nabla}_R|\Psi^{BO}_n\rangle\\ &= \langle\Psi^{BO}_m|\Big(\vec{\nabla}_R,\hat{H}^{BO}(\vec{R})\Big)|\Psi^{BO}_n\rangle-\underbrace{\delta_{m,n}\vec{\nabla}_R E^{BO}_n(\vec{R})}_{=0\text{for } m\neq n}+\Big(\hat{E}^{BO}_m(\vec{R})-E^{BO}_n(\vec{R})\Big)\langle\Psi^{BO}_m|\vec{\nabla}_R|\Psi^{BO}_n\rangle\end{aligned}$$

If the indices m, n differ, the middle term containing gradient of the Born-Oppenheimer surfaces vanishes. If the energies E^{BO}_m, E^{BO}_n furthermore differ, we can divide by their difference which brings us to the desired result.

$$\vec{A}_{m,n} \overset{\text{Eq. 2.14}}{=} \langle\Psi^{BO}_m|\frac{\hbar}{i}\vec{\nabla}_R|\Psi^{BO}_n\rangle=\frac{\langle\Psi^{BO}_m|\Big(\frac{\hbar}{i}\vec{\nabla}_R\hat{H}^{BO}(\vec{R})\Big)|\Psi^{BO}_n\rangle}{\hat{E}^{BO}_n(\vec{R})-E^{BO}_m(\vec{R})} \qquad \text{for } E_n\neq E_m \tag{2.25}$$

This is the proof for Eq. 2.24 given above. Unfortunately, no information can be extracted for the diagonal elements of the first-derivative couplings.

2.3.2 Topology of crossings of Born-Oppenheimer surfaces

The off-diagonal non-adiabatic effects Eq. 2.24 become important, when the different Born-Oppenheimer surfaces come close or even cross. Therefore we need to understand the topology of crossings of Born-Oppenheimer surfaces.

Non-crossing rule and conical intersections

The most important information on the topology of surface crossings is provided by the **non-crossing rule** of von Neumann and Wigner[15]. It is derived in appendix E on p. 255.

NON-CROSSING RULE

Consider a Hamiltonian $\hat{H}(Q_1,\ldots,Q_m)$ which depends on m parameters $Q_1,\ldots,Q_m$. For each accidental crossing of two eigenvalues, there is a three-dimensional subspace of the parameter space in which the degeneracy is lifted except for a single point.
If the hamiltonian is real for all parameters, the subspace where the degeneracy is lifted has two dimensions.

- In our case, the Hamiltonian is the Born-Oppenheimer Hamiltonian and the parameters are the nuclear positions.

---

[3] The Born-Oppenheimer wave functions depend on electronic and nuclear coordinates. The nuclear coordinates, however, play a different role, because they are treated like external parameters.

- Two energy surfaces that meet in a point within a two-dimensional space form two cones. Therefore the crossing is called a **conical intersection**.
- The hamiltonian is usually[4] real in the absence of magnetic fields and relativistic effects. Spin-orbit coupling, a relativistic effect, and magnetic fields that act on the orbital motion of charged particles introduce complex contributions. The question whether the Hamiltonian is real is related to time inversion symmetry discussed in Appendix L on p. L.
- The limitation of the non-crossing rule to accidental crossings implies that there are no limitations on crossings that occur because two states have different symmetry.
- The non-crossing rule implies that there are no accidental crossings for dimeric molecules

Let me consider a conical intersection related to the motion of a single atom in a molecule or solid. Let me furthermore consider real Hamiltonians. For this example, the conical intersection would form a one-dimensional line. The two displacements perpendicular to this line lift the degeneracy. If we plot the energy surfaces in the plane of these two distortions, the Born-Oppenheimer surfaces would have the shape of a double cone, a **conical intersection**. A change of a coordinate along the line will change the shape of the cone, but it will not affect the qualitative cone-like topology.

More generally, there is a two-dimensional plane of coordinates relevant to the conical intersection. Distortions of distan nuclei will not affect the general shape of the conical intersection, while they do have an effect on the total energy.

2.3.3 Avoided crossings

Often, the symmetries that are responsible for a degeneracy, which is a surface crossing, are only approximately present. A small matrix element lifts the degeneracy "a little". This is called an avoided **avoided crossing**. Due to an avoided crossing two Born-Oppenheimer sheets may be come so close that the non-adiabatic effects can no more be ignored. Thus, regions with avoided crossings are important for the relaxation of the system onto a Born-Oppenheimer surface with lower energy.

What is an avoided crossing? Let me start with a position-dependent Hamiltonian for a one-dimensional nuclear coordinate R, for which the energy levels cross for a certain position $R_0 = 0$. In addition we add non-diagonal terms Δ.

The Hamilton has the form

$$\boldsymbol{H} = \frac{\hat{P}^2}{2M} + \underbrace{\begin{pmatrix} |a\rangle \\ |b\rangle \end{pmatrix} \begin{pmatrix} E_0 + F\hat{Q} & -\Delta \\ -\Delta & E_0 - F\hat{Q} \end{pmatrix} \begin{pmatrix} \langle a| \\ \langle b| \end{pmatrix}}_{\hat{H}_{BO}} \tag{2.26}$$

The Born-Oppenheimer Hamiltonian is obtained by stripping away the nuclear kinetic energy and by converting the coordinate operator $\hat{Q}$ into a parameter.

$$\boldsymbol{H}^{BO}(Q) = \begin{pmatrix} |a\rangle \\ |b\rangle \end{pmatrix} \begin{pmatrix} E_0 + FQ & -\Delta \\ -\Delta & E_0 - FQ \end{pmatrix} \begin{pmatrix} \langle a| \\ \langle b| \end{pmatrix} \tag{2.27}$$

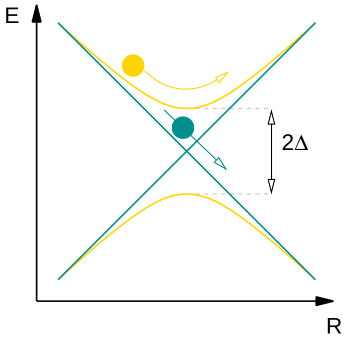
Diagonalization of the Born-Oppenheimer Hamiltonian yields the two Born-Oppenheimer surfaces

$$E^{BO}_{\pm}(Q) = E_0 \pm \Delta \sqrt{1 + \left(\frac{F}{\Delta} Q\right)^2} \tag{2.28}$$

where the minus sign applies for the lower and the plus sign for the upper Born-Oppenheimer surface.

---

[4] There are other interactions that are similar to a magnetic field. They are, however, not very important for solid state physics.



The eigenstates of the Born-Oppenheimer Hamiltonian, the Born-Oppenheimer states are

$$|\Phi^{BO}_{+}(Q)\rangle = \begin{pmatrix} |a\rangle \\ |b\rangle \end{pmatrix} \begin{pmatrix} -\sin(\gamma(Q)) \\ \cos(\gamma(Q)) \end{pmatrix} \qquad \text{and} \qquad |\Phi^{BO}_{-}(Q)\rangle = \begin{pmatrix} |a\rangle \\ |b\rangle \end{pmatrix} \begin{pmatrix} \cos(\gamma(Q)) \\ \sin(\gamma(Q)) \end{pmatrix} \tag{2.29}$$

where

$$\gamma(Q) = \arctan\left(\frac{F}{\Delta}Q + \sqrt{1+\left(\frac{F}{\Delta}Q\right)^2}\right) = \frac{\pi}{4} + \frac{1}{2}\arctan\left(\frac{F}{\Delta}Q\right) \tag{2.30}$$

is shown in Fig. 2.1.

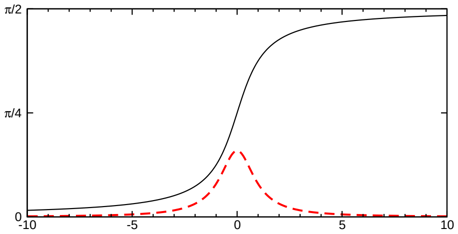


Fig. 2.1: $\gamma(Q)$ determining the eigenstates and derivative couplings for the avoided crossing. The coordinate is drawn in units of $\frac{\Delta}{F}$ The dashed line is the derivative of $\phi(Q)$ in units of $\frac{F}{\Delta}$.

The non-diagonal derivative couplings are

$$A_{-,+} = -A_{+,-} = \frac{\hbar}{i}\frac{d\gamma(Q)}{dQ} \tag{2.31}$$

and the diagonal derivative couplings vanish. From the red dashed line in Fig. 2.1, it is evident that the derivative couplings are large near the avoided crossing. The reason for this is that the wave function changes its character from one side of the derivative coupling to the other.

2.3.4 Conical intersections: Jahn-Teller model

Because I find the conical intersections a little mind-boggling, let me switch from the general and abstract to the more concrete case. That is will explain the conical intersection for a simple model, the **Jahn-Teller model**. The Jahn Teller model is the main model system for a **conical intersection**.

For a review on conical intersection see for example the paper of Yarkony [16]

The Jahn-Teller model can be used to describe the Jahn-Teller effect in doped manganites. Manganites such as $Ca_xLa_{1-x}MnO_3$ form a very interesting class of materials with a complex phase diagram and enormous variety of different physical properties that can be controlled by doping,

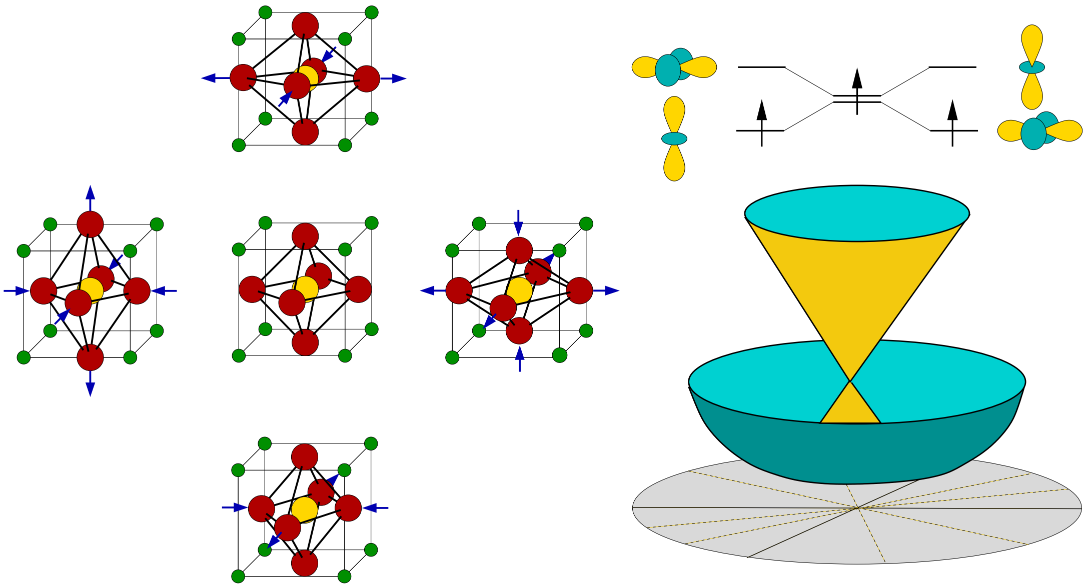


Fig. 2.2: Left: the two Jahn-Teller active modes of the octahedron in a perovskite ABO_3. The red spheres represent oxygen ions, the yellow atom in the center is the transition metal ion (B-type ion) on which the two e_g orbitals reside. The green spheres are the A-type ions. The undistorted octahedron is drawn in the center. Along the horizontal axis the positive and negative distotions of the $\mathcal{Q}_2$ mode are drawn. Along the vertical axis the positive and negative distotions of the $\mathcal{Q}_3$ mode are drawn. Right top: Jahn-Teller splitting of the two e_g orbitals under the influence of a positive and negative distortion of the $\mathcal{Q}_3$ mode. Left bottom: The two Born-Oppenheimer surfaces of the Jahn-Teller model in the two-dimensional coordinate space of the two Jahn-Teller active modes. The point of contact of the two Born-Oppenheimer surfaces is a conical intersection.

temperature, magnetic fields, etc. Their main structural element is a network of corner-sharing MnO_6 octrahedra. If Mn is in the 4+ charge state (Mn^{IV}) as in $CaMnO_3$, the Mn ions in the center of the octahedra exhibit two degenerate d-orbitals of e_g symmetry, which are unoccupied.

In the doped materials, electrons can be added to these pair of degenerate states. This electron in a given orbital induces a structural distortion of the octahedra, which lowers the filled orbital and raises the empty orbital. The distortions are the elongation and compression of the octahedra along one axis into a prolate (zigar-shaped) and an oblate (pancake-shaped) form.

One can say that the electron and the octahedral distortion (a phonon) form a bound state. This bound state is called a **polaron**.

The Jahn-Teller distortions can be expressed in terms of the outward displacements a_x, a_y and a_z of the oxygen ligands in x,y, and z directions, respectively. These can be combined into a vector (a_x, a_y, a_z). The two Jahn-Teller active displacements are[17] $\mathcal{Q}_2 = \frac{1}{\sqrt{2}}(1, -1, 0)$ and $\mathcal{Q}_3 = \frac{1}{\sqrt{6}}(-1, -1, 2)$. The positive and negative distortions $\mathcal{Q}_3$ can be described by a **prolate** (cigar-shaped) and an **oblate** (pancake-shaped) distortion. The Jahn-Teller distortions span a two dimensional parameter space. They can be described by an amplitude and an angle parameter α specifying the type of the distortion pattern.

In the following we discuss the conical intersection at a single octahedron. The relevant structural distortions of the octahedra are described by the generalized nuclear coordinates x, y, and the electronic wave function is described as a vector in a two-dimensional Hilbert space of the two e_g orbitals. I label the two e_g orbitals as $|a\rangle$ and $|b\rangle$.

The atomic coordinates, that is the generalized coordinates for the two octahedral distortions,

α	(a_x, a_y, a_z)	$(\mathcal{Q}_2, \mathcal{Q}_3)$	
$0°$	$\frac{1}{\sqrt{2}}(1, -1, 0)$	$(1, 0)$	
$30°$	$\frac{1}{\sqrt{6}}(1, -2, 1)$	$(\frac{1}{2}\sqrt{3}, \frac{1}{2})$	oblate along y
$60°$	$\frac{1}{\sqrt{2}}(0, 1, -1)$	$(\frac{1}{2}, \frac{1}{2}\sqrt{3})$	
$90°$	$\frac{1}{\sqrt{6}}(-1, -1, 2)$	$(0, 1)$	prolate along z
$120°$	$\frac{1}{\sqrt{2}}(-1, 0, 1)$	$(-\frac{1}{2}, \frac{1}{2}\sqrt{3})$	
$150°$	$\frac{1}{\sqrt{6}}(-2, 1, 1)$	$(-\frac{1}{2}\sqrt{3}, \frac{1}{2})$	oblate along x
$180°$	$\frac{1}{\sqrt{2}}(-1, 1, 0)$	$(-1, 0)$	
$210°$	$\frac{1}{\sqrt{6}}(-1, 2, -1)$	$(-\frac{1}{2}\sqrt{3}, -\frac{1}{2})$	prolate along y
$240°$	$\frac{1}{\sqrt{2}}(0, 1, -1)$	$(-\frac{1}{2}, -\frac{1}{2}\sqrt{3})$	
$270°$	$\frac{1}{\sqrt{6}}(1, 1, -2)$	$(0, 1)$	oblate along z
$300°$	$\frac{1}{\sqrt{2}}(1, 0, -1)$	$(\frac{1}{2}, -\frac{1}{2}\sqrt{3})$	
$330°$	$\frac{1}{\sqrt{6}}(2, -1, -1)$	$(\frac{1}{2}\sqrt{3}, -\frac{1}{2})$	prolate along x

Table 2.1: . Jahn-Teller active distortions of the octahedron as function of the angular coordinte α. The second row (a_x, a_y, a_z) describes the bond expansions in x,y,z directions and the third describes the mode amplitudes.

which couple to the two e$_g$ states are denoted by the two-dimensional vector[5] $\vec{R} = (X, Z)$.

The time-dependent Schrödinger equation

$$i\hbar\partial_t|\Phi(t)\rangle = \left[\frac{\hat{\vec{P}}^2}{2M} + w\hat{\vec{R}}^2 + g\hat{\vec{R}}\,\hat{\vec{\sigma}}\right]|\Phi(t)\rangle \tag{2.32}$$

where g is the electron-phonon coupling parameter and w is the force constant for the restoring force. $\hat{\vec{\sigma}}$ act on the electronic degrees of freedom, where they are represented as the Pauli matrices $\boldsymbol{\sigma}_x$ and $\boldsymbol{\sigma}_z$. $\hat{\vec{R}} = (\hat{X}, \hat{Z})$ and $\hat{\vec{P}}$ are the operators describing the nuclear octahedral distortions and their momenta.

As a little warning: In this problem only electrons of one spin direction take part. Because of that reason there is no spin coordinate, and we use the spatial coordinate $\vec{r}$ of the electrons. The Pauli matrices used here are no spin operators but matrices in the two-dimensional Hilbert space of the two e_g orbitals (both having the same spin).

In components, this equation has the form

$$\begin{aligned} i\hbar\partial_t\begin{pmatrix}\Phi_a(X,Z,t)\\ \Phi_b(X,Z,t)\end{pmatrix} &= \left[\frac{-\hbar^2}{2M}\left(\partial_X^2+\partial_Z^2\right) + w\left(X^2+Z^2\right)\right]\begin{pmatrix}\Phi_a(X,Z,t)\\ \Phi_b(X,Z,t)\end{pmatrix} \\ &\quad + g\begin{pmatrix}Z, & X\\ X & -Z\end{pmatrix}\begin{pmatrix}\Phi_a(X,Z,t)\\ \Phi_b(X,Z,t)\end{pmatrix} \end{aligned} \tag{2.33}$$

Here we used the matrix elements $\Phi_a(X, Z, t) = \int d^3r\ \phi_a(\vec{r})\Phi(\vec{r}, X, Z, t)$ with the one-particle orbitals ϕ_a and ϕ_b for the two e_g orbitals. The full wave functions including nuclear and electronic wave functions is $\Phi(\vec{r}, X, Z, t) = \phi_a(\vec{r})\Phi_a(X, Z, t) + \phi_b(\vec{r})\Phi_b(X, Z, t)$.

Born-Oppenheimer Hamiltonian and Born-Oppenheimer surfaces The Born-Oppenheimer Hamiltonian is obtained by stripping the terms containing the nuclear momenta from the many-particle

---

[5] I am using uppercase symbols as these are nuclear coordinates. X and Z are used rather than X and Y, because of the two non-imaginary Pauli matrices σ_x and σ_z.

Hamiltonian.[6]

$$\hat{H}^{BO}(\vec{R}) = g\begin{pmatrix} Z & X \\ X & -Z \end{pmatrix} + w\begin{pmatrix} X^2+Z^2, & 0 \\ 0, & X^2+Z^2 \end{pmatrix} = g\vec{R}\cdot\vec{\boldsymbol{\sigma}} + w\vec{R}^2\mathbf{1} \tag{2.34}$$

The first term describes the electronic Hamiltonian, that is the splitting of the degenerate electronic states upon distortion, while the second term is a simple restoring potential for the nuclear distortions.

The characteristic equation for the eigenvalues $E^{BO}_{\pm}(\vec{R})$ is

$$\begin{pmatrix} +gZ + w\vec{R}^2 - E^{BO}_{\pm}(\vec{R}) & gX \\ gX & -gZ + w\vec{R}^2 - E^{BO}_{\pm}(\vec{R}) \end{pmatrix}\begin{pmatrix} \phi_{a,\pm} \\ \phi_{b,\pm} \end{pmatrix} = 0 \tag{2.35}$$

This yields the two Born-Oppenheimer surfaces of the Jahn-Teller model

BORN-OPPENHEIMER SURFACES OF THE JAHN-TELLER MODEL

$$E^{BO}_{\pm}(\vec{R}) = \pm g|\vec{R}| + w\vec{R}^{\,2} \tag{2.36}$$

The Born-Oppenheimer surfaces $E^{BO}_{\pm}(\vec{R})$ are shown in Fig. 2.2. The two surfaces meet at a single point, which is the **conical intersection**.

Born-Oppenheimer wave functions To construct the eigenstates we take the first line of Eq. 2.35 and construct the orthogonal vector according to the rule $(u, v) \perp (-v, u)$, and then we normalize it. This yields the Born-Oppenheimer states $|\bar{\Psi}^{BO}_{\pm}\rangle$ with components $\bar{\Psi}^{BO}_{a,\pm} = \langle a|\bar{\Psi}^{BO}_{\pm}\rangle$ and $\bar{\Psi}^{BO}_{b,\pm} = \langle b|\bar{\Psi}^{BO}_{\pm}\rangle$.

$$\begin{aligned} \begin{pmatrix} \bar{\Psi}^{BO}_{a,\pm} \\ \bar{\Psi}^{BO}_{b,\pm} \end{pmatrix} &= \begin{pmatrix} -gX \\ gZ + w\vec{R}^2 \mp g|\vec{R}| - w\vec{R}^2 \end{pmatrix} \frac{1}{\sqrt{g^2X^2 + \left(gZ + w\vec{R}^2 \mp g|\vec{R}| - w\vec{R}^2\right)^2}} \\ &= \begin{pmatrix} -X \\ Z \mp |\vec{R}| \end{pmatrix} \frac{1}{\sqrt{X^2 + \left(Z \mp |\vec{R}|\right)^2}} \end{aligned} \tag{2.37}$$

The bar ontop of the symbol for the Born-Oppenheimer state is there to distinguish it from the final result, which is multiplied with an additional phase factor.

Now we switch to polar coordinates

$$\begin{aligned} X(R,\alpha) &= R\sin(\alpha) \\ Z(R,\alpha) &= R\cos(\alpha) \end{aligned} \tag{2.38}$$

which yields

$$\begin{pmatrix} \bar{\Psi}^{BO}_{a,\pm} \\ \bar{\Psi}^{BO}_{b,\pm} \end{pmatrix} = \begin{pmatrix} -\sin(\alpha) \\ \mp(1 \mp \cos(\alpha)) \end{pmatrix} \frac{1}{\sqrt{2(1 \mp \cos(\alpha))}} \tag{2.39}$$

We use the triginometric identities

$$\begin{aligned} \left|\sin\left(\frac{x}{2}\right)\right| &= \sqrt{\frac{1}{2}\Big(1 - \cos(x)\Big)} \\ \left|\cos\left(\frac{x}{2}\right)\right| &= \sqrt{\frac{1}{2}\Big(1 + \cos(x)\Big)} \end{aligned} \tag{2.40}$$

---

[6] $\vec{R}\vec{\boldsymbol{\sigma}} = X\boldsymbol{\sigma}_x + Y\boldsymbol{\sigma}_y$.

to obtain

$$\begin{aligned}
\begin{pmatrix}\bar{\Psi}^{BO}_{a,+}\\ \bar{\Psi}^{BO}_{b,+}\end{pmatrix} &= \begin{pmatrix}-2\sin\left(\frac{\alpha}{2}\right)\cos\left(\frac{\alpha}{2}\right)\\ -2\sin^2\left(\frac{\alpha}{2}\right)\end{pmatrix}\frac{1}{2|\sin\left(\frac{\alpha}{2}\right)|} = \begin{pmatrix}-\cos\left(\frac{\alpha}{2}\right)\\ -\sin\left(\frac{\alpha}{2}\right)\end{pmatrix}\mathrm{sgn}\Big(\sin(\alpha/2)\Big)\\
\begin{pmatrix}\bar{\Psi}^{BO}_{a,-}\\ \bar{\Psi}^{BO}_{b,-}\end{pmatrix} &= \begin{pmatrix}-2\sin\left(\frac{\alpha}{2}\right)\cos\left(\frac{\alpha}{2}\right)\\ 2\cos^2\left(\frac{\alpha}{2}\right)\end{pmatrix}\frac{1}{2|\cos\left(\frac{\alpha}{2}\right)|} = \begin{pmatrix}-\sin\left(\frac{\alpha}{2}\right)\\ \cos\left(\frac{\alpha}{2}\right)\end{pmatrix}\mathrm{sgn}\Big(\cos(\alpha/2)\Big) \qquad (2.41)
\end{aligned}$$

where the sign function "sgn" is positive for positive arguments and negative for negative arguments.[7] The resulting functions are sketched in Fig. 2.3.

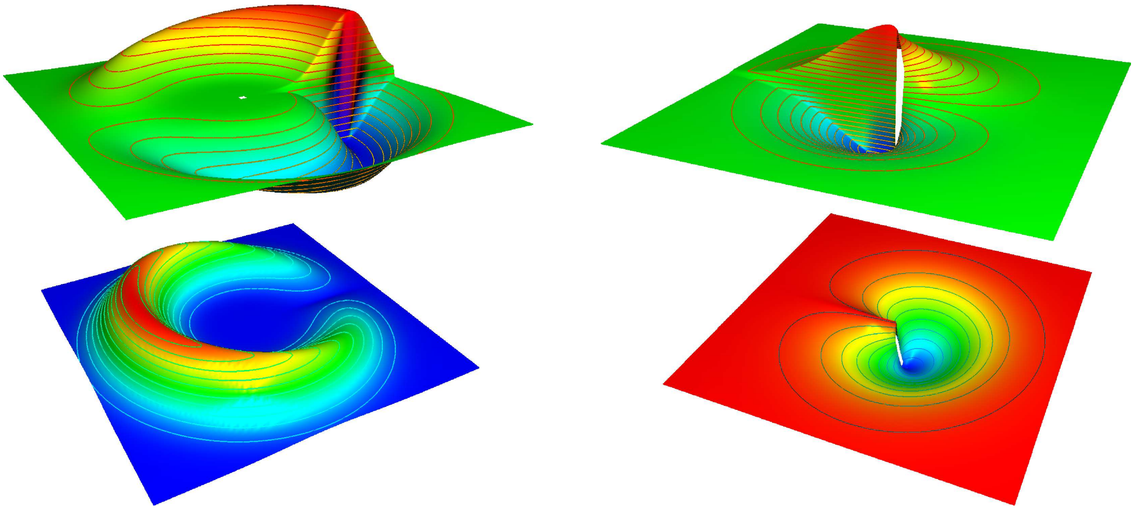


Fig. 2.3: Angular dependence of the discontinuous, real Born-Oppenheimer wave functions for the lower (left) and upper(right) Born-Oppenheimer surfaces. The top graph is the coefficient for the orbital $|a\rangle$ and the lower graph is that for orbital $|b\rangle$. The radial dependence is only schematic and is not calculated. Left and right wave functions have different orientations. The discontinuities are in opposite directions for the two Born-Oppenheimwe wave functions.

The Born-Oppenheimer wave functions are discontinuous because of the sign function. The term without the sign-term is periodic with 4π, that is, it changes sign with every full turn. The sign term, restores the original 2π periodicity of the problem. It does so, however, at the price of introducing a discontinuity.

The Born-Oppenheimer wave function are determined only up to a complex phase factor. This may lead to several alternative descriptions

- The Born-Oppenheimer wave functions are real, but discontinuous in the nuclear coordinates as in the example above.
- The Born-Oppenheimer wave functions are real and continuous, but double valued. Double valued means that they change sign with a full turn around the conical intersection. That is, they obtain the original value only after two full turns. This result would be obtained if one drops the sign term in the above equation.
- The Born-Oppenheimer wave function is continuous and single valued, but complex. This is obtained from the above result by replacing the sign term with a smooth phase factor with the same periodicity as the sign function. This representation is created below.

---

[7] For an argument zero, the sgn function vanishes. Here we will not discuss this isolated point which will not cause problems.

In the literature the first choice, namely real Born-Oppenheimer wave functions are usually assumed, explicitely or implicitely. In order to avoid discontinuities, half-integer spins are introduced. While this may have some intellectural appeal, I find the resulting double-valuedness of the wave function highly ambiguous. Therefore, I follow a different path and admit that the Born-Oppenheimer wave functions are complex-valued.

By replacing the sign function by[8] $\mp\mathrm{e}^{i\alpha/2}$, i.e.

$$\begin{aligned}|\Psi^{BO}_{+}\rangle &= -|\bar\Psi^{BO}_{+}\rangle \mathrm{sgn}\Big(\sin(\alpha/2)\Big)\mathrm{e}^{i\alpha/2}\\ |\Psi^{BO}_{-}\rangle &= +|\bar\Psi^{BO}_{-}\rangle \mathrm{sgn}\Big(\cos(\alpha/2)\Big)\mathrm{e}^{i\alpha/2}\end{aligned} \qquad (2.42)$$

we obtain[9] Born-Oppenheimer states

$$\begin{aligned}\begin{pmatrix}\Psi^{BO}_{a,+}\\ \Psi^{BO}_{b,+}\end{pmatrix} &= \begin{pmatrix}\cos(\alpha/2)\\ \sin(\alpha/2)\end{pmatrix}\mathrm{e}^{i\alpha/2} = \frac{1}{2}\begin{pmatrix}\mathrm{e}^{i\alpha}+1\\ -i(\mathrm{e}^{i\alpha}-1)\end{pmatrix}\\ \begin{pmatrix}\Psi^{BO}_{a,-}\\ \Psi^{BO}_{b,-}\end{pmatrix} &= \begin{pmatrix}-\sin(\alpha/2)\\ \cos(\alpha/2)\end{pmatrix}\mathrm{e}^{i\alpha/2} = \frac{1}{2}\begin{pmatrix}+i(\mathrm{e}^{i\alpha}-1)\\ \mathrm{e}^{i\alpha}+1\end{pmatrix}\end{aligned} \qquad (2.43)$$

BORN-OPPENHEIMER STATES OF THE JAHN-TELLER MODEL

$$\begin{pmatrix}\Psi^{BO}_{a,+}\\ \Psi^{BO}_{b,+}\end{pmatrix} = \frac{1}{2}\begin{pmatrix}\mathrm{e}^{i\alpha}+1\\ -i(\mathrm{e}^{i\alpha}-1)\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}\Psi^{BO}_{a,-}\\ \Psi^{BO}_{b,-}\end{pmatrix} = \frac{1}{2}\begin{pmatrix}+i(\mathrm{e}^{i\alpha}-1)\\ \mathrm{e}^{i\alpha}+1\end{pmatrix} \qquad (2.44)$$

The abstract states are obtained from these coefficients as

$$|\Psi^{BO}_{\pm}(\vec R)\rangle = |\vec a\rangle\Psi^{BO}_{a,\pm}(\vec R) + |\vec b\rangle\Psi^{BO}_{b,\pm}(\vec R) \qquad (2.45)$$

where α is determined by the nuclear coordinates $\vec R$ via Eq. 2.38

Eq. 2.42 is a gauge transformation, which removes the discontinuities except for the one at the conical intersection, i.e. for at the origin $R = 0$. At the conical intersection, the Born-Oppenheimer wave function is undetermined, which reflects that any unitary transformation of a degenerate set of eigenstates produces another, equally valid, set of eigenstates. We may fix this ambiguity by the choice that for $R = 0$ the value $\alpha = 0$ is taken.

Nevertheless the discontinuity at the origin remains. Discontinuities in the Born-Oppenheimer wave functions lead to δ-function-like contributions in the derivative couplings.

The important message is that the Born-Oppenheimer wave function near a conical intersection cannot be chosen real without introducing steps in the wave function. This invalidates many arguments which indicate that the derivative couplings may be neglegible.

Derivative couplings The next step towards the equation of motion for the nuclear wave functions is to evaluate the derivative couplings.

$$\vec A_{m,n} \overset{\text{Eq. 2.22}}{=} \left\langle \Psi^{BO}_{m}(\vec R)\right| \frac{\hbar}{i}\vec\nabla_{\vec R} \left|\Psi^{BO}_{n}(\vec R)\right\rangle \qquad (2.46)$$

with $m, n \in \{-,+\}$.

---

[8] The additional sign change has been introduced to make the equations appear simpler. It is permitted because it is a global sign change.

[9] Using $\cos(x) = \frac{1}{2}(\mathrm{e}^{ix} + \mathrm{e}^{-ix})$ and $\sin(x) = \frac{1}{2i}(\mathrm{e}^{ix} - \mathrm{e}^{-ix})$

We transform the gradient into polar coordinates defined in Eq. 2.38

$$
\begin{aligned}
\begin{pmatrix} \partial_R \\ \partial_\alpha \end{pmatrix} &= \begin{pmatrix} \frac{\partial_X}{\partial_R}; \frac{\partial_Z}{\partial_R} \\ \frac{\partial_X}{\partial_\alpha}; \frac{\partial_Z}{\partial_\alpha} \end{pmatrix} \begin{pmatrix} \partial_X \\ \partial_Z \end{pmatrix} = \begin{pmatrix} \sin(\alpha); & \cos(\alpha) \\ R\cos(\alpha); & -R\sin(\alpha) \end{pmatrix} \begin{pmatrix} \partial_X \\ \partial_Z \end{pmatrix} \\
\Rightarrow \vec{\nabla}_{\vec{R}} = \begin{pmatrix} \partial_X \\ \partial_Z \end{pmatrix} &= \begin{pmatrix} \sin(\alpha); & +\frac{1}{R}\cos(\alpha) \\ \cos(\alpha); & -\frac{1}{R}\sin(\alpha) \end{pmatrix} \begin{pmatrix} \partial_R \\ \partial_\alpha \end{pmatrix}
\end{aligned}
\tag{2.47}
$$

$$
\begin{aligned}
\vec{A}_{m,n} &\overset{\text{Eq. 2.22}}{=} \left\langle \Psi_m^{BO}(\vec{R}) \right| \frac{\hbar}{i}\vec{\nabla} \left| \Psi_n^{BO}(\vec{R}) \right\rangle \\
&\overset{\text{Eq. 2.44}}{=} \frac{\hbar}{iR} \begin{pmatrix} \cos(\alpha) \\ -\sin(\alpha) \end{pmatrix} \left\langle \Psi_m^{BO}(\vec{R}) \right| \partial_\alpha \left| \Psi_n^{BO}(\vec{R}) \right\rangle
\end{aligned}
\tag{2.48}
$$

With the Born-Oppenheimer wave functions from Eq. 2.44 we obtain

$$
\begin{aligned}
\left\langle \Psi_+^{BO}(\vec{R}) \right| \partial_\alpha \left| \Psi_+^{BO}(\vec{R}) \right\rangle &= \left(\frac{1}{2}(\mathrm{e}^{i\alpha}+1)\right)^* \left(\frac{i}{2}\mathrm{e}^{i\alpha}\right) + \left(-\frac{i}{2}(\mathrm{e}^{i\alpha}-1)\right)^* \left(\frac{1}{2}\mathrm{e}^{i\alpha}\right) = \frac{i}{2} \\
\left\langle \Psi_+^{BO}(\vec{R}) \right| \partial_\alpha \left| \Psi_-^{BO}(\vec{R}) \right\rangle &= \left(\frac{1}{2}(\mathrm{e}^{i\alpha}+1)\right)^* \left(-\frac{1}{2}\mathrm{e}^{i\alpha}\right) + \left(-\frac{i}{2}(\mathrm{e}^{i\alpha}-1)\right)^* \left(\frac{i}{2}\mathrm{e}^{i\alpha}\right) = -\frac{1}{2} \\
\left\langle \Psi_-^{BO}(\vec{R}) \right| \partial_\alpha \left| \Psi_+^{BO}(\vec{R}) \right\rangle &= \left(\frac{i}{2}(\mathrm{e}^{i\alpha}-1)\right)^* \left(\frac{i}{2}\mathrm{e}^{i\alpha}\right) + \left(\frac{1}{2}(\mathrm{e}^{i\alpha}+1)\right)^* \left(\frac{1}{2}\mathrm{e}^{i\alpha}\right) = \frac{1}{2} \\
\left\langle \Psi_-^{BO}(\vec{R}) \right| \partial_\alpha \left| \Psi_-^{BO}(\vec{R}) \right\rangle &= \left(\frac{i}{2}(\mathrm{e}^{i\alpha}-1)\right)^* \left(-\frac{1}{2}\mathrm{e}^{i\alpha}\right) + \left(\frac{1}{2}(\mathrm{e}^{i\alpha}+1)\right)^* \left(\frac{i}{2}\mathrm{e}^{i\alpha}\right) = \frac{i}{2}
\end{aligned}
\tag{2.49}
$$

which yields the first-derivative couplings

$$
\begin{aligned}
\vec{A}_{+,+} &= \frac{\hbar}{2R} \begin{pmatrix} \cos(\alpha) \\ -\sin(\alpha) \end{pmatrix} = \frac{\hbar}{2R^2} \begin{pmatrix} Z \\ -X \end{pmatrix} \\
\vec{A}_{+,-} &= \frac{i\hbar}{2R} \begin{pmatrix} \cos(\alpha) \\ -\sin(\alpha) \end{pmatrix} = \frac{i\hbar}{2R^2} \begin{pmatrix} Z \\ -X \end{pmatrix} \\
\vec{A}_{-,+} &= \frac{-i\hbar}{2R} \begin{pmatrix} \cos(\alpha) \\ -\sin(\alpha) \end{pmatrix} = \frac{-i\hbar}{2R^2} \begin{pmatrix} Z \\ -X \end{pmatrix} \\
\vec{A}_{-,-} &= \frac{\hbar}{2R} \begin{pmatrix} \cos(\alpha) \\ -\sin(\alpha) \end{pmatrix} = \frac{\hbar}{2R^2} \begin{pmatrix} Z \\ -X \end{pmatrix}
\end{aligned}
\tag{2.50}
$$

FIRST-DERIVATIVE COUPLINGS OF THE JAHN-TELLER MODEL

$$
\begin{pmatrix} \vec{A}_{++} & \vec{A}_{+-} \\ \vec{A}_{-+} & \vec{A}_{--} \end{pmatrix} = \frac{\hbar}{2R^2} \begin{pmatrix} Z \\ -X \end{pmatrix} \otimes \begin{pmatrix} 1 & +i \\ -i & 1 \end{pmatrix}
\tag{2.51}
$$

The outer product between a vector and a matrix is defined as $(\vec{a} \otimes \boldsymbol{B})_{i,j,k} = a_i B_{j,k}$. It can be written a matrix

$$
\vec{a} \otimes \boldsymbol{B} = \begin{pmatrix} \vec{a}\, B_{1,1} & \vec{a}\, B_{1,2} \\ \vec{a}\, B_{2,1} & \vec{a}\, B_{2,2} \end{pmatrix}
\tag{2.52}
$$

with vectors as matrix elements.

Nuclear wave function At this point, we know the Born-Oppenheimer surfaces from Eq. 2.36, the electronic Born-Oppenheimer wave functions from Eq. 2.44 and the derivative couplings from Eq. 2.50. Thus we can set up the Schrödinger equation Eq. 2.20 for the nuclei.

$$\begin{aligned}
i\hbar\partial_t\phi_\pm \overset{\text{Eq. 2.20}}{=} & \left\{\frac{1}{2M}\left[\frac{\hbar}{i}\vec{\nabla}+\vec{A}_{\pm,\pm}\right]^2+\frac{1}{2M}\vec{A}_{\pm,\mp}\vec{A}_{\mp,\pm}+E^{BO}_{\pm}\right\}\phi_\pm \\
& +\frac{1}{2M}\left\{\left[\frac{\hbar}{i}\vec{\nabla}+\vec{A}_{\pm,\pm}\right]\vec{A}_{\pm,\mp}+\vec{A}_{\pm,\mp}\left[\frac{\hbar}{i}\vec{\nabla}+\vec{A}_{\mp,\mp}\right]\right\}\phi_\mp
\end{aligned} \tag{2.53}$$

Let me translate the Schrödinger equation into polar coordinates defined in Eq. 2.38 with Eq. 2.47. This yields

$$\frac{\hbar}{i}\vec{\nabla}+\vec{A}_{n,n}=\overbrace{\underbrace{\begin{pmatrix}\sin(\alpha)\\ \cos(\alpha)\end{pmatrix}\frac{\hbar}{i}\partial_R}_{(a)}+\underbrace{\begin{pmatrix}\cos(\alpha)\\ -\sin(\alpha)\end{pmatrix}\frac{\hbar}{iR}\partial_\alpha}_{(b)}}^{\frac{\hbar}{i}\vec{\nabla}}+\overbrace{\underbrace{\frac{\hbar}{2R}\begin{pmatrix}\cos(\alpha)\\ -\sin(\alpha)\end{pmatrix}}_{(c)}}^{\vec{A}_{n,n}} \tag{2.54}$$

so that

$$\begin{aligned}
\left(\frac{\hbar}{i}\vec{\nabla}+\vec{A}_{n,n}\right)^2 &= \overbrace{\underbrace{-\hbar^2\partial_R^2}_{(aa)}\underbrace{-\frac{\hbar^2}{R}\partial_R}_{ba}\underbrace{-\frac{\hbar^2}{R^2}\partial_\alpha^2}_{bb}}^{-\hbar^2\vec{\nabla}^2}+\overbrace{\underbrace{\frac{\hbar^2}{iR^2}\partial_\alpha}_{bc+cb}+\underbrace{\frac{\hbar^2}{4R^2}}_{cc}}^{\text{nonBO}} \\
&= \frac{-\hbar^2\left(R^2\partial_R^2+R\partial_R+\partial_\alpha^2\right)}{R^2}+\frac{\hbar^2}{R^2}\Big(-i\partial_\alpha+\frac{1}{4}\Big) \\
\vec{A}_{\pm,\mp}\Big(\frac{\hbar}{i}\vec{\nabla}+\vec{A}_{\mp,\mp}\Big) &= \pm\left(\frac{\hbar^2}{2R^2}\partial_\alpha+\frac{i\hbar^2}{4R^2}\right)=\pm i\frac{\hbar^2}{2R^2}\left(-i\partial_\alpha+\frac{1}{2}\right) \\
\Big(\frac{\hbar}{i}\vec{\nabla}+\vec{A}_{\pm,\pm}\Big)\vec{A}_{\pm,\mp} &= \pm\left(\frac{\hbar^2}{2R^2}\partial_\alpha+\frac{i\hbar^2}{4R^2}\right)=\pm i\frac{\hbar^2}{2R^2}\left(-i\partial_\alpha+\frac{1}{2}\right) \\
\vec{A}_{\pm,\mp}\vec{A}_{\mp,\pm} &= \frac{\hbar^2}{4R^2}
\end{aligned} \tag{2.55}$$

In polar coordinates, Eq. 2.53 has thus the form

$$\begin{aligned}
i\hbar\partial_t\phi_\pm = & \left\{\left[\underbrace{\frac{-\hbar^2\left(R^2\partial_R^2+R\partial_R+\partial_\alpha^2\right)}{2MR^2}}_{-\hbar^2\vec{\nabla}^2/2M}+\underbrace{\frac{\hbar^2\left(-i\partial_\alpha+\frac{1}{4}\right)}{2MR^2}}_{\text{nonBO}}\right]\underbrace{+\frac{\hbar^2}{8MR^2}}_{\text{nonBO}}\underbrace{\pm gR+wR^2}_{E^{BO}_\pm}\right\}\phi_\pm(R,\alpha,t) \\
& \underbrace{\pm i\frac{\hbar^2\left(-i\partial_\alpha+\frac{1}{2}\right)}{2MR^2}}_{\text{nonBO}}\phi_\mp(R,\alpha,t)
\end{aligned} \tag{2.56}$$

This equation can be rewritten in two-component form

$$\begin{aligned}
\Rightarrow \qquad i\hbar\partial_t\begin{pmatrix}\phi_+\\ \phi_-\end{pmatrix} = & \underbrace{\frac{-\hbar^2\left(R^2\partial_R^2+R\partial_R+\partial_\alpha^2\right)}{2MR^2}}_{-\hbar^2\vec{\nabla}^2/(2M)}\begin{pmatrix}\phi_+\\ \phi_-\end{pmatrix}+\begin{pmatrix}\left(+gR+wR^2\right)\phi_+\\ \left(-gR+wR^2\right)\phi_-\end{pmatrix} \\
& \underbrace{+\frac{\hbar^2\left(-i\partial_\alpha+\frac{1}{2}\right)}{2MR^2}\begin{pmatrix}1 & +i\\ -i & 1\end{pmatrix}}_{\text{nonBO}}\begin{pmatrix}\phi_+\\ \phi_-\end{pmatrix}
\end{aligned} \tag{2.57}$$

Now, I shift a factor i into ϕ_+, because this will lead to a radial Schrödinger equation with real coefficients. $i\phi_+$ being a real function implies that ϕ_+ is purely imaginary.

$$\begin{aligned} i\hbar\partial_t \begin{pmatrix} i\phi_+ \\ \phi_- \end{pmatrix} &= \underbrace{\frac{-\hbar^2\left(R^2\partial_R^2 + R\partial_R + \partial_\alpha^2\right)}{2MR^2}}_{-\hbar^2\vec{\nabla}^2/(2M)} \begin{pmatrix} i\phi_+ \\ \phi_- \end{pmatrix} + \begin{pmatrix} \left(+gR + wR^2\right) i\phi_+ \\ \left(-gR + wR^2\right)\phi_- \end{pmatrix} \\ &\quad + \frac{\hbar^2\left(-i\partial_\alpha + \frac{1}{2}\right)}{2MR^2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} i\phi_+ \\ \phi_- \end{pmatrix} \\ &= \Bigg\{ \underbrace{\frac{-\hbar^2\left(R^2\partial_R^2 + R\partial_R + \partial_\alpha^2\right)}{2MR^2}\mathbf{1}}_{-\hbar^2\vec{\nabla}^2/(2M)} + gR\boldsymbol{\sigma}_z + wR^2\mathbf{1} + \underbrace{\frac{\hbar^2\left(-i\partial_\alpha + \frac{1}{2}\right)}{2MR^2}\left(\mathbf{1} - \boldsymbol{\sigma}_x\right)}_{\text{nonBO}} \Bigg\} \begin{pmatrix} i\phi_+ \\ \phi_- \end{pmatrix} \end{aligned} \tag{2.58}$$

This nuclear Schrödinger equation has a rotational symmetry, because the angular coordinate does not enter in the differential operator. Only the angular derivatives enter. Therefore, I decompose the nuclear wave function into angular-momentum eigenstates. The angular-momentum eigenstates are further divided into the eigenstates of the nuclear Hamiltonian. [10]

$$\phi_\pm(\alpha, R, t) = \sum_m \mathrm{e}^{im\alpha} \sum_j \mathrm{e}^{-\frac{i}{\hbar}\mathcal{E}_{j,m}t} \mathcal{R}_{\pm,j,m}(R) c_{j,m} \tag{2.59}$$

The energies $\mathcal{E}_{m,j}$ are the those of the complete system including electronic and nuclear degrees of freedom. The angular-momentum quantum numbers m extend from $-\infty$ to $+\infty$. They are integer, so that the wave function is single valued.

Once the radial functions $\mathcal{R}_{\pm,m,j}(R)$ and it energies $\mathcal{E}_{m,j}$ are known, we can reconstruct the full wave function my multiplying this nuclear wave function with the Born-Oppenheimer wave functions.

$$|\Phi(t)\rangle = \sum_{\pm,m,j} \int dX \int dZ \, \Big(|a, X, Z\rangle \Psi^{BO}_{a,\pm}(R,\alpha) + |b, X, Z\rangle \Psi^{BO}_{b,\pm}(R,\alpha) \Big) \mathcal{R}_{\pm,j,m}(R) \mathrm{e}^{im\alpha} \mathrm{e}^{-\frac{i}{\hbar}\mathcal{E}_{j,m}t} c_{j,m} \tag{2.60}$$

Here, $R = \sqrt{X^2 + Z^2}$ and $\alpha = \mathrm{atan}\left(\frac{X}{Z}\right)$ are direct functions of the nuclear coordinates (X, Z).

We insert the Born-Oppenheimer wave functions from Eq. 2.44 to otain the final form of the electronic-nuclear wave function.

---

[10] Another way of looking at this is: After multiplication with $2MR^2$, the Hamiltonian is the sum of an angular and a radial Hamiltonien. This allows to separate the angular and radial variables (see ΦSX:Quantum physics) and use a product ansatz.

ELECTRONIC-NUCLEAR WAVE FUNCTION FOR THE JAHN-TELLER MODEL

The complete electronic nuclear wave functions expressed in the two electronic orbitals $|\vec{a}\rangle$ and $|\vec{a}\rangle$ describing the two e_g orbitals and the generalized nuclear coordinates (X, Z), which define the angular coordinate α.

$$|\Phi(t)\rangle = \sum_{m,j} |\Phi_{j,m}\rangle \mathrm{e}^{-\frac{i}{\hbar}\mathcal{E}_{j,m}t} c_{j,m} \tag{2.61}$$

The complex coefficients $c_{j,m} = \langle \Phi_{j,m}|\Phi(0)\rangle$ are defined by the initial conditions. The eigenstates of the Hamiltonian are

$$\begin{aligned} |\Phi_{j,m}\rangle = \int dX \int dZ\, \frac{1}{2} \bigg(& |a, X, Z\rangle \Big[(\mathrm{e}^{i\alpha}+1)\mathcal{R}_{+,j,m}(R) + i(\mathrm{e}^{i\alpha}-1)\mathcal{R}_{-,j,m}(R)\Big] \\ & + |b, X, Z\rangle \Big[-i(\mathrm{e}^{i\alpha}-1)\mathcal{R}_{+,j,m}(R) + (\mathrm{e}^{i\alpha}+1)\mathcal{R}_{-,j,m}(R)\Big]\bigg) \mathrm{e}^{im\alpha} \end{aligned} \tag{2.62}$$

The radial components are obtained from Eq. 2.58 with the ansatz Eq. 2.59 for the nuclear wave functions. Thus the defining equation is

$$\begin{aligned} \Bigg\{ & \underbrace{\frac{-\hbar^2\left(R^2\partial_R^2 + R\partial_R - m^2\right)}{2MR^2}\mathbf{1}}_{-\hbar^2\vec{\nabla}^2/(2M)} + \underbrace{gR\boldsymbol{\sigma}_z + wR^2\mathbf{1}}_{E^{BO}} - \mathcal{E}_{j,m}\mathbf{1} \\ & + \underbrace{\frac{\hbar^2\left(m+\frac{1}{2}\right)}{2MR^2}(\mathbf{1}-\boldsymbol{\sigma}_x)}_{\text{nonBO}}\Bigg\} \begin{pmatrix} i\mathcal{R}_{+,j,m}(R) \\ \mathcal{R}_{-,j,m}(R) \end{pmatrix} \overset{\text{Eq. 2.58}}{=} 0 \end{aligned} \tag{2.63}$$

The radial equation is a two-component system of coupled ordinary differential equations. It can be solved either by a power series ansatz, or by numerical integration of the differential equation.

The radial nuclear wave functions are shown in Fig. 2.4. For low energies the wave packet lies in the circular valley of the ground-state Born-Oppenheimer surface. With increasing energy the number of radial nodes increases. At the same time also the contribution of the excited-state Born-Oppenheimer surface grows. At higher energies, the wave function looks like a combination of a ground-state and an excited-state Born-Oppenheimer wave function at about the same energy. Thus the ground state contribution is spead over a wide region and has short oscillations which result from the high kinetic energy. The excited state surface has no or less nodes and is localized around the center, which is the minimum of the excited state Born-Oppenheimer surface.

What can be learned from this equation?

- At first we see that the eigenstates have contribution on both the upper and the lower sheet of the Born-Oppenheimer surface. This coupling, indicated in Eq. 2.63 as "nonBO", is scaled with the inverse mass. This seems to indicate that non-adiabatic effects become less important when the nuclear masses are large (in comparison to the electron mass.) However, the non-adiabatic coupling also diverges at the position of the conical intersection. That is, it cannot be ignored however large the nuclear masses are.

- An important observation is that the eigenstates are no angular-momentum eigenstates. In addition to the term $\mathrm{e}^{im\alpha}$ related to the nuclear wave function, there are also terms $\mathrm{e}^{i\alpha}$ from the Born-Oppenheimer wave function. Taken together, this implies that the eigenstates contain components from two angular momenta, namely $L_z = \hbar m$ and $L_z = \hbar(m+1)$. It is as if the conical intersection would have its own pseudo-spin of $\pm\frac{1}{2}\hbar$, which may interact with the nuclear "orbital motion".

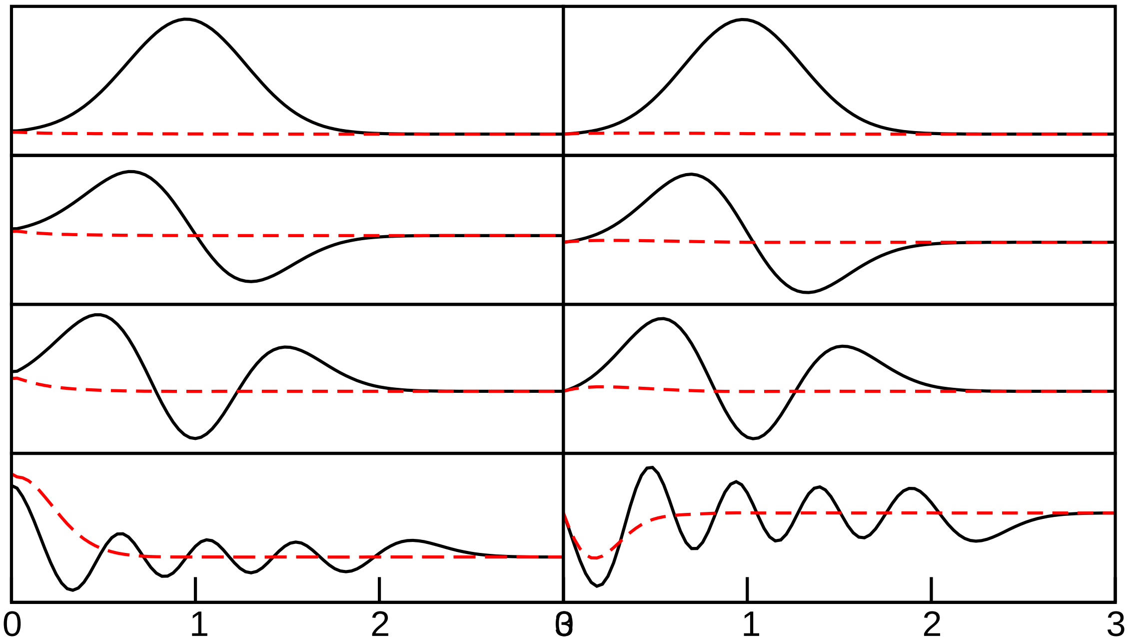


Fig. 2.4: Radial nuclear wave functions $\mathcal{R}_{\pm,j,m}(R)$ for the Jahn-Teller model for $j=0,1,2$ and $j=9$ from to to bottom and for $m=0$ left and $m=1$ right. The ground-state component is shown as black full line, and the component on the excited Born-Oppenheimer surface is shown as red dshed line. The energies for $m=0$ are $E_{j,m}=-0.89705,-0.69574,-0.49290,+0.80876$ for $j=0,1,2,9$ and $m=0$ and $E_{j,m}=-0.87415,-0.66431,-0.44920,+0.98900$ for $j=0,1,2,9$ and $m=1$. The model parameters are $M=\frac{1}{2}$, $g_{JT}=2$ and $w_{JT}=1$. The conical intersection is at $E=0$ and the minimum of the ground-state Born-Oppenheimer surface lies at $E=-1$ and $R=1$.

- Even though the Schrödinger equation is purely real, the Born-Oppenheimer wave functions at a conical intersection cannot be chosen real and at the same time continuous.
- The stationary wave functions for m and $-m-1$ are complex conjugates of each other, i.e.

  $$\Phi_{j,-m-1}(\vec{r},X,Z)=\Phi^{*}_{j,m}(\vec{r},X,Z) \tag{2.64}$$

  and the same holds for the radial functions

  $$\Big(\mathcal{R}_{\pm,j,-m-1}(R)\Big)^{*}=\mathcal{R}_{\pm,j,m}(R) \qquad \Leftrightarrow \qquad \begin{cases}\mathcal{R}_{-,j,-m-1}(R)=-\mathcal{R}_{-,j,m}(R)\\ \mathcal{R}_{+,j,-m-1}(R)=+\mathcal{R}_{+,j,m}(R)\end{cases} \tag{2.65}$$

  and the energies are related by $\mathcal{E}_{j,-m-1}=\mathcal{E}_{j,m}$. This implies that the corresponding solutions of the stationary Schrödinger equation are complex conjugates of each other. The derivation of Eq. 2.64 and Eq. 2.65 is given in Appendix G.1 on p. 263.

2.3.5 From a conical intersection to an avoided crossing

We have seen that avoided crossings and conical intersections are the points where non-adiabatic effects become important. In order to learn about the relation between the two, it is possible within the Jah-Teller model to change the conical intersection into an avoided crossing with a single parameter. Here, I only point out how one arrives at it. Its investigation is left to the interested reader.

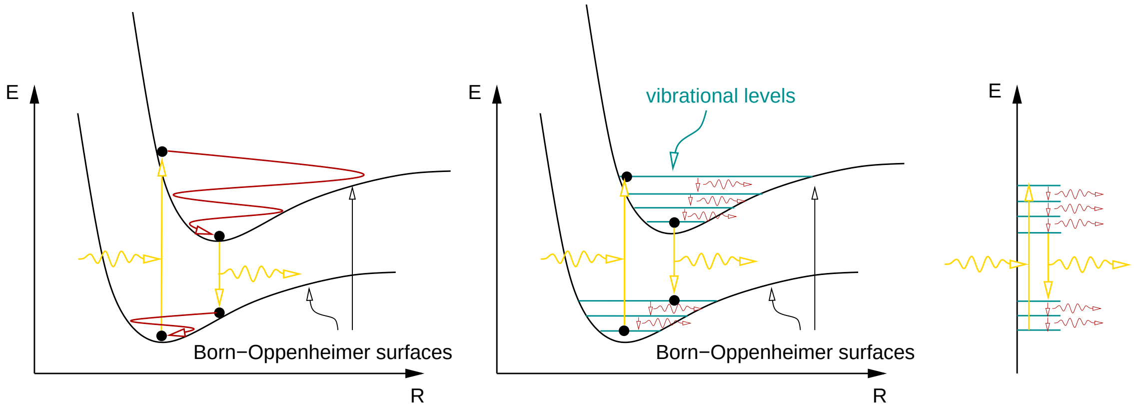


Fig. 2.5: Schematic for the optical absorption and desorption of a photon. Let R be one spatial coordinate for the nuclei, such as a bond distance. A photon excites an electron from the ground-state sheet of the total energy surface to an excited state sheet. The atoms vibrate around the equilibrium structure of the excited energy surface. During that process the system dissipates energy. In other words, it thermalizes by emitting phonons. At some point the electron drops again on the ground-state energy surface, while emitting a photon. The emitted photon has a lower energy than the absorbed photon, because some of the energy has been lost by creating for example phonos, i.e. heat. The system vibrates again about the equilibrium structure until it has dissipated its energy. Note that the transition proceeds between different energy levels of the total energy, which also include the vibrational levels. Thus vibrations or phonons are excited as well. Thus, to determine the absorption probability not only the optical transition matrix elements of the electronic subsystem must be considered, but also the overlap of the nuclear part of the wave function. Note that the figure describes the case at zero temperature. At finite temperature, the system can be initially already in a excited vibrational level. The left figure is appropriate for a classical description of the nuclear motion, while the right figure is an abstract energy level diagram appropriate for a quantum description.

The Jahn-Teller model can easily be extended to a three-dimensional parameter space by adding the third Pauli matrix in teh Schrödinger equation Eq. 2.32.

$$i\hbar\partial_t|\Phi(t)\rangle = \left[\frac{\hat{P}_x^2+\hat{P}_y^2+\hat{P}_z^2}{2M} + w\hat{\vec{R}}^2 + g\hat{X}\,\hat{\sigma}_x + g\hat{Y}\,\hat{\sigma}_y + g\hat{Z}\,\hat{\sigma}_z\right]|\Phi(t)\rangle \tag{2.66}$$

The conical intersection in the Born-Oppenheimer surface is at $(X,Y,Z)=\vec{0}$.

In the next step we can treat Y as a parameter, so that we obtain a two-dimensional model in the coordinate space (X,Z) but an additional parameter $Y=\lambda$

$$i\hbar\partial_t|\Phi(t)\rangle = \left[\frac{\hat{P}_x^2+\hat{P}_y^2}{2M} + w\hat{\vec{R}}^2 + g\hat{X}\,\hat{\sigma}_x + g\hat{Z}\,\hat{\sigma}_z + \lambda\,\hat{\sigma}_y\right]|\Phi(t)\rangle \tag{2.67}$$

For $\lambda=0$ we obtain the Jahn-Teller model with the conical intersection. For non-zero values of λ the conical intersection is converted into an avoided crossing with a gap of $2|\lambda|$.

2.4 Further reading

Here, a few reviews and other articles related to non-adiabatic effects. I have not read all of them carefully yet myself.

- Article by Butler[18]. Interesting is also the PhD thesis of Florian Dufey[11].

- A good introduction given in the special volume of the Journal "Advances in Chemical Physics" from 2002[19]. In particular, I liked the article of Child[20] and that by Worth and Robb[9]
- the book "Beyond Born-Oppenheimer: Conical Intersections and Electronic Nonadiabatic Couling Terms" by Michael Baer[21]. Baer uses a very compact notation, but he describes the topological effects in a very elegant manner.
- Review article by Worth and Cederbaum[22]
- Herzberg[23]
- For the role of the so-called **geometrical phase** or **Berry phase** see Berry's original paper[24]. For the role of the Berry phase in chemical reactions see the short article by Clary[25].

Chapter 3

Many-particle wave functions

In the previous section we have shown how we can split the description into two parts, one for the electrons and another one for the nuclei. The dynamics of the nuclei requires the knowledge of the electronic problem, namely the Born-Oppenheimer surfaces $E_n^{BO}(\vec{R})$ and the derivative couplings $\vec{A}_{m,n}(\vec{R})$ obtained from the Born-Oppenheimer wave functions $|\Psi^{BO}(\vec{R})\rangle$. The Born-Oppenheimer wave functions are states in the electronic Hilbert space and the atomic positions $\vec{R}$ are parameters and not coordinates in the quantum mechanical sense.

The next difficulty is to obtain the Born-Oppenheimer wave functions. For non-interacting electrons this is of the same complexity as a one-particle problem, while the interaction makes this problem usually untractable. Therefore we will start with non-interacting electrons before we turn to the more complicated interacting systems.

In the present section, I will describe the basics of many-particle wave functions. I will start out how to describe one-particle wave functions wityh spin. Then I will describe the underlying symmetry of identical particles, which resuls in the concept of Fermions and Bosons. Finally, I will make the turn back to one-particle wave functions, by showing how the wave function of a noninteracting system of fermions and bosons can be expressed by one-particle wave functions.

This will form the basis of the following sections that adress non-interacting electrons.

3.1 Spin orbitals

Before I continue with many-particle wave functions, let me discuss one-particle orbitals. I will introduce a notation for treating the electron spin which, on the one hand, simplifies the expressions, and on the other hand it is more rigorous.

When I first learned about spins and magnetic moments, I was puzzled by the special role, which was attributed to the z-coordinate. I learned that an electron can have a spin pointing parallel or antiparallel to the z-axis. The direction of the z-axis, however, is arbitrary. It seemed as if the physics changed, when I turned my head.

Indeed, there is nothing special about the z-coordinate. Here, I present a more general formulation, which describes electrons by two-component spinor wave functions. With this formulation, one can form wave functions with a spin pointing in an arbitrary direction in each point in space. This description restores the rotational symmetry that appears to be broken in the simple-minded formulation used commonly.

We use here **two-component spinor wave functions**, which are also called **spin-orbitals**

$$\phi(\vec{x}) = \phi(\vec{r}, \sigma) = \langle \vec{r}, \sigma | \phi \rangle$$

Spin-orbitals actually consists of two complex wave functions, one for the spin-down $\phi(\vec{r}, \downarrow)$ and one for the spin up contribution $\phi(\vec{r},\uparrow)$. One can write the spin orbitals also as two-component spinor

$$\begin{pmatrix} \phi(\vec{r},\uparrow) \\ \phi(\vec{r},\downarrow) \end{pmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} \langle \vec{r},\uparrow|\phi\rangle \\ \langle \vec{r},\downarrow|\phi\rangle \end{pmatrix}$$

PHYSICAL MEANING OF A SPIN WAVEFUNCTION

The square of a component of a spin wavefunction is the probability density $P_\sigma(\vec{r})$ for finding a particle with the specified spin orientation at a specific position $\vec{r}$.

This rule is completely analogous to the rule that the absolute square of a scalar wave function is the probability density $P(\vec{r}) = \psi^*(\vec{r})\psi(\vec{r})$ of finding a particle at a given position $\vec{r}$.

Note, however, that the wave function contains much more information than that just mentioned. Below, we will see that the spin wave function does not only provide the information on the probabilities for the spin projection on the z-axis but also for any other axis.[1]

Relation to spin eigenstates

This notation is a little puzzling at first, because one usually works with spin-eigenfunctions, for which one of the components vanishes. The spin is defined as

$$\hat{\vec{S}} = \frac{\hbar}{2}\hat{\vec{\sigma}} \tag{3.1}$$

where the three components of the vector $\hat{\vec{\sigma}}$ are

$$\hat{\sigma}_i = \int d^3r \sum_{\sigma,\sigma'=1}^{2} |\vec{r},\sigma\rangle\sigma_{i,\sigma,\sigma'}\langle\vec{r},\sigma'|$$

Somewhat confusing are the different indices for the different vector spaces: The index i can have the values x, y, z and refers to the components of the vector $\vec{\sigma}$. Each component of the vector $\hat{\vec{\sigma}}$ is a scalar operator. The indices σ, σ' of $\sigma_{i,\sigma,\sigma'}$ can assume the values $\uparrow$ or $\downarrow$, which are the coordinates of the two-dimensional spinor space. That is, $\boldsymbol{\sigma}_i$ is a (2×2) matrix for each value of i. These (2×2) matrices are the **Pauli matrices**[2]

$$\boldsymbol{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \boldsymbol{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \boldsymbol{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.2}$$

---

[1] This is analogous to the fact that a wave function in configuration space contains information also on the momentum.

[2] The matrices are made of the matrix elements $\sigma_{i,\sigma,\sigma'} = \langle\sigma|\hat{\sigma}_i|\sigma'\rangle$ of the operator $\hat{\sigma}_i$ with $\sigma \in \{\uparrow,\downarrow\}$, $\sigma' \in \{\uparrow,\downarrow\}$ and $i \in \{x,y,z\}$.

As a worked example, let us determine the expectation value of the operator $\hat{S}_x$ for a one-particle state $|\phi\rangle$. The states $|\vec{r},\uparrow\rangle$ and $|\vec{r},\uparrow\rangle$ are still eigenstates of $\hat{S}_z$.

$$\begin{aligned}
\langle\phi|\hat{S}_x|\phi\rangle &= \langle\phi|\left[\frac{\hbar}{2}\int d^3r \sum_{\sigma,\sigma'=1}^{2} |\vec{r},\sigma\rangle\sigma_{x,\sigma,\sigma'}\langle\vec{r},\sigma'|\right]|\phi\rangle \\
&= \frac{\hbar}{2}\int d^3r \left[\sum_{\sigma,\sigma'=1}^{2}\langle\phi|\vec{r},\sigma\rangle\sigma_{x,\sigma,\sigma'}\langle\vec{r},\sigma'|\phi\rangle\right] \\
&= \frac{\hbar}{2}\int d^3r \left[\begin{pmatrix}\langle\phi|\vec{r},\uparrow\rangle \\ \langle\phi|\vec{r},\downarrow\rangle\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\langle\vec{r},\uparrow|\phi\rangle \\ \langle\vec{r},\downarrow|\phi\rangle\end{pmatrix}\right] \\
&= \frac{\hbar}{2}\int d^3r \left[\langle\phi|\vec{r},\uparrow\rangle\langle\vec{r},\downarrow|\phi\rangle + \langle\phi|\vec{r},\downarrow\rangle\langle\vec{r},\uparrow|\phi\rangle\right] \\
&= \frac{\hbar}{2}\int d^3r \left[\phi^*(\vec{r},\uparrow)\phi(\vec{r},\downarrow) + \phi^*(\vec{r},\downarrow)\phi(\vec{r},\uparrow)\right]
\end{aligned}$$

We obtain the spin density, that is the probability that we find a particle at position $\vec{r}$ multiplied with the average spin expectation value in x-direction of that particle.

For a spin orbital that is an eigenstate of S_z, one of the spinor components vanishes.

$$\begin{aligned}
\hat{S}_z|\phi_\uparrow\rangle &= |\phi_\uparrow\rangle\left(+\frac{\hbar}{2}\right) &\Rightarrow&\quad \begin{pmatrix}\langle\vec{r},\uparrow|\phi_\uparrow\rangle \\ \langle\vec{r},\downarrow|\phi_\uparrow\rangle\end{pmatrix} = \begin{pmatrix}\phi_\uparrow(\vec{r},\uparrow) \\ 0\end{pmatrix} \\
\hat{S}_z|\phi_\downarrow\rangle &= |\phi_\downarrow\rangle\left(-\frac{\hbar}{2}\right) &\Rightarrow&\quad \begin{pmatrix}\langle\vec{r},\uparrow|\phi_\downarrow\rangle \\ \langle\vec{r},\downarrow|\phi_\downarrow\rangle\end{pmatrix} = \begin{pmatrix}0 \\ \phi_\downarrow(\vec{r},\downarrow)\end{pmatrix}
\end{aligned}$$

Note that the indices $\uparrow,\downarrow$ indicate the quantum number, while the pointers to the spinor components are treated as argument, so that they can be distinguished from the quantum numbers. The quantum number indicates that the orbital is an eigenstate to some symmetry operator.

The spin-orbitals have the advantage that they allow to describe orbitals, for which the spin does not point along the z-axis, but it can also point, for example, to the right or in any other direction. For a spin orbital the spin direction can actually vary in space. Such an orbital is called non-collinear, because the spins are not aligned.

Problem with conventional notation

In the literature, one usually uses another notation, namely

$$\phi_\sigma(\vec{r})\stackrel{\text{def}}{=}\phi(\vec{r},\sigma)$$

Here, it is difficult to distinguish the role of σ as a coordinate and as quantum number. One usually uses one-particle orbitals with only one spin component, for which σ is a quantum number. In our notation this would be – for a spin-up particle –

$$\phi_\uparrow(\vec{r},\sigma) = \langle\vec{r},\sigma|\phi_\uparrow\rangle\hat{=}\begin{pmatrix}\langle\vec{r},\uparrow|\phi_\uparrow\rangle \\ 0\end{pmatrix}$$

The orbital still has two components, but one of them, namely $\phi_\uparrow(\vec{r},\downarrow)$, vanishes. With the common notation, it is difficult to write down expressions that do not break the rotational symmetry for the spin direction.

Magnetization

So-far we used the spin operator to determine the expectation value of the spin. Now we would like to obtain the spin density or the magnetization respectively. In order to describe the principles, let us work out the charge density as a trivial example:

We know the projector $\hat{P}(\vec{r})$ onto a certain point in space $\vec{r}$, which is

$$\hat{P}(\vec{r}) \stackrel{\text{def}}{=} \sum_{\sigma} |\vec{r},\sigma\rangle\langle\vec{r},\sigma| = |\vec{r},\downarrow\rangle\langle\vec{r},\downarrow| + |\vec{r},\uparrow\rangle\langle\vec{r},\uparrow|$$

We obtain the charge density as expectation value of $\hat{P}(\vec{r})$, multiplied with the electron charge $q_e = -e$.

$$\begin{aligned}
\rho(\vec{r}) &= \langle\phi|\left[q\hat{P}(\vec{r})\right]|\phi\rangle = q\langle\phi|\left[\sum_{\sigma}|\vec{r},\sigma\rangle\langle\vec{r},\sigma|\right]|\phi\rangle \\
&= q\sum_{\sigma}\langle\phi|\vec{r},\sigma\rangle\langle\vec{r},\sigma|\phi\rangle \\
&= q\Big(\langle\phi|\vec{r},\downarrow\rangle\langle\vec{r},\downarrow|\phi\rangle + \langle\phi|\vec{r},\uparrow\rangle\langle\vec{r},\uparrow|\phi\rangle\Big)
\end{aligned}$$

Hence, the charge density is, up to the factor q, the sum of the spin-up and spin-down densities. After this introduction, we can analogously determine the magnetization.

The magnetization operator is obtained as product of the factor[3] $\frac{q}{m}$, the spin operator and the projection operator onto a point in space.

$$\begin{aligned}
\hat{m}_i(\vec{r}) &\stackrel{\text{def}}{=} \frac{q}{m}\hat{P}(\vec{r})\hat{S}_i = \frac{q\hbar}{2m}\hat{P}(\vec{r})\hat{\sigma}_i \\
&= \frac{q\hbar}{2m}\underbrace{\left[\sum_{\sigma}|\vec{r},\sigma\rangle\langle\vec{r},\sigma|\right]}_{\hat{P}(\vec{r})}\underbrace{\left[\int d^3r'\sum_{\sigma',\sigma''}|\vec{r'},\sigma'\rangle\sigma_{i,\sigma',\sigma''}\langle\vec{r'},\sigma''|\right]}_{\hat{\sigma}_i} \\
&= \frac{q\hbar}{2m}\sum_{\sigma}\int d^3r'\sum_{\sigma',\sigma''}|\vec{r},\sigma\rangle\underbrace{\langle\vec{r},\sigma|\vec{r'},\sigma'\rangle}_{\delta(\vec{r}-\vec{r'})\delta_{\sigma,\sigma'}}\sigma_{i,\sigma',\sigma''}\langle\vec{r'},\sigma''| \\
&= \frac{q\hbar}{2m}\sum_{\sigma,\sigma''}|\vec{r},\sigma\rangle\sigma_{i,\sigma,\sigma''}\langle\vec{r},\sigma''|
\end{aligned}$$

The factor $\mu_B \stackrel{\text{def}}{=} \frac{e\hbar}{2m_e}$ is the **Bohr magneton**, which is approximately equal to the magnetic moment of an electron. Spin and magnetic moment of the electron point in opposite directions due to the negative charge of the electron.

MAGNETIZATION OPERATOR

$$\hat{\vec{m}}(\vec{r}) = \frac{q\hbar}{2m}\sum_{\sigma,\sigma'}|\vec{r},\sigma\rangle\vec{\sigma}_{\sigma,\sigma'}\langle\vec{r},\sigma'| \tag{3.3}$$

with the Pauli matrices $\vec{\boldsymbol{\sigma}}$ defined in Eq. 3.2.

---

[3] The classical ratio of magnetic moment and angular momentum is $\frac{q}{2m}$, which however assumes a constant ratio the mass- and charge distribution in space. While the "distribution" of electrons leads to such a constant ratio, the properties of a single electron may differ in a classical picture. In a quantum mechanical description the electron is point-like with a certain gyromagnetic ratio, that follows directly from the relativistic Dirac equation.

More explicitly we obtain

$$
\begin{aligned}
\rho(r) &= q \begin{pmatrix} \langle\phi|\vec{r},\uparrow\rangle \\ \langle\phi|\vec{r},\downarrow\rangle \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \langle\vec{r},\uparrow|\phi\rangle \\ \langle\vec{r},\downarrow|\phi\rangle \end{pmatrix} \\
&= q\left[\phi^*(\vec{r},\uparrow)\phi(\vec{r},\uparrow) + \phi^*(\vec{r},\downarrow)\phi(\vec{r},\downarrow)\right] \\
m_x(r) &= \frac{q\hbar}{2m} \begin{pmatrix} \langle\phi|\vec{r},\uparrow\rangle \\ \langle\phi|\vec{r},\downarrow\rangle \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \langle\vec{r},\uparrow|\phi\rangle \\ \langle\vec{r},\downarrow|\phi\rangle \end{pmatrix} \\
&= \frac{q\hbar}{2m_e}\left[\phi^*(\vec{r},\uparrow)\phi(\vec{r},\downarrow) + \phi^*(\vec{r},\downarrow)\phi(\vec{r},\uparrow)\right] \\
m_y(r) &= \frac{q\hbar}{2m} \begin{pmatrix} \langle\phi|\vec{r},\uparrow\rangle \\ \langle\phi|\vec{r},\downarrow\rangle \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} \langle\vec{r},\uparrow|\phi\rangle \\ \langle\vec{r},\downarrow|\phi\rangle \end{pmatrix} \\
&= -i\frac{q\hbar}{2m}\left[\phi^*(\vec{r},\uparrow)\phi(\vec{r},\downarrow) - \phi^*(\vec{r},\downarrow)\phi(\vec{r},\uparrow)\right] \\
m_z(r) &= \frac{q\hbar}{2m} \begin{pmatrix} \langle\phi|\vec{r},\uparrow\rangle \\ \langle\phi|\vec{r},\downarrow\rangle \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \langle\vec{r},\uparrow|\phi\rangle \\ \langle\vec{r},\downarrow|\phi\rangle \end{pmatrix} \\
&= \frac{q\hbar}{2m}\left[\phi^*(\vec{r},\uparrow)\phi(\vec{r},\uparrow) - \phi^*(\vec{r},\downarrow)\phi(\vec{r},\downarrow)\right]
\end{aligned}
$$

The two-component spinor descriptions follows from the **Pauli equation**. The Pauli equation is the non-relativistic limit of the **Dirac equation**, the relativistic one-particle equation for electrons. In the Dirac equation each particle has four components. Two components describe spin-up and spin-down electrons, and the two other components describe spin-up and spin-down positrons. The **positron**[26] is the anti-particle of the electron. In the non-relativistic limit, the electronic and positronic components become independent of each other, so that the electrons can be described by a two-component spinor wave function.

A Slater determinant of N spin-orbitals corresponds to 2^N Slater determinants of one-component one-particle orbitals, according to the N spin indices.

3.2 Symmetry and quantum mechanics

In the following, we will be concerned with the symmetry of the Hamilton operator under permutation of particles. Therefore, I will revisit the main **symmetry** arguments discussed in section 9 of Φ*SX:Quantum Theory*[27]. This is a series of arguments may be worthwhile to keep in mind.

A symmetry of an object is a transformation, that leaves the appearance of the object unchanged. For example a square is symmetric under four-fold rotation.

A physical system is characterized[4] by its Hamiltonian, which determines the dynamics of the system by the Schrödinger equation. The system is symmetric under a given transformation, if every solution of the Schrödinger equation, is transformed into another solution of the same Schrödinger equation.

Let me formalize this:

1. Definition of a **transformation operator**: An operator $\hat{S}$ can be called a transformation, if it conserves the norm for any state, that is if

$$\forall_{|\psi\rangle} \qquad \langle\psi|\psi\rangle \overset{|\phi\rangle=\hat{S}|\psi\rangle}{=} \langle\phi|\phi\rangle \tag{3.4}$$

2. A transformation operator $\hat{S}$ is unitary, that is

$$\hat{S}^\dagger\hat{S} = 1 \tag{3.5}$$

---

[4] The argument is not limited to quantum systems. The Schrödinger equation can be replaced by any other equation of motion governing the dynamics or other properties.

Proof:

$$\forall_{|\psi\rangle}\quad \langle\psi|\hat{S}^\dagger\hat{S}|\psi\rangle \overset{\text{Eq. 3.4}}{=} \langle\psi|\psi\rangle \qquad \Rightarrow \qquad \hat{S}^\dagger\hat{S}=1$$

3. Definition of a **symmetry**: A system is called symmetric under the transformation $\hat{S}$, if, for any solution of the Schrödinger equation describing that system, also $\hat{S}|\Psi\rangle$ is a solution of the same Schrödinger equation. That is, if

$$\Big(\ i\hbar\partial_t|\Psi\rangle=\hat{H}|\Psi\rangle\ \Big) \overset{|\Phi\rangle\overset{\text{def}}{=}\hat{S}|\Psi\rangle}{\Rightarrow} \Big(\ i\hbar\partial_t|\Phi\rangle=\hat{H}|\Phi\rangle\ \Big) \tag{3.6}$$

4. The commutator of the Hamilton operator with its symmetry operator vanishes, that is $[\hat{H},\hat{S}]_-=0$:

   Proof:

$$\begin{aligned} i\hbar\partial_t|\Phi\rangle &= \hat{H}|\Phi\rangle \\ \overset{|\Phi\rangle\overset{\text{def}}{=}\hat{S}|\Psi\rangle}{\Rightarrow}\ i\hbar\partial_t\hat{S}|\Psi\rangle &= \hat{H}\hat{S}|\Psi\rangle \\ \overset{\partial_t\hat{S}=0}{\Rightarrow}\ \hat{S}\left(i\hbar\partial_t|\Psi\rangle\right) &= \hat{H}\hat{S}|\Psi\rangle \\ \overset{i\hbar\partial_t|\Psi\rangle=\hat{H}|\Psi\rangle}{\Rightarrow}\ \hat{S}\hat{H}|\Psi\rangle &= \hat{H}\hat{S}|\Psi\rangle \\ \underbrace{\left(\hat{H}\hat{S}-\hat{S}\hat{H}\right)}_{[\hat{H},\hat{S}]_-}|\Psi\rangle &= 0 \end{aligned}$$

   Because this equation holds for any solution of the Schrödinger equation, it holds for any wave function, because any function can be written as superposition of solutions of the Schrödinger equation. (The latter form a complete set of functions.) Therefore

SYMMETRY AND COMMUTATOR

The commutator between the Hamilton operator $\hat{H}$ with its (time-independent[a]) symmetry operators $\hat{S}$ vanishes.

$$[\hat{H},\hat{S}]_-=0$$

Thus, one usually identifies a symmetry by working out the commutator with the Hamiltonian.

---

[a] For time-dependent symmetry operators, the more general rule is $[\hat{H},\hat{S}]_-=i\hbar\partial_t\hat{S}$. An example for a time-dependent symmetry is that between two relatively moving frames of inertia.

5. The matrix elements of the Hamilton operator between two eigenstates of the symmetry operator with different eigenvalues vanish. That is

$$\Big(\ \hat{S}|\Psi_s\rangle=|\Psi_s\rangle s \quad\wedge\quad \hat{S}|\Psi_{s'}\rangle=|\Psi_{s'}\rangle s' \quad\wedge\quad s\neq s'\ \Big) \quad\Rightarrow\quad \langle\Psi_s|\hat{H}|\Psi_{s'}\rangle=0$$

   Thus, the Hamilton operator is block diagonal in a representation of eigenstates of its symmetry operators. The eigenstates of the Hamilton operator can be obtained for each block individually. Because the effort to diagonalize a matrix increases rapidly with the matrix size, this simplifies the calculation of Hamilton eigenstates substantially, both for analytical and for numerical calculations.

   If the Hamiltonian is block diagonal, the time dependent Schrödinger equation does not mix states from different subspaces. A wave function that starts out as an eigenstate of a symmetry operator with a given eigenvalue, will always be an eigenstate with the same eigenvalue. In other words, the eigenvalue of the symmetry operator is a conserved quantity. (Note, however, that the eigenvalue of a symmetry operator is usually complex.)

The eigenvalues of the symmetry operators are related to the **quantum numbers**.

Proof: In the following we will need an expression for $\langle\psi_s|\hat{S}$, which we will work out first:

- We start by showing that the absolute value of an eigenvalue of a unitary operator is equal to one, that is $s = \mathrm{e}^{i\phi}$ where ϕ is real. With an eigenstate $|\psi_s\rangle$ of $\hat{S}$ we obtain

$$s^*\langle\psi_s|\psi_s\rangle s \overset{\hat{S}|\psi_s\rangle=|\psi_s\rangle s}{=} \langle\hat{S}\psi_s|\hat{S}\psi_s\rangle = \langle\psi_s|\underbrace{\hat{S}^\dagger\hat{S}}_{\hat{1}}|\psi_s\rangle = \langle\psi_s|\psi_s\rangle \qquad \Rightarrow \qquad |s| = 1 \tag{3.7}$$

- Next, we show that the eigenvalues of the Hermitian conjugate operator $\hat{S}^\dagger$ of a unitary operator $\hat{S}$ are the complex conjugates of the eigenvalues of $\hat{S}$.

$$\begin{aligned}
\Rightarrow |\psi_s\rangle &\overset{\text{Eq. 3.5}}{=} \underbrace{\hat{S}^\dagger\hat{S}}_{=\hat{1}}|\psi_s\rangle \overset{\hat{S}|\psi_s\rangle=|\psi_s\rangle s}{=} \hat{S}^\dagger|\psi_s\rangle s \\
\Rightarrow \qquad \hat{S}^\dagger|\psi_s\rangle s &\overset{\text{Eq. 3.7}}{=} |\psi_s\rangle\underbrace{s^*s}_{=1} \\
\Rightarrow \qquad \hat{S}^\dagger|\psi_s\rangle &= |\psi_s\rangle s^* \\
\Rightarrow \qquad \langle\psi_s|\hat{S} &= s\langle\psi_s|
\end{aligned}$$

- With this, we are ready to show that the matrix elements of the Hamilton operator between two eigenstates of the symmetry operator with different eigenvalues vanish.

$$\begin{aligned}
0 &\overset{[\hat{H},\hat{S}]_-=0}{=} \langle\Psi_s|[\hat{H},\hat{S}]_-|\Psi_{s'}\rangle = \langle\Psi_s|\hat{H}\hat{S}|\Psi_{s'}\rangle - \langle\Psi_s|\hat{S}\hat{H}|\Psi_{s'}\rangle \\
&\overset{\hat{S}|\Psi_s\rangle=|\Psi_s\rangle s, etc.}{=} \langle\Psi_s|\hat{H}|\Psi_{s'}\rangle s' - s\langle\Psi_s|\hat{H}|\Psi_{s'}\rangle = \langle\Psi_s|\hat{H}|\Psi_{s'}\rangle(s'-s) \\
&\overset{s\neq s'}{\Rightarrow} \langle\Psi_s|\hat{H}|\Psi_{s'}\rangle = 0
\end{aligned}$$

q.e.d

6. For a operator $\hat{S}$ with the property $\hat{S}^N = \hat{1}$ for a finite N, one can decompose any wave function $|\chi\rangle$ into eigenstates $|\psi_\alpha\rangle$ of the symmetry operation $\hat{S}$ with eigenvalue $s_\alpha = \mathrm{e}^{i\frac{2\pi}{N}\alpha}$

$$|\psi_\alpha\rangle = \sum_{n=0}^{N}\hat{S}^n|\chi\rangle s_\alpha^{-n} = \sum_{n=0}^{N}\left(\hat{S}^\dagger\right)^n|\chi\rangle s_\alpha^{+n} \tag{3.8}$$

3.3 Identical particles

Electrons with the same spin are indistinguishable. This says that there is no conceivable experiment that discriminates between two electrons, except for their spin. Thus, there is a symmetry with respect to exchange of two particles, and the Hamiltonian for indistinguishable particles commutes with the permutation operator of two particles.

Editor: introduce symbol $\hat{\mathcal{P}}$ for the permutation operator to avoid confusion with a projection operator $\hat{P}(\vec{r})$ used later.

The **two-particle permutation operator** is defined as

$$\hat{\mathcal{P}}^{(2)}_{i,j}\Psi(\ldots,\vec{x}_i,\ldots,\vec{x}_j\ldots) = \Psi(\ldots,\vec{x}_j,\ldots,\vec{x}_i\ldots)$$

The eigenstates of the Hamiltonian are eigenstates of the permutation operator. Because $\left(\hat{\mathcal{P}}^{(2)}_{i,j}\right)^2 = \hat{1}$, the permutation operator has the two eigenvalues, namely $+1$ and -1. Thus, the wave functions are fully symmetric or fully antisymmetric with respect to the permutation of two particle coordinates. Particles with a symmetric wave function are called **Bosons** and particles with an antisymmetric wave function are called **Fermions**. Bosons are particles with integer spin such as photons, mesons, gluons, gravitons, etc, which are usually related to an interaction, while Fermions are particles with half-integer spin such as electrons, protons, neutrons, quarks, etc.

Unlike other symmetries, this symmetry is not a property of a specific Hamiltonian, but it is a property of the particles themselves. If the particles are indistinguishable, there is no conceivable Hamiltonian, that is not symmetric under permutation of two of these particles. If we could construct only one Hamiltonian that is not symmetric, we could design an experiment, just by realizing this Hamiltonian, that distinguishes two of such particles

The electron exchange is illustrated in Fig. 3.1. Except for the color, which distinguishes the two electrons, left and right situations are identical. Note, that for the particle exchange both spatial and spin-indices have to be exchanged simultaneously. Since the electrons are indistinguishable, no distinguishing property like the color indicated can exist.

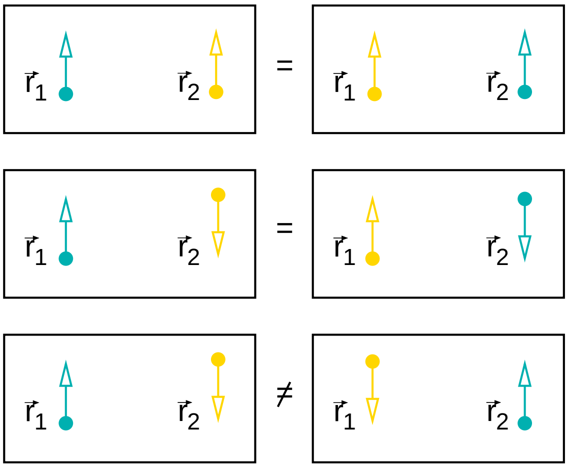


Fig. 3.1: Demonstration of the particle exchange. For the two figures at the top, left and right situations are identical except for a color of the particles. For identical particles there is no property such as the color that allows one to distinguish them. Thus $\Psi(\vec{x}_1,\vec{x}_2) = -\Psi(\vec{x}_2,\vec{x}_1)$ or $\Psi(\vec{r}_1,\sigma_1,\vec{r}_2,\sigma_2) = -\Psi(\vec{r}_2,\sigma_2,\vec{r}_1,\sigma_1)$. In the bottom figure, the particle positions are exchanged but not their spin. Thus, the two configurations are not identical. In this case we may also view electrons with up and down spin as distinguishable particles, because they have different spin.

Fermions, such as electrons, have an antisymmetric wave function, that is

$$\hat{\mathcal{P}}^{(2)}_{i,j}|\Psi\rangle = -|\Psi\rangle$$

3.3.1 Levi-Civita Symbol or the fully antisymmetric tensor

In the following, I will represent wave functions by determinants. In order to work with determinants, the Levi-Civita symbol, also called the fully antisymmetric tensor, will be introduced. Here, I define the fully antisymmetric tensor and derive some formulas, which will be needed later.

Definition

The **Levi-Civita symbol**[5], also called the **fully antisymmetric tensor**

---

[5] Tullio Levi-Civita (1873-1941): Italian mathematician. Invented the covariant derivative. Made tensor-algebra popular, which was used in Einstein's theory of general relativity.

is a **rank**[6] N Tensor defined by the following properties

- For ascending indices the Levi-Civita symbol has the value one:

$$\epsilon_{1,2,3,\ldots,N}=1$$

- The Levi-Civita symbol changes sign with any pairwise permutation of its indices
- The Levi-Civita symbol vanishes whenever at least two indices are pairwise identical

Relation to determinants

The determinant of a $N\times N$ matrix $\boldsymbol{A}$ is defined by the Levi-Civita symbol $\varepsilon_{i,j,k,l,\ldots}$ as

$$\det[\boldsymbol{A}]=\sum_{i,j,k,\ldots}\varepsilon_{i,j,k,\ldots}A_{1,i}A_{2,j}A_{3,k}\cdots$$

Vector product (not needed)

A common and useful application of the Levi-Civita Symbol is to express the vector product in three dimensions by its components.

$$\vec{a}=\vec{b}\times\vec{c}\qquad\Leftrightarrow\qquad a_i=\sum_{j,k}\varepsilon_{i,j,k}b_jc_k$$

3.3.2 Permutation operator

The most simple way to construct a many-particle wave function, is to take the product of **one-particle wave functions** $\phi_j(\vec{x})$, such as

$$\Psi(\vec{x}_1,\vec{x}_2,\ldots)=\phi_a(\vec{x}_1)\phi_b(\vec{x}_2)\cdots$$

This is a **product wave function**.

Such a product wave function is neither symmetric nor antisymmetric with respect to permutation. Therefore, we need to antisymmetrize it. We use the N-particle **permutation operator** $\hat{\mathcal{P}}_{i_1,\ldots,i_N}$ that is defined as follows

N-PARTICLE PERMUTATION OPERATOR

$$\hat{\mathcal{P}}_{i_1,\ldots,i_N}:=\int d^4x_1\cdots\int d^4x_N\;|\vec{x}_{i_1},\ldots,\vec{x}_{i_N}\rangle\langle\vec{x}_1,\ldots,\vec{x}_N| \tag{3.9}$$

The permutation operator reorders the coordinates so that the first particle is placed onto the position of the particle i_1 and so on. This implies that the coordinates of the i_1-th particle are placed on the first position, which is the one of the first particle.

In order to make the function of the operator more transparent, let us rewrite it as an operator of real-space functions. Consider a state $|\vec{y}_1,\ldots,\vec{y}_N\rangle$, which describes a particle distribution where the first particle is at position $\vec{y}_2$, the second at $\vec{y}_2$ and so on. Application of the permutation operator $\hat{\mathcal{P}}_{i_1,\ldots,i_N}$ yields

$$\begin{aligned}\hat{\mathcal{P}}_{i_1,\ldots,i_N}\Psi(\vec{x}_1,\ldots,\vec{x}_N)&=\langle\vec{x}_1,\ldots,\vec{x}_N|\hat{\mathcal{P}}_{i_1,\ldots,i_N}|\Psi\rangle &(3.10)\\ &\overset{\text{Eq. 3.9}}{=}\int d^4y_1\cdots\int d^4y_N\;\underbrace{\langle\vec{x}_1,\ldots,\vec{x}_N|\vec{y}_{i_1},\ldots,\vec{y}_{i_N}\rangle}_{\delta(\vec{x}_1-\vec{y}_{i_1})\cdots\delta(\vec{x}_N-\vec{y}_{i_N})}\underbrace{\langle\vec{y}_1,\ldots,\vec{y}_N|\Psi\rangle}_{\Psi(\vec{y}_1,\ldots,\vec{y}_N)} &(3.11)\end{aligned}$$

---

[6] The rank is the number of indices that define a tensor element. Note, that there is another more limited definition of the word rank.

Let me rename the coordinates so that the δ-functions can be resolved more conveniently

$$\hat{\mathcal{P}}_{i_1,\ldots,i_N}\Psi(\vec{x}_{i_1},\ldots,\vec{x}_{i_N}) = \int d^4y_1\cdots\int d^4y_N\;\underbrace{\langle\vec{x}_{i_1},\ldots,\vec{x}_{i_N}|\vec{y}_{i_1},\ldots,\vec{y}_{i_N}\rangle}_{\delta(\vec{x}_{i_1}-\vec{y}_{i_1})\cdots\delta(\vec{x}_{i_N}-\vec{y}_{i_N})}\underbrace{\langle\vec{y}_1,\ldots,\vec{y}_N|\Psi\rangle}_{\Psi(\vec{y}_1,\ldots,\vec{y}_N)} \tag{3.12}$$

$$= \Psi(\vec{x}_1,\ldots,\vec{x}_N) \tag{3.13}$$

In the new state the first particle (first vector in the state) is at $\vec{y}_{i_1}$, that is at the position where the particle i_1 has been before the permutation.

In the following it will be more convenient to work with the adjoint permutation operator.

$$\hat{\mathcal{P}}^{\dagger}_{i_1,\ldots,i_N} \overset{\text{Eq. 3.9}}{=} \int d^4x_1\cdots\int d^4x_N\;|\vec{x}_1,\ldots,\vec{x}_N\rangle\langle\vec{x}_{i_1},\ldots,\vec{x}_{i_N}| \tag{3.14}$$

Because the permutation operator is unitary, its adjoint is at the same time its inverse. The application of the adjoint permutation operator on an arbitrary wave function yields

$$\begin{aligned}
\hat{\mathcal{P}}^{\dagger}_{i_1,\ldots,i_N}\Psi(\vec{y}_1,\ldots,\vec{y}_N) &= \langle\vec{y}_1,\ldots,\vec{y}_N|\hat{\mathcal{P}}^{\dagger}_{i_1,\ldots,i_N}|\Psi\rangle\\
&\overset{\text{Eq. 3.14}}{=} \int d^4x_1\cdots\int d^4x_N\;\underbrace{\langle\vec{y}_1,\ldots,\vec{y}_N|\vec{x}_1,\ldots,\vec{x}_N\rangle}_{\delta(\vec{y}_1-\vec{x}_{i_1})\cdots\delta(\vec{y}_N-\vec{x}_{i_N})}\underbrace{\langle\vec{x}_{i_1},\ldots,\vec{x}_{i_N}|\Psi\rangle}_{\Psi(\vec{x}_1,\ldots,\vec{x}_N)}\\
&= \Psi(\vec{y}_{i_1},\ldots,\vec{x}_{i_N})
\end{aligned}$$

3.3.3 Antisymmetrize wave functions

The exchange or permutation operator $\hat{\mathcal{P}}_{i_1,\ldots,i_N}$ applied to a fermionic wave function has the eigenvalue -1 if the sequence $i_1,\ldots,i_N$ is obtained by an odd number of permutations from regular order $1,2,3,\ldots N$, and it has the eigenvalue $+1$ for an even number of permutations. These values are those of the fully antisymmetric tensor or the Levi-Civita symbol.

EIGENVALUES OF THE PERMUTATION OPERATOR

A many-fermion state is an eigenstate of the permutation operators

$$\hat{\mathcal{P}}_{i_1,\ldots,i_N}|\Psi^F\rangle = |\Psi^F\rangle\epsilon_{i_1,\ldots,i_n} \tag{3.15}$$

with eigenvalues given by the fully antisymmetric tensor defined in section 3.3.1.
A many-boson state can be defined similarly, but the eigenvalues are $+1$ for any permutation of the particle coordinates.

$$\hat{\mathcal{P}}_{i_1,\ldots,i_N}|\Psi^B\rangle = |\Psi^B\rangle\epsilon^2_{i_1,\ldots,i_n} \tag{3.16}$$

Using Eq. 3.8 one can construct symmetry eigenstates. In a simlar way we will construct Fermionic and Bosonic wave functions from a product state. In this case, however, we do not only have a single symmetry operation with a given multiplicity, but a set of non-commutating operations that from a symmetry group. Thus, Eq. 3.8 can not be applied directly: after having symetrized the wave function with respect to one permutation, the symmetrization with respect to another permutation may destroy the first symmetrization.

`Editor: Here we need a proper derivation for symmetrized states using non-abelian symmetry groups. I have started an appendix on symmetry groups for that purpose.`

$$|\Phi^F\rangle = \sum_{i_1,\ldots,i_N=1}^{N}\hat{\mathcal{P}}^{\dagger}_{i_1,\ldots,i_N}|\chi\rangle\epsilon_{i_1,\ldots,i_N} \tag{3.17}$$

3.3.4 Slater determinants

The most simple many-particle wave function is a product state $|\phi_1,\ldots,\phi_N\rangle$, which can be written as

$$\langle\vec{x}_1,\ldots,\vec{x}_N|\phi_1,\ldots,\phi_N\rangle=\langle\vec{x}_1|\phi_1\rangle\cdots\langle\vec{x}_N|\phi_N\rangle=\phi_1(\vec{x}_1)\cdots\phi_N(\vec{x}_N) \tag{3.18}$$

A product state is not yet antisymmetric with respect to particle exchange, but with Eq. 3.8 it can be antisymmetrized. Using the N-particle permutation operator $\hat{\mathcal{P}}_{i_1,\ldots,i_N}$ and the fully antisymmetric tensor $\epsilon_{i_1,\ldots,i_N}$ we can antisymmetrize a product wave function.

$$|\Phi^F\rangle\overset{\text{Eq. 3.19}}{=}\frac{1}{\sqrt{N!}}\sum_{i_1,\ldots,i_N=1}^{N}\hat{\mathcal{P}}^\dagger_{i_1,\ldots,i_N}|\phi_1,\phi_2,\ldots,\phi_N\rangle\epsilon_{i_1,\ldots,i_N} \tag{3.19}$$

The pre-factor $\frac{1}{\sqrt{N!}}$ is the required normalization for an orthonormal set of one-particle wave functions: there are $N!$ distinct permutations of the N orbitals. Product states that differ by a permutation are orthonormal, if the one-particle states are orthonormal.

The corresponding wave function is

$$\begin{aligned}\Phi^F(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_N)&=\frac{1}{\sqrt{N!}}\sum_{i_1,\ldots,i_N}^{N}\epsilon_{i_1,\ldots,i_N}\hat{\mathcal{P}}^\dagger_{i_1,\ldots,i_N}\phi_1(\vec{x}_1)\cdots\phi_N(\vec{x}_N)\\&=\frac{1}{\sqrt{N!}}\sum_{i_1,\ldots,i_N}^{N}\epsilon_{i_1,\ldots,i_N}\phi_1(\vec{x}_{i_1})\cdots\phi_N(\vec{x}_{i_N})\end{aligned} \tag{3.20}$$

When one exploits the connection between Levi-Civita Symbol and the determinant discussed in section 3.3.1, the wave function Eq. 3.20 can also be written as a **Slater determinant**[28, 29][7]

$$\Psi^F(\vec{x}_1,\ldots,\vec{x}_N)=\frac{1}{\sqrt{N!}}\det\begin{vmatrix}\phi_1(\vec{x}_1)&\phi_1(\vec{x}_2)&\ldots&\phi_1(\vec{x}_N)\\\phi_2(\vec{x}_1)&\phi_2(\vec{x}_2)&\ldots&\phi_2(\vec{x}_N)\\\vdots&\vdots&\vdots&\vdots\\\phi_N(\vec{x}_1)&\phi_N(\vec{x}_2)&\ldots&\phi_N(\vec{x}_N)\end{vmatrix} \tag{3.21}$$

where $\epsilon_{i,j,k,\ldots}$ is the fully antisymmetric tensor defined in Sec. 3.3.1. The Slater determinant exploits the property of the determinant, that it changes its sign unter permutation of two columns of a matrix. In the Slater determinant the exchange of two columns corresponds to the exchange of two particle coordinates.

An antisymmetric wave function can be constructed even from a non-orthogonal set of one-particle orbitals by forming the corresponding determinant. However, unless the one-particle orbitals are orthonormal, the evaluation of matrix elements between Slater determinants is a nearly hopeless undertaking. Therefore, Slater determinants are usually built from an orthonormal set of one-particle orbitals.

As an example, the two-particle Slater determinant $|S_{a,b}\rangle$ of two one-particle states $\phi_a(\vec{x})$ and $\phi_b(\vec{x})$ is

$$S_{a,b}(\vec{x}_1,\vec{x}_2)=\langle\vec{x}_1,\vec{x}_2|S_{a,b}\rangle=\frac{1}{\sqrt{2}}\Big(\phi_a(\vec{x}_1)\phi_b(\vec{x}_2)-\phi_b(\vec{x}_1)\phi_a(\vec{x}_2)\Big)$$

---

[7] According to a remark on Wikipedia, Heisenberg and Dirac proposed already in 1926 to write the antisymmetric wave function in a form of a determinant.

3.3.5 Bose wave functions

Analogous to the Slater determinant we can represent symmetrized product wave functions as a **permanent**. The permanent of a matrix $[A]$ can be written as

$$\mathrm{perm}[\boldsymbol{A}] = \sum_{i_1,\dots,i_N=1}^{N} \epsilon^2_{i_1,\dots,i_N} A_{1,i_1} A_{2,i_2} \cdots A_{N,i_N}$$

We see from Eq. 3.16 that the permanent plays for bosons the same role as the determinant for Fermions.

A Bose wave function obtained by symmetrization of a product wave function has the form

$$\Psi^B(\vec{x}_1,\dots,\vec{x}_N) = \frac{1}{\sqrt{N!}} \sum_{i_1,\dots,i_N=1}^{N} \epsilon^2_{i_1,\dots,i_N} \phi_1(\vec{x}_{i_1}) \cdots \phi_{i_N}(\vec{x}_N)$$

If the one-particle states are orthonormal, also the resulting “Permanent-wave functions” built from the same one-particle basisset are orthonormal.

3.3.6 General many-fermion wave function

It is important to realize that not every antisymmetric wave function can be represented as a Slater determinant. Slater determinants are derived from product wave functions and thus have a rather restricted form.

However, the Slater determinants that can be constructed from a complete, orthonormal set of one-particle wave functions $\{|\phi_1\rangle, |\phi_1\rangle, \dots\}$ form a complete orthonormal basis for antisymmetric many-particle states. (without proof)[8]

A general N-particle state can be constructed as follows: Let us consider a complete and orthonormal one-particle basis set $|\phi_1\rangle, |\phi_2\rangle, \dots$. A subset of N such basis functions, namely $|\phi_{i_1}\rangle, |\phi_{i_2}\rangle, \dots, |\phi_{i_N}\rangle$ defines a particular Slater determinant $|S_{i_1,i_2,\dots,i_N}\rangle$. A general antisymmetric N-particle state $|\Psi\rangle$ can be written as a sum[9]. over all Slater determinants

$$|\Psi\rangle = \sum_{i_1,i_2,\dots,i_N=1; i_1<i_2<i_3\dots}^{\infty} |S_{i_1,i_2,\dots}\rangle c_{i_1,i_2,\dots}$$

where the $c_{i_1,i_2,\dots}$ are the (complex) expansion coefficients.

3.3.7 Number representation

We have learned that a Slater determinant is defined by a subset of N states from a given one-particle basis. Thus, we can form a vector with one-component for each one-particle orbital. If we set the components equal to one for the orbitals in the set and equal to zero for all orbitals not in the set, we arrive at the number representation for Slater determinants

$$|S_{i_1,\dots,i_n}\rangle = |0, 0, \underbrace{1}_{pos\,i_1}, 0, 0, 0, \underbrace{1}_{pos\,i_2}, \underbrace{1}_{pos\,i_3}, 0, \dots\rangle$$

This allows to write a general fermionic state in the form

$$|\Psi\rangle = \sum_{\sigma_1=0}^{1} \sum_{\sigma_2=0}^{1} \dots \sum_{\sigma_\infty=0}^{1} |\sigma_1, \sigma_2, \dots, \sigma_\infty\rangle c_{\sigma_1,\sigma_2,\cdots,\sigma_\infty} = \sum_{\vec{\sigma}} |\vec{\sigma}\rangle c_{\vec{\sigma}}$$

---

[8] The one-particle states need not be orthonormal to fulfill the completeness condition. The requirement of orthonormal wave functions is a convenience to make the evaluation of matrix elements for many-particle states constructed from these one-particles states efficient. These will become clear later when the Slater-Condon rules are derived.

[9] The sum goes over all N-tuples $(i_1,\dots,i_N)$ with $1 \leq i_1 < i_2 < \dots < i_N$. The indices can go up to infinity.

Chapter 4

Non-interacting electrons

Because the electrons interact strongly by their Coulomb repulsion, their description is a most difficult quantum mechanical problem.

Before we enter into the difficulties of the many-particle physics, let us for the moment ignore the interaction and study one-particle problems. At first, this sounds like intolerable approximation. However, if we consider, instead of bare electrons, quasi-particles that carry with them a distortion of the surrounding electrons, the effective interaction becomes screened. As a result, the interaction becomes sufficiently small that the picture og non-interacting quasi electrons becomes appropriate for most physical systems. This picture breaks down in a fundamental way for materials with so-called **strong correlations**. Examples of materials with strong correlations are the High-T_c superconductors, manganates which exhibit colossal magnetoresistance, heavy Fermion systems, and many more.

4.1 Dispersion relations

In solid state physics we often deal with crystals, which exhibit discrete translation symmetry. As a consequence of Bloch's theorem, this is sufficient to make the momentum a conserved quantity for non-interacting electrons. Thus we can characterize the states by their crystal momentum or wave vector.

The one-particle energy as function of the crystal momentum or the wave vector is the **dispersion relation** or the **band structure** .

quasiparticles: The dispersion relation is a very general concept in solid state physics, which applies to

- electrons, as free particles or ina material
- phonons, i.e. lattice vibrations,
- photons, i.e. light,
- holes, i.e. "missing electrons",
- excitons, i.e. electron-hole pairs,
- polariton, i.e. a photon interacting with electrons etc.,
- plasmons, i.e. charge oscillations,
- magnons, i.e. oscillations of the magnetization,
- polarons, i.e. electron-phonon pairs

- Cooper pairs, i.e. electron pairs which are responsible for superconductivity.

All these entities are called **quasiparticles**, which behave as particles with a defined energy and momentum. While we deal in solid-state physics with a small number of basic particles, electrons, nuclei with their electrodynamic interactions, the interplay between them produces a large zoo of quasiparticles, which exceeds the complexity of real particles by far.

Free particles: Let us consider a free relativistic particle, which has the dispersion relation given by $E^2 - p^2c^2 = m_0^2c^4$, that is

$$E_\pm(\vec{p}) = \pm\sqrt{(m_0c^2)^2 + (pc)^2} \tag{4.1}$$

which is shown in Fig. 4.1 alongside with the bandstructure of electrons in silicon.

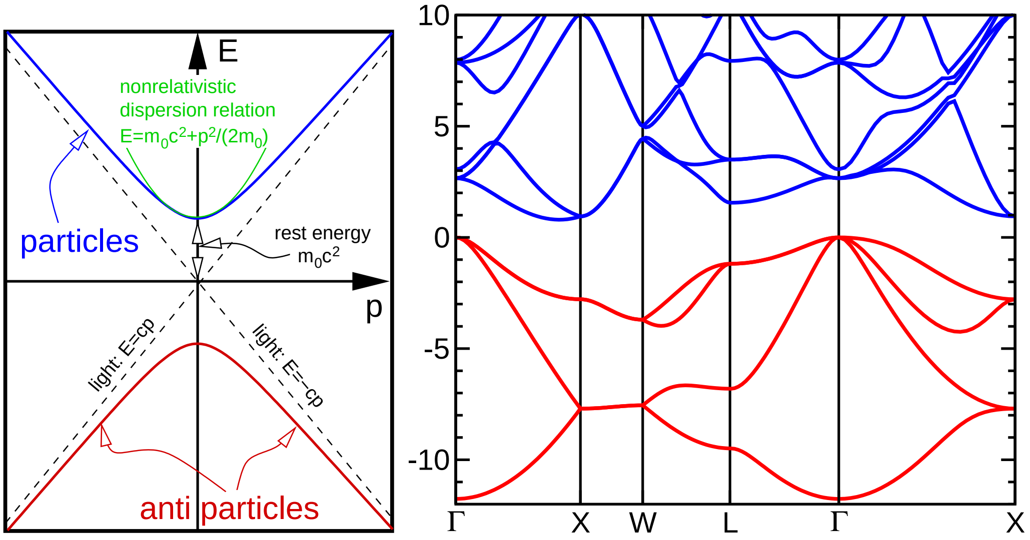


Fig. 4.1: Left: Dispersion relation of a relativsistic particle. Right: bandstructure of electrons in silicon as calculated with density-functional theory. (CP-PAW and PBE functional)

The dispersion relation of a relativistic particle has two branches, one for particles and one for antiparticles. For small momenta, the non-relativistic dispersion relation is obtained by a Taylor expansion to quadratic order. For large momenta the dispersion relation approaches a straight line with slope equal to the speed of light.

The energy as function of momentum can be considered as the **Hamilton function** $H(\vec{p}, \vec{r})$ of the corresponding particle[1]. Because of translation symmetry of the free particle the Hamilton function has no spatial dependence. We can use **Hamilton's equations**

$$\dot{\vec{r}} = \vec{v} = \vec{\nabla}_p H(\vec{p}, \vec{r}) \tag{4.2a}$$

$$\dot{\vec{p}} = \vec{F} = -\vec{\nabla}_r H(\vec{p}, \vec{r}) \tag{4.2b}$$

Thus, we can obtain the **velocity** v of the particle with a given momentum graphically as the slope of the dispersion relation. We observe that

- the particle is at rest for zero momentum, and that
- the velocity approaches the speed of light c for large momenta, but never exceeds it.

---

[1] Position and momenta are three dimensional.

Observe also that the slope in the anti-particle branch, that is the velocity, is opposite to the direction of the momentum. This is something to be taken seriously.[2]

The second Hamilton's equation is Newton's law. If the system is translationally symmetric, there are no forces and the momentum is a constant of motion, i.e. it does not change with time. However, we may also consider external forces. A constant **force** implies that the momentum changes linearly with time. By observing the change of the slope, that is the velocity, over time while the momentum is shifted to the right, we can extract the dynamics of the system under a constant force.

As an example, we may follow a skyrocket, which accelerates under the constant thrust created by its rocket engines. Initially, that is for small momenta, the dispersion relation is quadratic, so that the velocity changes linear with momentum. Under a constant force, the velocity increases linearly with time with an acceleration a obtained from $\vec{F} = m_0\vec{a}$. The curvature of the dispersion relation is the inverse **mass**. As the rocket continues to accelerate, its speed grows less and less, while the momentum still grows continuously. This captures that the speed of light is the absolute maximum velocity for particles.

The dispersion relation of the relativistic particle has two branches. The upper one is identified with particles and the other is identified with antiparticles. In order to avoid the problem, that an antiparticle could accelerate indefinitely while emitting light, the concept of the **Dirac sea** has been introduced. The problem finds a more satisfactory solution in the context of quantm field theory. However, here we employ it, because it is analogous to the conventional view in solid state physics. The idea is that all states of the antiparticle band are filled with electrons under normal circumstances. Due to Pauli principle a fermionic particle cannot lower its energy by occupying an already filled state of the antiparticle band and thus emit light. Thus an antiparticle is a "missing electron in the antiparticle band" or a hole.

A particle anti-particle pair can annihilate by emitting to quanta of light, which carry away energy and momentum of the particle antiparticle pair. This is an example of the interplay of two band structures, one for electrons and one for photons, i.e. light. Their only difference is the mass which vanishes for photons. Let us consider the process in detail. Initially there is a particle, say with momentum zero and a "hole" in the antiparticle branch. The particle can drop in the hole. In that process the rest energy $2m_0c^2$ of both particles and the momentum $p' = 0$ needs to be carried away. Therefore we turn to the bandstructure for photons, that is for mass-less particles. We observe that the energy cannot be placed into a single photon, because that would also absorb a finite momentum. However it can be placed into two photons with opposite momentum and each having an energy equal to the rest mass of the annihilated particle. This is an example for a scattering process of particles as they are present in solids.

Real band structure: The interpretation of the band structure of a real material is analogous to that of the dispersion relation of the free particle. Therefore, the band structure of electrons in silicon are shown in Fig. 4.1 alongside the dispersion relation of the free particle.

There are two apparent difference to the free particle:

- there are even more more branches
- the momenta are drawn into several finite sections

The reason is the reduced symmetry of the solid compared to a particle in vacuum. The dispersion relation of a free particle is isotropic, so that it is suffient to draw it in one direction. In a crystal, the bands depend on the direction of the momentum.

Because the translational symmetry of a lattice is discrete, the coordinate axis for the momentum is folded up into a finite region. The resulting section of momentum space is thus finite. Furthermore the folded bands add to the bands already present in this finite region. Each segment describes the band structure along a so-called high-symmetry line in momentum space.

Now, we can translate the observations made for the free particle to the band structure of silicon:

---

[2] However, the interpretation of antiparticles can be changed so that the energies are positive.

- The bands at negative energies shown in red in Fig. 4.1 describe the valence electrons. These bands are filled with electrons as the antiparticle bands in the Dirac sea.
- The bands at positive energies shown in blue in Fig. 4.1 describe the conduction electrons.
- Valence and conduction states are separated by a band gap, a region where no electrons can exist in a perfect lattice.
- By heating up the crystal, a small number of electrons are created in the conduction band and a small number of holes, i.e. missing electrons, are formed in the valence band. In the following we often simply refer to electrons and **holes**. The electrons will relax towards the minima of the conduction band, near the point X.
- The curvature of the energy band at the minimum is related to the inverse **effective mass** of the conduction electron.

  $$\begin{aligned}\epsilon(\vec{p}) &\approx \frac{1}{2}\vec{p}\,\boldsymbol{m}_{\text{eff}}^{-1}\,\vec{p} \\ \boldsymbol{m}_{eff} &= \hbar^2\Big(\vec{\nabla}_k \otimes \vec{\nabla}_k\big|_{\vec{k}_0}\,\epsilon\Big)^{-1} \qquad (4.3)\end{aligned}$$

  The derivative is taken at the band extremum and for the band that forms it. This effective mass is anisotropic, which can be seen from the curvature in $\Gamma - X$ direction and the perpendicular direction $X - W$. The holes are located at the maximum of the valence band, which lies at Γ. Flat bands indicate that the particle is heavy, while bands with large curvature indicate light particles. The effective masses of holes are obtained analytically, with the only difference that the sign in the definition changes.
- For metals, the slope of the bands at the Fermi level is a measure for the velocity of n electron, the **Fermi velocity**.

  $$\vec{v}^{\text{Fermi}}(\vec{k}) = \frac{1}{\hbar}\vec{\nabla}_k\epsilon(\vec{k}) \qquad (4.4)$$

  The Fermi velocity depends strongly on the partular band and the position at the Fermi surface. Typical Fermi velocity are $10^6\ m/sec$, which is only a factor 300 slower than the speed of light. [3]
- The number of thermally created holes and electrons and their effective masses determine the **electrical conductivity**.
- A photon can be absorbed by the formation of electron-hole pairs, if its energy exceeds the band gap. Silicon is transparent to infrared radiation, but it reflects visible light. The band gap of silicon is about $E_g = 1.17$ eV, while the photon energy of visible light ranges from 1.65 to 3.26 eV. Because of the large speed of light, a photon has nearly infinite slope. Thus a excitation or recombination is a **vertical process**.
- A lattice vibration can also be represented by a band structure. Opposite to the photon, the phonon has small energies and can have large momenta. This is consisten with the small speed of sound. Thus, the absorption of a phonon shifts the particle nearly horizontally in the band structure.

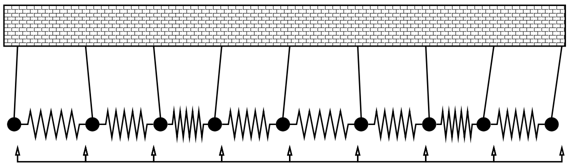
---

[3] Consider a free electron gas with dispersion relation $\epsilon(\vec{p}) = \frac{\vec{p}^2}{2m_e}$ As shown below the particle number N in a volume V is given by $N = V\sum_\sigma \int_{|\vec{k}|<k_F} \frac{d^3k}{(2\pi)^2} = \frac{1}{3\pi^2}k_F^3 V$, so that $k_F = \sqrt[3]{3\pi^3\frac{N}{V}}$. The velocity is $v_F = (\hbar|k_F|)/m_e = \frac{\hbar}{m_e}\sqrt[3]{3\pi^3\frac{N_e}{V}}$. Let us estimate the quantities: Aluminium has a lattice constant of $a_{lat} = 4.05$ Å. Aluminium has the face centered cubic structure, with 4 atoms in the cubic unit cell. Each Al atom contributes three valence electrons. Thus the electron density is $N/V = 12a_{lat}^{-3} = 0.18 \times 10^{30}$ m^{-3} With the electron mass of 0.911×10^{-30} kg. Thus we obtain

$$v_F = \frac{\hbar}{m_e}\sqrt[3]{3\pi^3\frac{N_e}{V}} \approx 2 \times 10^6\ \frac{m}{s} \qquad (4.5)$$

Correspondence principle: So far we have argued on the basis of classical particles. However, wave packets as in quantum mechanics behave completely analogous[4]. For waves, the dispersion relation is usually drawn in terms of frequency $\omega(\vec{k})$ as function of the wave vector $\vec{k}$.

An example, which is often useful is the hanging linear chain. The hanging linear chain is a set of pendula coupled by springs. In the continuous limit, the hanging linear chain has the same dispersion relation as the relativistic particle.



The dispersion relation $\omega(\vec{k})$ has the same interpretation as $E(\vec{p})$:

- the **group velocity** $\vec{v}_g$ of a wave packet is given as the slope of the dispersion relation

$$\vec{v}_g = \vec{\nabla}_k \omega(\vec{k}) \tag{4.6}$$

- for a system that does not change with time, its frequency remains constant over time. If we now consider a dispersion relation $\omega(\vec{k},\vec{r})$ with a weak spatial dependence, we can investigate the motion of a wave packet centered at $\vec{r}(t)$ having a wave vector $\vec{k}(t)$

$$0 = \frac{d}{dt}\omega\Big(\vec{k}(t),\vec{r}(t)\Big) = \dot{\vec{k}}\underbrace{\vec{\nabla}_k\omega(\vec{k},\vec{r})}_{\dot{\vec{r}}} + \dot{\vec{r}}\vec{\nabla}_r\omega(\vec{k},\vec{r}) = \dot{\vec{r}}\Big(\dot{\vec{k}} + \vec{\nabla}_r\omega(\vec{k},\vec{r})\Big)$$

$$\Rightarrow \quad \dot{\vec{k}} = -\vec{\nabla}_r\omega(\vec{k},\vec{r}) \tag{4.7}$$

Tus we arrive at equations of motion for a wave packet that is completely analogous to Hamilton's equations for particles.

$$\dot{\vec{r}} = \vec{\nabla}_k\omega(\vec{k}) \tag{4.8a}$$

$$\dot{\vec{k}} = -\vec{\nabla}_r\omega(\vec{k},\vec{r}) \tag{4.8b}$$

This is the basis for the **correspondence principle**, which says that localized wave packets behave like classical particles with energy and momentum given by

$$E = \hbar\omega \qquad \text{and} \qquad \vec{p} = \hbar\vec{k} \tag{4.9a}$$

The constant $\hbar$ is at first an arbitrary constant. Experiments determine it for quantuym particles to be equal to Plancks quantum. The correspondence principle, i.e. the analogy of particles and wave packets can be considered as one of the main pillars for the formulation of quantum physics.

4.2 Crystal lattices

In solid state physics, one deals mostly with materials that are periodic in space, so-called crystals. Crystals exhibit a discrete translational symmetry

$$V(\vec{r}+\vec{t}) = V(\vec{r}) \tag{4.10}$$

---

[4] See Chapter 3 of *"ΦSX: Quantum Physics."*[27]

where the translation vectors $\vec{t}$ form a regular lattice. More specifically, we should require for the one-particle Hamilton operator acting on the electrons

$$\langle \vec{r}+\vec{t}|\hat{H}|\vec{r}+\vec{t}\rangle = \langle \vec{r}|\hat{H}|\vec{r}\rangle \tag{4.11}$$

In addition to the discrete translation symmetry, a crystal has also a **point-group symmetry**, which allows to divide lattices into 230 **crystallographic space groups** . A set of useful tools for crystallography can be found on the Bilbao crystallograhic server `http://www.cryst.ehu.es`.

Besides crystals, there are also solids without translational symmetry. Among them are quasicrystals, amorphous material and glasses. Even crystals are rarely perfect.

They usually have

- **point defects** such as substitutional defects (foreign atoms), vacancies (missing atoms), interstitials (additional atoms)

- **line defects** that are so-called **dislocations**

- **planar defects** such as **grain boundaries**, twin boundaries and interfaces

- other defects such as **precipitates**, i.e. small crystallites within another material.

4.2.1 Reciprocal space revisited

From ΦSX: *Quantum mechanics of the chemical bond*[30]. Shortened version of section *"Real and reciprocal lattice in three dimensions"*

Let me explain my notation and remind you of the some of the concepts related to real and reciprocal lattices.

Fig. 4.2.1 demonstrates the concepts developed in the following. Corresponding to the three spatial directions there are now three primitive lattice vectors, which we denote by $\vec{T}_1, \vec{T}_2, \vec{T}_3$. Note that in two or three dimensions there is no unique choice of primitive lattice vectors. The three primitive lattice vectors span the **primitive unit cell**.

A general lattice vector $\vec{t}_{i,j,k}$ can be expressed by the **primitive lattice vectors** $\vec{T}_1, \vec{T}_2, \vec{T}_3$ as

$$\vec{t}_{i,j,k} = i\vec{T}_1 + j\vec{T}_2 + k\vec{T}_3 \hat{=} \underbrace{\begin{pmatrix} T_{x,1} & T_{x,2} & T_{x,3} \\ T_{y,1} & T_{y,2} & T_{y,3} \\ T_{z,1} & T_{z,2} & T_{z,3} \end{pmatrix}}_{\boldsymbol{T}} \begin{pmatrix} i \\ j \\ k \end{pmatrix}$$

where i, j, k are arbitrary integers.

It is often convenient to combine the lattice vectors into a 3×3 matrix $\boldsymbol{T}$ as shown above. Often the atomic positions $\vec{R}_n$ are provided in **relative positions** $\vec{X}$ which are defined as

$$\vec{R}_n = \boldsymbol{T}\vec{X}_n \qquad \Leftrightarrow \qquad \vec{X}_n = \boldsymbol{T}^{-1}\vec{R}_n$$

A potential is called periodic with respect to these lattice translations, if $V(\vec{r}+\vec{t}_{i,j,k}) = V(\vec{r})$.

RECIPROCAL LATTICE

The reciprocal lattice is given by those values of the wave vector $\vec{G}$, for which the corresponding plane waves $e^{i\vec{G}\vec{r}}$ have the same periodicity as the real-space lattice.
The primitive reciprocal-lattice vectors $\vec{g}_n$ for $n = 1, 2, 3$ are defined in three dimensions as

$$\vec{g}_n \vec{T}_m = 2\pi\delta_{n,m} \tag{4.12}$$

The reciprocal lattice vectors can be obtained as

$$\vec{g}_1 = 2\pi \frac{\vec{T}_2 \times \vec{T}_3}{\vec{T}_1\left(\vec{T}_2 \times \vec{T}_3\right)} \quad , \quad \vec{g}_2 = 2\pi \frac{\vec{T}_3 \times \vec{T}_1}{\vec{T}_1\left(\vec{T}_2 \times \vec{T}_3\right)} \quad , \quad \vec{g}_3 = 2\pi \frac{\vec{T}_1 \times \vec{T}_2}{\vec{T}_1\left(\vec{T}_2 \times \vec{T}_3\right)}$$

The general reciprocal lattice vectors $\vec{G}_{i,j,k}$ are multiples of the primitive lattice vectors $\vec{g}_1$, $\vec{g}_2$, $\vec{g}_3$.

$$\vec{G}_{i,j,k} = i\vec{g}_1 + j\vec{g}_2 + k\vec{g}_3$$

- The reciprocal lattice vectors are perpendicular to the lattice planes of the real-space lattice.

$$\vec{g}_1 || \vec{t}_2 \times \vec{t}_3 \qquad \text{and} \qquad \vec{g}_2 || \vec{t}_3 \times \vec{t}_1 \qquad \text{and} \qquad \vec{g}_3 || \vec{t}_1 \times \vec{t}_2$$

- The length of the *primitive* reciprocal lattice vectors is 2π divided by the distance Δ_n of the lattice planes.

$$|\vec{g}_n| = \frac{2\pi}{\Delta_n}$$

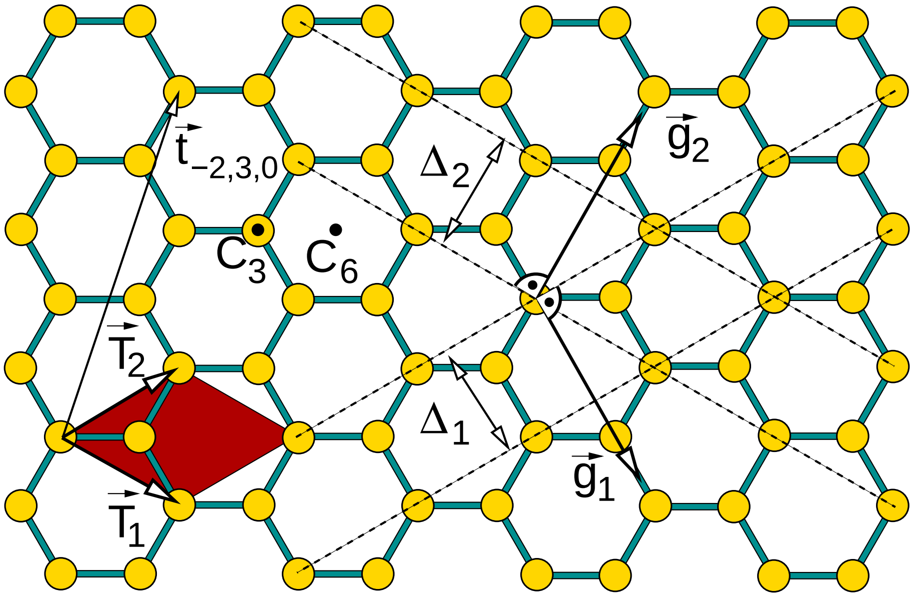


Fig. 4.2: Translational and rotational symmetry of a single graphite sheet and demonstration of reciprocal lattice vectors. Graphite is a layered material, which consists of sheets as the one shown. The yellow balls indicate the carbon atoms and the blue sticks represent the bonds. In graphite these two-dimensional sheets are stacked on top of each other, where every second sheet is shifted such that an atom falls vertically below the point indicated by C_6. Here, we only consider the symmetry of a single sheet. The elementary unit cell is indicated by the red parallelogram, which is spanned by the elementary lattice vectors $\vec{T}_1$ and $\vec{T}_2$. The third lattice vector points perpendicular to the sheet towards you. An example for a general lattice vector is $\vec{t}_{3,-2,0} = 3\vec{T}_1 - 2\vec{T}_2 + 0\vec{T}_3$. The lattice planes indicated by the dashed lines. The lattice planes are perpendicular to the sheet. The distance of the lattice planes are indicated by Δ_1 and Δ_2. The elementary reciprocal lattice vectors are $\vec{g}_1$ and $\vec{g}_2$. The third reciprocal lattice vector points perpendicular out of the plane towards you. Note that the reciprocal lattice vectors have the unit "inverse length". Thus, their length is considered irrelevant in this real space figure. The axis through C_3 standing perpendicular to the plane is a three-fold rotation axis of the graphite crystal. The axis through C_6 perpendicular to the plane is a 6-fold rotation axis of the graphite sheet, but only a three-fold rotation axis of the graphite crystal. (Note that a rotation axis for the graphite crystal must be one for both sheets of the crystal). In addition there is a mirror plane lying in the plane. Furthermore there are several mirror planes perpendicular to the plane: One passing through every atom with one of the three bonds lying in the plane and one perpendicular each bond passing through the bond center.

4.2.2 Bloch Theorem

From section 10.5.3 of Φ*SX: Quantum Theory*[27].

Symmetry: For every lattice-translation vector $\vec{t}$, we define a lattice-translation operator

$$\hat{S}(\vec{t}) \stackrel{\text{def}}{=} \int d^3r \; |\vec{r}\rangle \langle \vec{r} - \vec{t}| \tag{4.13}$$

which satisfies

$$\langle \vec{r}|\hat{S}(\vec{t})|\psi\rangle = \langle \vec{r} - \vec{t}|\psi\rangle \; . \tag{4.14}$$

Let us investigate the eigenstates of the translation operator. Because the lattice-translation operator is unitary, the absolute value of the eigenvalue is one, that is

$$\hat{S}(\vec{t})|\psi_\varphi\rangle = |\psi_\varphi\rangle \mathrm{e}^{i\varphi} \tag{4.15}$$

where the phase shift depends on the translation vector.

Because $\hat{S}(n\vec{t}) = \hat{S}^n(\vec{t})$, a similar relation must hold for the eigenvalues, i.e. $\varphi(n\vec{t}) = n\varphi(\vec{t})$. This requirement can be satisifed by choosing the form $\varphi(\vec{t}) = \vec{k}\vec{t}$. In addition, this form works for superpositions of linearly indepdendent lattice vectors. The vector-parameter $\vec{k}$ is the so-called **Bloch vector** or simply wave vector.

Thus, the eigenvalue equation for the lattice-translation operator is

$$\hat{S}(\vec{t})|\phi_{\vec{k}}\rangle = |\phi_{\vec{k}}\rangle \mathrm{e}^{i\vec{k}\vec{t}} \tag{4.16}$$

Let us now construct the eigenstates of the lattice translation operator:

$$\langle\vec{r}|\hat{S}(\vec{t})|\phi_{\vec{k}}\rangle = \langle\vec{r}|\phi_{\vec{k}}\rangle \mathrm{e}^{i\vec{k}\vec{t}} \tag{4.17}$$

$$\langle\vec{r}-\vec{t}|\phi_{\vec{k}}\rangle = \langle\vec{r}|\phi_{\vec{k}}\rangle \mathrm{e}^{i\vec{k}\vec{t}}$$

$$\Rightarrow \underbrace{\langle\vec{r}-\vec{t}|\phi_{\vec{k}}\rangle\rangle \mathrm{e}^{-i\vec{k}\vec{t}}}_{u_{\vec{k}}(\vec{r}-\vec{t})} = \underbrace{\langle\vec{r}|\phi_{\vec{k}}\rangle\rangle\, \underbrace{\mathrm{e}^{-i\vec{k}\vec{0}}}_{=1}}_{u_{\vec{k}}(\vec{r})} \tag{4.18}$$

Thus, the function

$$u_{\vec{k}}(\vec{r}) \stackrel{\text{def}}{=} |\phi_{\vec{k}}\rangle \mathrm{e}^{-i\vec{k}\vec{r}} \tag{4.19}$$

which is defined for an eigenstate of the lattice translation operator $\hat{S}(\vec{t})$ with eigenvalue $\mathrm{e}^{i\vec{k}\vec{t}}$ is invariant with the lattice translations. That is

$$u_{\vec{k}}(\vec{r}-\vec{t}) = u_{\vec{k}}(\vec{r})\,. \tag{4.20}$$

Hence the eigenstates of the lattice-translation operators can be witten as product of a periodic function and a plane wave. Without limitation of generality, this Bloch vector can be chosen from within the first reciprocal unit cell.

Lattice symmetry implies that the eigenstates of the lattice Hamiltonian can be chosen to be eigenstates of the lattice-translation operator, which immediately implies Bloch theorem.

BLOCH THEOREM

The eigenstates of a system with lattice symmetry can be written as product of a periodic function $u_{\vec{k}}(\vec{r})$, i.e.

$$u_k(\vec{r}+\vec{t}) = u_{\vec{k}}(\vec{r})\,,\text{‘} \tag{4.21}$$

and a modulating plane wave $\mathrm{e}^{i\vec{k}\vec{r}}$, that is

$$\psi_k(\vec{r}) = u_{\vec{k}}(\vec{r})\mathrm{e}^{i\vec{k}\vec{r}} \tag{4.22}$$

where the wave vector $\vec{k}$, the Bloch vector, is a good quantum number of the system.

4.2.3 Non-interacting electrons in an external potential

Another way to obtain the Bloch theorem is to set up the Hamiltonian in a plane wave basis. On the one hand, this approach to the problem connects to some of the principles used in a plane wave pseudopotential method. On the other hand it shows that lattice potential leads to a coupling between points in the reciprocal space, which are separated by reciprocal lattice vectors.

Let us now consider the consequences of a weak external potential.

$$V(\vec{r}) = \sum_{\vec{G}} V_{\vec{G}} \mathrm{e}^{i\vec{G}\vec{r}} \tag{4.23}$$

Since the potential shall be periodic, only reciprocal lattice vectors $\vec{G}$ occur in the sum. The Hamiltonian can be written as

$$\hat{H} = \sum_{\sigma} \int d^3r\, |\vec{r}, \sigma\rangle \left(\frac{-\hbar^2}{2m_e} \vec{\nabla}^2 + \sum_{G} V_G \mathrm{e}^{i\vec{G}\vec{r}} \right) \langle \vec{r}, \sigma|$$

Let us now work out the matrix elements of the Hamiltonian between two plane waves,

$$\phi_{\vec{k},\sigma}(\vec{r}, \sigma') \overset{\text{Eq. 4.91}}{=} \langle \vec{r}, \sigma' | \phi_{\vec{k},\sigma}\rangle = \frac{\delta_{\sigma,\sigma'}}{\sqrt{\Omega}} \mathrm{e}^{i\vec{k}\vec{r}}, \tag{4.24}$$

which are the eigenstates of the free electron gas.

$$\begin{aligned}
H_{\sigma,\sigma'}(\vec{k}, \vec{k'}) &:= \langle \phi_{\vec{k},\sigma} | \hat{H} | \phi_{\vec{k'},\sigma'} \rangle \\
&= \sum_{\sigma''} \int_{\Omega} d^3r\, \underbrace{\langle \phi_{\vec{k},\sigma} | \vec{r}, \sigma'' \rangle}_{\phi^*_{\vec{k},\sigma}(\vec{r},\sigma'')} \left(\frac{-\hbar^2}{2m_e} \vec{\nabla}^2 + \sum_{G} V_G \mathrm{e}^{i\vec{G}\vec{r}} \right) \underbrace{\langle \vec{r}, \sigma'' | \phi_{\vec{k'},\sigma'} \rangle}_{\phi_{\vec{k},\sigma}(\vec{r},\sigma'')} \\
&\overset{Eq.\ 4.91}{=} \frac{\delta_{\sigma,\sigma'}}{\Omega} \int_{\Omega} d^3r\, \mathrm{e}^{-i\vec{k}\vec{r}} \left(\frac{-\hbar^2}{2m_e} \vec{\nabla}^2 + \sum_{G} V_G \mathrm{e}^{i\vec{G}\vec{r}} \right) \mathrm{e}^{i\vec{k'}\vec{r}} \\
&= \frac{(\hbar\vec{k})^2}{2m_e} \delta_{\sigma,\sigma'} \underbrace{\frac{1}{\Omega} \int_{\Omega} d^3r\, \mathrm{e}^{i(\vec{k'}-\vec{k})\vec{r}}}_{\delta_{\vec{k'},\vec{k}}} + \sum_{\vec{G}} V_{\vec{G}} \delta_{\sigma,\sigma'} \underbrace{\frac{1}{\Omega} \int_{\Omega} d^3r\, \mathrm{e}^{-i\vec{k}\vec{r}} \mathrm{e}^{i\vec{G}\vec{r}} \mathrm{e}^{i\vec{k'}\vec{r}}}_{\delta_{\vec{k'}+\vec{G},\vec{k}}} \\
&= \delta_{\sigma,\sigma'} \sum_{\vec{G}} \left[\frac{(\hbar\vec{k})^2}{2m_e} \delta_{\vec{G},\vec{0}} + V_{\vec{G}} \right] \delta_{\vec{k'}+\vec{G},\vec{k}}
\end{aligned} \tag{4.25}$$

We have used here the normalization condition[5] corresponding to a finite, even though very large, volume Ω with side-lengths L_x, L_y, L_z . Correspondingly, the spectrum is discrete with wave vectors $\vec{k}_{i,j,k} = \frac{2\pi}{L_x} i + \frac{2\pi}{L_y} j + \frac{2\pi}{L_z} k$.

Thus, there is only a coupling between states that are separated by a reciprocal lattice vector. This is one way to prove **Bloch's theorem**.

BLOCH'S THEOREM

Bloch's theorem says that the eigenstates in a periodic potential can be written as a product of a function $u_{\vec{k},\sigma}(\vec{r}, \sigma')$ that is periodic with the periodicity of the potential and a phase factor $\mathrm{e}^{i\vec{k}\vec{r}}$.

This can be shown as follows: Once we diagonalize the Hamiltonian, the eigenstates will be a superposition of the basis states, the plane waves, that are connected by a Hamiltonian matrix element. Thus, the final wave function has the general form

$$\psi_{\vec{k}}(\vec{r}) = \sum_{\vec{G}} \mathrm{e}^{i(\vec{k}+\vec{G})\vec{r}} c_{k,G} = \underbrace{\left[\sum_{\vec{G}} \mathrm{e}^{i\vec{G}\vec{r}} c_{k,G} \right]}_{u_{\vec{k}}(\vec{r})} \mathrm{e}^{i\vec{k}\vec{r}} = u_{\vec{k}}(\vec{r}) \mathrm{e}^{i\vec{k}\vec{r}} \tag{4.26}$$

---

[5] Other choices would normalize the wave function within the unit cell, or they would set the average density equal to one.

Because the function $u_k(\vec{r})$ has only Fourier components at the reciprocal lattice vectors, it is periodic with the real space lattice symmetry.

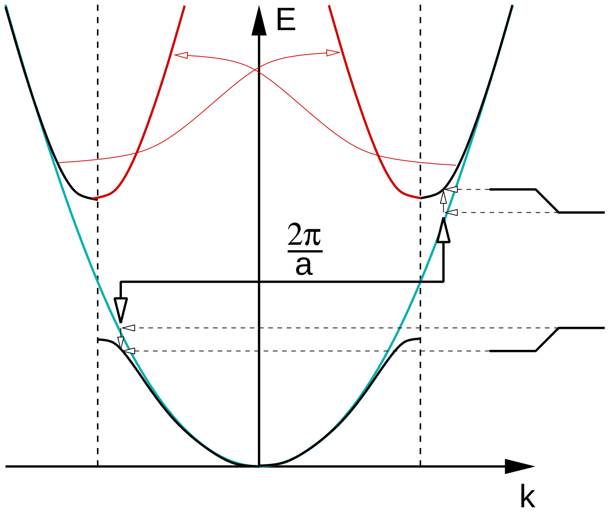


Fig. 4.3: Schematic demonstration how a periodic potential results in a coupling between states with the same wave vector in the reduced zone scheme

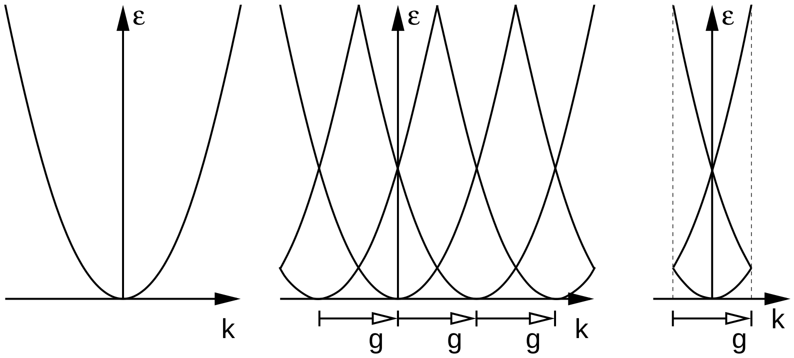


Fig. 4.4: Extended, periodic and reduced zone scheme for a one-dimensional free particle.

This is the reason why band structures are not represented in an **extended zone scheme** but in a **reduced zone scheme**.

- in the extended zone scheme the band structure of free electrons would be a single parabola, and would extend to infinity in reciprocal space.

- in the **periodic zone scheme**, the free electron would consists of many parabola centered at the reciprocal lattice points. Thus, the reciprocal zone scheme is periodic with the periodicity of the reciprocal lattice.

- in the reduced zone scheme, only the irreducible part of the periodic zone scheme is shown, that is one single repeat unit. The band structures shown in Fig. 4.9 are in a reduced zone scheme. There are several choices for the repeat units of the reciprocal lattice. One choice is simply the primitive unit cell of the reciprocal lattice. However, the shape of the unit cell does not have the point-group symmetry of the reciprocal lattice. Therefore one instead chooses the **Wigner Seitz cell**[6] of the reciprocal lattice, which is called the **Brillouin zone**.

In the reduced zone scheme, all wave vectors connected by a reciprocal lattice vector, fall on top of the same point in the Brillouin zone. Thus, a Hamiltonian that has lattice periodicity couples all points that lie at the same point in the Brillouin zone. This is demonstrated in Fig. 4.3.

From perturbation theory we know that the splitting of two interacting states is large when the two states are close or even degenerate. Thus the largest effect in the band structure occurs right at the boundary of the irreducible zone, where the degeneracy of the free electron gas is lifted by the periodic potential. Thus, local band gaps appear at the surface of the Brillouin zone. This band gap, however, is usually warped, so that there is no energy window, that completely lies in all local band gaps. In that case the material remains a metal.

4.3 Bands and orbitals

From ΦSX:Quantum mechanics of the chemical bond[30]

So far we discussed the band structures as a variation of the dispersion relation of the free-electron bands. However, this point of view obscures the relation to chemical binding and the local picture of the wave functions.

The band structure can also be constructed starting from local orbitals. This is demonstrated here for the two-dimensional square lattice with atoms having s- and p-orbitals in the plane.

In order to simplify the work, we restrict ourselves to the high symmetry points Γ, X and M of the two-dimensional reciprocal unit cell.

The real space lattice vectors are

$$\vec{T}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} a \qquad \text{and} \qquad \vec{T}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} a$$

The corresponding reciprocal space lattice vectors are

$$\vec{g}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \frac{2\pi}{a} \qquad \text{and} \qquad \vec{g}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \frac{2\pi}{a}$$

The high-symmetry points are

$$k_\Gamma = 0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \frac{2\pi}{a} \quad ; \quad k_X = \frac{1}{2}\vec{g}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \frac{\pi}{a} \quad \text{and} \quad k_M = \frac{1}{2}\left(\vec{g}_1 + \vec{g}_2\right) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \frac{\pi}{a}$$

The first step is to construct Bloch waves out of the atomic orbitals, by multiplying them with $e^{i\vec{k}\vec{t}}$, where $\vec{t}$ is the lattice translation of the atom relative to the original unit cell. This is illustrated in Fig. 4.5 on page 69.

- At the Γ, the phase factor for a lattice translation of the orbital is one. Thus, we repeat the orbital from the central unit cell in all other unit cells.
- At the X point, the phase factor for a lattice translation along the first lattice vector is $e^{i\vec{k}_X\vec{T}_1} = e^{i\pi} = -1$ and the phase factor for the lattice translation along the second lattice vector $\vec{T}_2$ is $e^{i\vec{k}_X\vec{T}_2} = +1$. Thus, if we go to the right the sign of the orbital alternates, while in the vertical direction the orbital remains the same.

---

[6] The Wigner Seitz cell of a lattice consists of all points that are closer to the origin than to any other lattice point. Thus, it is enclosed by planes perpendicular to the lattice vectors cutting the lattice vector in half.

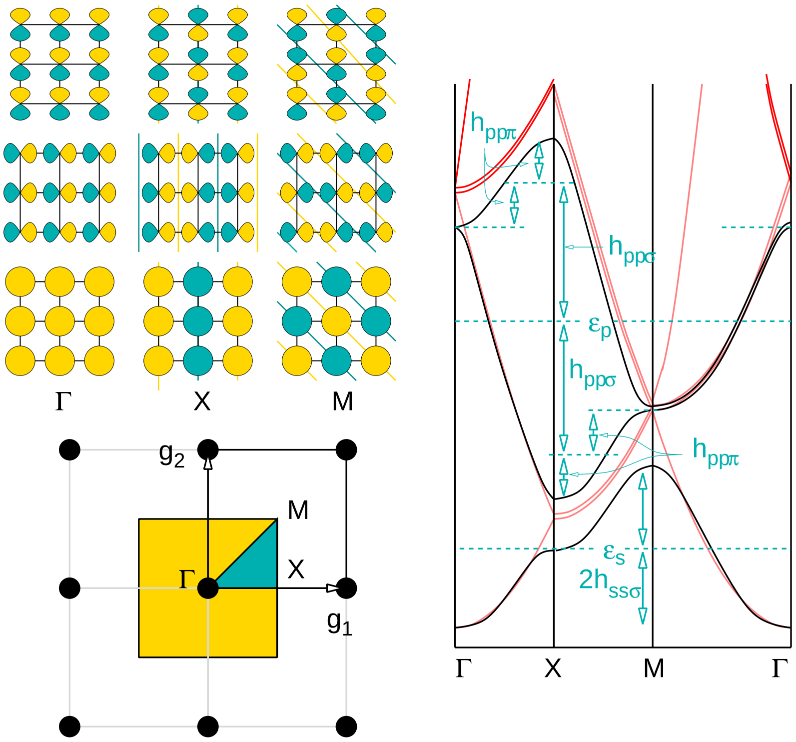


Fig. 4.5: Illustration of a band structure in real and reciprocal space for a planar square lattice. Bottom left: The reciprocal lattice is spanned by the reciprocal lattice vectors $\vec{g}_1$ and $\vec{g}_2$.The yellow square is the corresponding Brillouin zone. The green inset is the Brillouin zone, whose corners are the high symmetry points Γ, X and M. On the top left, the corresponding basis functions made from one s-orbital and two p-orbitals in a Bloch-basis are shown. These states are also eigenstates of the Hamiltonian, because the coupling vanishes due to point group symmetry at the high-symmetry points. On the top right, a schematic band structure is shown and compared to a free-electron band (red).

- At the M point, the phase factor alternates in each lattice direction, resulting in a checkerboard pattern.

Next we need to set up the Hamilton matrix elements for each k-point, using the Slater-Koster parameters. For a back-on-the envelope estimate the nearest neighbor matrix elements will be sufficient. The matrix elements are firstly ϵ_s and ϵ_p, the energies of the orbitals without any hybridization. Secondly we need the hopping matrix elements $h_{ss\sigma}$, $h_{pp\sigma}$, $h_{pp\pi}$. In principle we would also need $h_{sp\sigma}$, but that matrix element will not be needed in our example.

In our simple example, the Bloch states are already eigenstates of the Hamiltonian so that we only need to calculate their energy. We obtain[7]

Γ	X	M
$\epsilon_p + \|h_{pp\sigma}\| - \|h_{pp\pi}\|$	$\epsilon_p + \|h_{pp\sigma}\| + \|h_{pp\pi}\|$	$\epsilon_p - \|h_{pp\sigma}\| + \|h_{pp\pi}\|$
$\epsilon_p + \|h_{pp\sigma}\| - \|h_{pp\pi}\|$	$\epsilon_p - \|h_{pp\sigma}\| - \|h_{pp\pi}\|$	$\epsilon_p - \|h_{pp\sigma}\| + \|h_{pp\pi}\|$
$\epsilon_s - 2\|h_{ss\sigma}\|$	ϵ_s	$\epsilon_s + 2\|h_{ss\sigma}\|$

---

[7] Note, that the hopping matrix elements have the opposite sign than the overlap matrix elements. Thus, a positive overlap leads to a lowering of the energy level of the bond. In order to avoid confusion, I am introducing explicitely the absolute value of the hopping matrix elements, which have the same sign as the corresponding overlap matrix elements.

If there is a band structure available, we can estimate the Hamilton matrix elements from that. Otherwise we just make an educated guess.

4.4 Calculating band structures in the tight-binding model

From ΦSX:Quantum mechanics of the chemical bond[30]

Now we wish to calculate the band structure. Let us begin with a tight-binding basisset with the usual assumption that the tight-binding orbitals shall be orthogonal. The tight-binding orbitals shall be represented by kets $|\vec{t},\alpha\rangle$, where $\vec{t}$ denotes a discrete lattice-translation vector and α denotes the type of the orbital such as $s, p_x, p_y, p_z, \ldots$ and a spin quantum number, as well as a site index R of an atom in the first unit cell of the lattice.

A Bloch wave can be represented as

$$
\begin{aligned}
|\Psi_{\vec{k},n}\rangle &= \sum_{\vec{t},\alpha} |\vec{t},\alpha\rangle\langle\vec{t},\alpha|\Psi_{\vec{k},n}\rangle \\
&\overset{\text{Eq. 4.22}}{=} \sum_{\vec{t},\alpha} |\vec{t},\alpha\rangle\langle\vec{t},\alpha| \int d^4x\; |\vec{x}\rangle\; \underbrace{\mathrm{e}^{i\vec{k}\vec{r}}\langle\vec{x}|u_{\vec{k},n}\rangle}_{\psi_{\vec{k}}(\vec{x})=u_{\vec{k}}(\vec{x})\mathrm{e}^{i\vec{k}\vec{r}}} \\
&= \sum_{\vec{t},\alpha} |\vec{t},\alpha\rangle\mathrm{e}^{i\vec{k}\vec{t}}\langle\vec{t},\alpha| \int d^4x\; |\vec{x}\rangle\mathrm{e}^{i\vec{k}(\vec{r}-\vec{t})}\langle\vec{x}|u_{\vec{k},n}\rangle \\
&\overset{u(\vec{x})=u(\vec{r}-\vec{t},\sigma)}{=} \sum_{\vec{t},\alpha} |\vec{t},\alpha\rangle\mathrm{e}^{i\vec{k}\vec{t}}\langle\vec{t},\alpha| \sum_{\sigma} \int d^3r\; |\vec{r},\sigma\rangle\mathrm{e}^{i\vec{k}(\vec{r}-\vec{t})}\langle\vec{r}-\vec{t},\sigma|u_{\vec{k},n}\rangle \\
&\overset{\vec{r}\to\vec{r}+\vec{t}}{=} \sum_{\vec{t},\alpha} |\vec{t},\alpha\rangle\mathrm{e}^{i\vec{k}\vec{t}} \sum_{\sigma} \int d^3r\; \underbrace{\langle\vec{t},\alpha|\vec{r}+\vec{t},\sigma\rangle}_{\langle\vec{0},\alpha|\vec{r},\sigma\rangle} \underbrace{\mathrm{e}^{i\vec{k}\vec{r}}\langle\vec{r},\sigma|u_{\vec{k},n}\rangle}_{\langle\vec{r},\sigma|\Psi_{\vec{k},n}\rangle} \\
&= \sum_{\vec{t},\alpha} |\vec{t},\alpha\rangle\mathrm{e}^{i\vec{k}\vec{t}}\langle\vec{0},\alpha|\psi_{\vec{k},n}\rangle
\end{aligned}
\tag{4.27}
$$

This shows us how we can represent the entire wave function with a vector $\vec{c}_n$ with elements $c_{\alpha,n} = \langle\vec{0},\alpha|\Psi_{k,n}\rangle$, which has the dimension of the number of orbitals in the elementary unit cell.

Our next goal is to find an equation for this vector. We will use the matrix elements of the Hamilton operator in the basis of tight-binding orbitals

$$
\langle\vec{t},\alpha|\hat{H}|\vec{t}',\beta\rangle = h_{\vec{t},\alpha,\vec{t}',\beta}
\tag{4.28}
$$

Let us insert the ansatz Eq. 4.27 into the Schrödinger equation

$$
\begin{aligned}
&\hat{H}|\Psi_{k,n}\rangle = |\Psi_{k,n}\rangle\epsilon_{k,n} \\
\overset{\text{Eq. 4.27}}{\Rightarrow}\quad & \hat{H}\sum_{\vec{t},\alpha} |\vec{t},\alpha\rangle\mathrm{e}^{i\vec{k}\vec{t}}\langle\vec{0},\alpha|\Psi_{k,n}\rangle = \sum_{\vec{t},\alpha} |\vec{t},\alpha\rangle\mathrm{e}^{i\vec{k}\vec{t}}\langle\vec{0},\alpha|\Psi_{k,n}\rangle\epsilon_{k,n} \\
\overset{\langle\vec{0},\beta|}{\Rightarrow}\quad & \sum_{\vec{t},\alpha} \underbrace{\langle\vec{0},\beta|\hat{H}|\vec{t},\alpha\rangle}_{h_{\vec{0},\beta,\vec{t},\alpha}} \mathrm{e}^{i\vec{k}\vec{t}}\langle\vec{0},\alpha|\Psi_{k,n}\rangle = \sum_{\vec{t},\alpha} \underbrace{\langle\vec{0},\beta|\vec{t},\alpha\rangle}_{\delta_{\vec{0},\vec{t}}\delta_{\alpha,\beta}} \mathrm{e}^{i\vec{k}\vec{t}}\langle\vec{0},\alpha|\Psi_{k,n}\rangle\epsilon_{k,n} \\
\Rightarrow\quad & \underbrace{\left[\sum_{\vec{t},\alpha} h_{\vec{0},\beta,\vec{t},\alpha}\mathrm{e}^{i\vec{k}\vec{t}}\right]}_{h_{\alpha,\beta}(\vec{k})} \underbrace{\langle\vec{0},\alpha|\Psi_{k,n}\rangle}_{c_{\alpha,n}(\vec{k})} = \underbrace{\langle\vec{0},\beta|\Psi_{k,n}\rangle}_{c_{\beta,n}(\vec{k})}\epsilon_{k,n} \\
\Rightarrow\quad & \boldsymbol{h}(\vec{k})\vec{c}_n(\vec{k}) = \vec{c}_n(\vec{k})\epsilon_n(\vec{k})
\end{aligned}
\tag{4.29}
$$

Thus, we obtain

SCHRÖDINGER EQUATION IN K-SPACE

$$\boldsymbol{h}(\vec{k})\vec{c}_n(\vec{k}) = \vec{c}_n(\vec{k})\epsilon_n(\vec{k}) \tag{4.30}$$

with $c_{\alpha,n}(\vec{k}) \overset{\text{def}}{=} \langle\vec{0},\beta|\Psi_{\vec{k},n}\rangle$, respectively $|\Psi_{\vec{k},n}\rangle \overset{\text{Eq. 4.27}}{=} \sum_{\vec{t}\alpha,n} |\vec{t},\alpha\rangle \mathrm{e}^{i\vec{k}\vec{t}} c_{\alpha,n}(\vec{k})$

and

$$h_{\alpha,\beta}(\vec{k}) \overset{\text{def}}{=} \sum_{\vec{t},\alpha} h_{\vec{0},\beta,\vec{t},\alpha}\mathrm{e}^{i\vec{k}\vec{t}} \tag{4.31}$$

Thus, we obtained a k-dependent eigenvalue equation in matrix form with a finite and usually small dimension. This problem can be broken down into smaller subproblems, if we exploit the symmetry arguments that we used earlier for the molecules. Symmetry can be used especially at the high-symmetry points and lines in reciprocal space, where the band structure is usually represented.

Worked example

Let us consider the example of a linear chain atoms with one s-type orbital. There shall be two atoms per unit cell. The lattice constant shall be a. The orbitals are denoted by $(t/a, \alpha)$, where $\alpha \in \{1, 2\}$ denote the two orbitals and t/a is the number of lattice displacements.

The two atoms have orbital energies $\bar{\epsilon}_1$ and $\bar{\epsilon}_1$. We denote the hopping matrix element between orbitals $(i, 1)$ and $(i, 2)$ by $h_{ss\sigma}$ and the one between $(i, 2)$ and $(i + 1, 1)$ by $h'_{ss\sigma}$.

For the sake of simplicity we only consider wave functions with one spin direction. The bands of the two spin direction are identical, because there is no magnetic field. When we consider one spin direction, the wave function component of the other spin direction is zero.

The infinite Hamiltonian has the form

$$\mathbf{h} = \left(\begin{array}{c|cc|cc|cc|c}
 & \vdots & \vdots & \vdots & & & & \\
\hline
\dots & \bar{\epsilon}_1 & h_{ss\sigma} & 0 & 0 & \dots & & \\
\dots & h_{ss\sigma} & \bar{\epsilon}_2 & h'_{ss\sigma} & 0 & \dots & & \\
\hline
\dots & 0 & h'_{ss\sigma} & \bar{\epsilon}_1 & h_{ss\sigma} & 0 & 0 & \dots \\
\dots & 0 & 0 & h_{ss\sigma} & \bar{\epsilon}_2 & h'_{ss\sigma} & 0 & \dots \\
\hline
 & & \dots & 0 & h'_{ss\sigma} & \bar{\epsilon}_2 & h_{ss\sigma} & \dots \\
 & & \dots & 0 & 0 & h_{ss\sigma} & \bar{\epsilon}_1 & \dots \\
\hline
 & & & \vdots & \vdots & \vdots & \vdots &
\end{array}\right)$$

The Hamilton matrix elements can also be written as

$$\boldsymbol{h}_{i,j} = \begin{pmatrix} \bar{\epsilon}_1 & h_{ss\sigma} \\ h_{ss\sigma} & \bar{\epsilon}_2 \end{pmatrix}\delta_{i,j} + \begin{pmatrix} 0 & 0 \\ h'_{ss\sigma} & 0 \end{pmatrix}\delta_{i,j+1} + \begin{pmatrix} 0 & h'_{ss\sigma} \\ 0 & 0 \end{pmatrix}\delta_{i,j-1} \tag{4.32}$$

where i and j are the indices of the lattice translations $\vec{t}_j = aj$. The components of the 2×2 and the orbital in the unit cell.

When we transform the Hamiltonian into Bloch representation using Eq. 4.31, we obtain the k-dependent Hamiltonian with $\vec{t}_j = aj$

$$\boldsymbol{h}(\vec{k}) \overset{\text{Eq. 4.31}}{=} \sum_j \boldsymbol{h}(0,j)\mathrm{e}^{ikaj} = \begin{pmatrix} \bar\epsilon_1 & h_{ss\sigma} + h'_{ss\sigma}\mathrm{e}^{ika} \\ h_{ss\sigma} + h'_{ss\sigma}\mathrm{e}^{-ika} & \bar\epsilon_2 \end{pmatrix}$$

The eigenvalues $\epsilon_n(k)$ and eigenvectors $c_\alpha(k)$ are obtained from the characteristic equation $\det|\boldsymbol{h}(k) - \epsilon\boldsymbol{S}| = 0$.

$$\begin{aligned} (\bar\epsilon_1 - \epsilon)(\bar\epsilon_2 - \epsilon) - \left(h_{ss\sigma} + h'_{ss\sigma}\mathrm{e}^{+ika}\right)\left(h_{ss\sigma} + h'_{ss\sigma}\mathrm{e}^{-ika}\right) &= 0 \\ \epsilon^2 - 2\frac{\bar\epsilon_1+\bar\epsilon_2}{2}\epsilon + \left(\frac{\bar\epsilon_1+\bar\epsilon_2}{2}\right)^2 - \left(\frac{\bar\epsilon_1+\bar\epsilon_2}{2}\right)^2 + \bar\epsilon_1\bar\epsilon_2 &= h^2_{ss\sigma} + h'^2_{ss\sigma} + 2h_{ss\sigma}h'_{ss\sigma}\cos(ka) \\ \left(\epsilon - \frac{\bar\epsilon_1+\bar\epsilon_2}{2}\right)^2 = \left(\frac{\bar\epsilon_1-\bar\epsilon_2}{2}\right)^2 &= h^2_{ss\sigma} + h'^2_{ss\sigma} + 2h_{ss\sigma}h'_{ss\sigma}\cos(ka) \\ \epsilon_n(k) = \frac{\bar\epsilon_1+\bar\epsilon_2}{2} \pm \sqrt{\left(\frac{\bar\epsilon_1-\bar\epsilon_2}{2}\right)^2} &\overline{+ h^2_{ss\sigma} + h'^2_{ss\sigma} + 2h_{ss\sigma}h'_{ss\sigma}\cos(ka)} \end{aligned} \tag{4.33}$$

The first band $n = 1$ is obtained with $\pm = -$ and the second band is obtained with $\pm = +$.

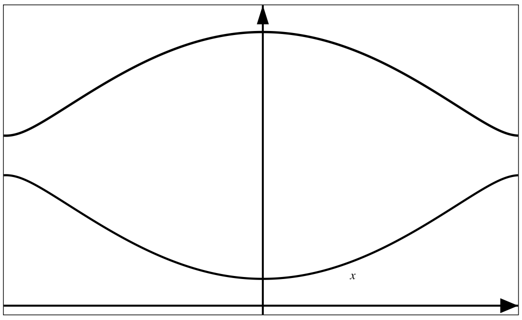


Fig. 4.6: Bandstructure of the alternating linear chain. (Energy vs. k)

Let us investigate some special cases

- chain of decoupled molecules: We set one hopping parameter to zero, i.e. $h'_{ss\sigma} = 0$. The result are two k-independent bands for the molecular orbitals

$$\epsilon_n(k) = \frac{\bar\epsilon_1+\bar\epsilon_2}{2} \pm \sqrt{\left(\frac{\bar\epsilon_1-\bar\epsilon_2}{2}\right)^2 + h^2_{ss\sigma}} \tag{4.34}$$

- both atoms and hoppings are identical, i,e, $\bar\epsilon_1 = \bar\epsilon_2$ and $h_{ss\sigma} = h'_{ss\sigma}$.

$$\begin{aligned} \epsilon_n(k) &= \bar\epsilon_1 \pm \sqrt{2h^2_{ss\sigma} + 2h^2_{ss\sigma}\cos(ka)} \\ &= \bar\epsilon_1 \pm |h_{ss\sigma}| \underbrace{\sqrt{2+2\cos(ka)}}_{\sqrt{4cos^2(ka/2)}} \\ &= \bar\epsilon_1 \pm \left|2h_{ss\sigma}\cos\left(\frac{1}{2}ka\right)\right| \end{aligned} \tag{4.35}$$

This is the bandstructure of a mono-atomic chain

$$\epsilon(k) = \bar\epsilon_1 + 2h_{ss\sigma}\cos(ka/2)$$

folded back into the first Brillouin zone.

- for alternating orbital energies or alternating hopping matrix elements, The band gap at the zone boundary opens, so that an insulator is obtained if the orbitals are half filled.

Note that $h_{ss\sigma}$ is negative, because it is approximately equal to the overlap matrix element multiplied with the potential in the bond-center. The potential is negative.

4.5 Thermodynamics of non-interacting electrons

What makes the many-electron problem so complicated is the interaction between the electrons. A similar problem is known from classical mechanics. While the Kepler problem, a two-body problem, can be solved analytically, no such solution exists for the three-body problem, where the gravitational attraction between the planets is considered as well.

4.5.1 Thermodynamics revisited

General setup

For a system described by a Hamiltonian $\hat{H}$, which is coupled to a heat bath and reservoirs for the observables $\{X_i\}$, the free energy $F(T,\{f_j\})$ is (Eq. 1.61 of [31])

$$F(T,\{f_j\}) = -k_BT\ln\Big[\mathrm{Tr}\Big(\mathrm{e}^{-\beta(\hat{H}+\sum_j f_j\hat{X}_j)}\Big)\Big] \tag{4.36}$$

The heat bath is characterized by the temperature T and the other reservoirs are characterized by the other intrinsic variables f_j.

From the free energy, we can obtain (Eqs. 1.65, 1.66, 1.67 of [31]) the expectation values for the observables X and the internal energy U, the expectation value of the energy of the system by

$$S(T,\{f_k\}) = -\frac{\partial F(T,\{f_k\})}{\partial T} \tag{4.37}$$

$$X_i(T,\{f_k\}) = +\frac{\partial F(T,\{f_k\})}{\partial f_i}$$

$$U(T,\{f_k\}) = F(T,\{f_k\}) + TS(T,\{f_k\}) - \sum_i f_iX_i(T,\{f_k\}) \tag{4.38}$$

Specific relations

For our purposes we will start from a **grand canonical ensemble**, which describes a system in thermal equilibrium with a heat bath and a particle reservoir.

extensive variable	X	f	intensive variable
volume	V	p	pressure
particle number	N	$-\mu$	chemical potential

The thermodynamic potential for the grand canonical ensemble is the **grand potential** $\Omega(T,V,\mu)$.

$$\Omega(T,V,\mu) = -k_BT\ln\Big[\mathrm{Tr}\Big(\mathrm{e}^{-\beta(\hat{H}+\mu\hat{N})}\Big)\Big] \tag{4.39}$$

which yields

$$S(T,V,\mu) = -\frac{\partial\Omega(T,V,\mu)}{\partial T} \tag{4.40}$$

$$N(T,V,\mu) = -\frac{\partial\Omega(T,V,\mu)}{\partial\mu} \tag{4.41}$$

$$U(T,V,\mu) = \Omega(T,V,\mu) + TS(T,V,\mu) + \mu N(T,V,\mu) \tag{4.42}$$

These equations can be put together to a Legendre transform to a quantity depending on the extensive variables

$$U(S,V,N) = \underset{T,\mu}{\text{stat}}\Big(\Omega(T,V,\mu) + TS + \mu N\Big) \tag{4.43}$$

which yields the **fundamental relation of thermodynamics**

$$dU = \underbrace{TdS}_{\text{heat}} + \overbrace{\underbrace{\frac{d\Omega}{dV}}_{-p} dV + \mu dN}^{\text{Work}} \tag{4.44}$$

which in turn defines **Heat** and **Work**.

The pressure is defined as the derivative of the internal energy with respect to volume at fixed entropy and particle number

$$p \overset{\text{def}}{=} -\frac{\partial U(S,V,N)}{\partial V} = -\frac{\partial \Omega(T,V,\mu)}{\partial V} \tag{4.45}$$

4.5.2 Maximum entropy, minimum free energy principle and the meaning of the grand potential

The grand potential is related to the total energy of a composite system which consists of the material under study with an internal energy $U(S,V,N)$, the the heat-bath with energy U_{hb} and a particle reservoir with energy $U_{pr} = \mu N_{pr}$.

Consider a starting configuration where the entropy and the particle number on the system under study vanish initially, that is $(S = 0, N = 0)$. Also the particle number of the particle reservoir and the energy of the heat bath vanish, i.e. $(U_{hb} = 0, N_{pr} = 0)$.

As the system is brought into contact with the heat bath and its reservoirs, particles are transferred from the reservoir to the system and heat $T_{hb}dS_{hb}$ is transferred from the heat bath to the system. Thus the energy of the heat bath is

$$U_{hb}(S_{hb}) = \int_0^{S_{hb}} dS'\, T_{hb} = T_{hb}S_{hb} \tag{4.46}$$

The energy conservation between heat bath and our system has the form

$$\begin{aligned}
& \frac{\partial U}{\partial S} dS + T_{hb}dS_{hb} = 0 \\
\Rightarrow \quad & \frac{\partial U}{\partial S} dS - T_{hb}dS + T_{hb}(dS_{hb} + dS) = 0 \\
\overset{dS_{tot} := S + S_{hb}}{\Rightarrow} \quad & dU_{hb} = -\frac{\partial U}{\partial S} dS = -T_{hb}dS + T_{hb}dS_{tot}
\end{aligned} \tag{4.47}$$

where dS_{tot} is the total amount of entropy produced in the process. This entropy production is what drives the system into thermal equilibrium with its reservoirs. The entropy is produced during the heat transfer between heat bath and our system. Entropy is generally generated when heat is transferred along a temperature gradient or between two systems at different temperatures.

The total energy E_{tot} and total number of particles is conserved, i.e. $dE_{tot} = 0$, where the total energy is

$$E_{tot}(T_{hb}, V, \mu_{pr}, S, N, U_{hb}, N_{pr}) = \underbrace{U(S,V,N)}_{\text{system}} + \underbrace{U_{hb}}_{\text{heat bath}} + \underbrace{\mu_{pr}N_{pr}}_{\text{particle reservoir}} \tag{4.48}$$

Using the particle conservation $N_{pr} = -N$ and Eq. 4.47, the total energy obtains the form

$$E_{tot}(T_{hb}, V, \mu_{pr}; S, S_{tot}, N) = U(S,V,N) - T_{hb}S + T_{hb}S_{tot} - \mu_{pr}N \tag{4.49}$$

As the system evolves, the total entropy dS_{tot} increases until it reaches its maximum. While it evolves, the process obeys energy conservation, that is $dE_{tot} = 0$. Hence the state evolves on a hypersurface defined by $E_{tot} =$ const in the S, N, S_{tot} space. As a result, the total entropy can be expressed as function of S and N.

$$-T_{hb}S_{tot}(S, V, N) = U(S, V, N) - T_{hb}S - \mu_{pr}N - U(0, V, 0) \tag{4.50}$$

This is only valid though, if the state (S, V, N) is connected by an energy conserving process to the initial state $(S = 0, V, N = 0)$ defined above!

The thermal process comes to halt at a maximum of the total entropy. This is when an infinitesimal heat transfer from the heat bath does no more produce entropy in the system, which in turn implies that our system has reached the same temperature as the heatbath. The **maximum entropy principle** translates into a **minimum energy principle** for the grand potential

$$\begin{aligned}\Omega(T_{hb}, V, \mu_{pr}) &= U(0, V, 0) - TS_{tot}(S, V, N)\\ &= \min_{S,N}\Big(U(S, V, N) - T_{hb}S - \mu_{pr}N\Big) - E_{tot}\end{aligned} \tag{4.51}$$

The grand potential is therefore the total energy of the system under study and the reservoirs (heat bath and particle reservoir) which are in contact with our system, minus the heat generated respectively transported out of the entire system.

$$\Omega(T_{HB}, V, \mu_{pr}) = \min_{S,N}\Big(U(S, V, N) - T_{hb}S - \mu_{pr}N\Big) \tag{4.52}$$

4.5.3 Many-particle states of non-interacting electrons

Nevertheless, we should start with the most simple case, namely the non-interacting electron gas. The Hamiltonian for a non-interacting electron gas in a potential $v(\vec{r})$ has the form

$$\begin{aligned}\hat{H} &= \sum_{i=1}^{N} \hat{h}_i = \sum_{i=1}^{N} \frac{\hat{p}_i^2}{2m_e} + v(\hat{\vec{r}}_i)\\ &= \sum_{i=1}^{N} \int d^4x_1 \cdots \int d^4x_N\, |\vec{x}_1, \ldots, \vec{x}_N\rangle \left(\frac{-\hbar^2}{2m_e}\vec{\nabla}_i^2 + v(\vec{r}_i)\right) \langle \vec{x}_1, \ldots, \vec{x}_N|\end{aligned} \tag{4.53}$$

Here, $\hat{\vec{p}}_i$ is the momentum operator acting on the i-th particle, and $\hat{\vec{r}}_i$ is the position operator of the i-th particle.

Even though the electrons do not interact, they do not move completely independently. The reason is, again, the Pauli-principle or the antisymmetry requirement for the wave function. If one electron is at a certain position, no other electron with the same spin can approach it. There is no repulsive potential. The two electrons simply do not "like" each other.

In practice, this means that we have to use an antisymmetric wave function. Luckily, the eigenstates of a system of non-interacting electrons (with a non-degenerate ground state) is a single **Slater determinant**. In contrast, the wave function for interacting electrons is, in general, a superposition of many Slater determinants.

Here, we mention a few facts about the states of non-interacting electrons without proving them. Some are fundamental facts of statistical mechanics and others will be shown later in this book, in chapter 6.

Hamilton eigenstates: The eigenstates are Slater determinants made of eigenstates $\phi_n(\vec{x})$ of the one-particle Hamiltonian $\hat{h} = \frac{\hat{\vec{p}}^2}{2m} + v(\hat{\vec{r}})$

$$\underbrace{\left[\frac{-\hbar^2}{2m}\vec{\nabla}^2 + v(\vec{r})\right]\phi_n(\vec{x})}_{\langle\vec{x}|\hat{h}|\phi_n\rangle} = \phi_n(\vec{x})\epsilon_n \tag{4.54}$$

The Slater determinant, which is an eigenstate of the many-particle system, can be represented by its **number representation**. The **occupations numbers**, which may be zero or one, are combined into a vector $\vec{\sigma}$. The i-th element of $\vec{\sigma}$ vector is the occupation number of the i-th eigenstate of the one-particle Hamiltonian.

Thus the eigenstates of the many-particle Hamiltonian can be labeled by the vector $\vec{\sigma}$

$$|\Phi_{\vec{\sigma}}\rangle = |\sigma_1, \sigma_2, \dots\rangle \tag{4.55}$$

Energy eigenvalues: A Slater determinant $|\Phi_{\vec{\sigma}}\rangle$ has the energy

$$E_{\vec{\sigma}} = \sum_{n=1}^{\infty} \sigma_n \epsilon_n$$

where the ϵ_{n_i} are the one-particle energies defined above in Eq. 4.54.

Particle number: A Slater determinant $|\Phi_{\vec{\sigma}}\rangle$ has the particle number

$$N_{\vec{\sigma}} = \sum_{n=1}^{\infty} \sigma_n$$

`Editor: We came up to here on Wednesday, Dec. 2, 2015`

Grand potential: The **grand canonical ensemble** describes a system in thermal equilibrium which is in contact with a heat bath and a particle reservoir. A heat bath is characterized by a temperature T and the particle reservoir is characterized by a **chemical potential** μ. The temperature is often represented by the **thermodynamic beta** $\beta = \frac{1}{k_B T}$. k_B is the **Boltzmann constant**.

The thermodynamic potential for the grand-canonical ensemble is the **grand potential**. The grand potential for a system of non-interacting fermions is

$$\begin{aligned}
\Omega(T,\mu) &\overset{\text{def}}{=} -k_B T \ln \mathrm{Tr}[\mathrm{e}^{-\beta(\hat{H}-\mu\hat{N})}] \\
&= -k_B T \ln \underbrace{\Bigg[\sum_{\vec{\sigma}} \exp\Big(-\beta\Big(E_{\vec{\sigma}} - \mu N_{\vec{\sigma}}\Big)\Big)\Bigg]}_{\text{partition sum}} \\
&= -k_B T \ln \Bigg[\sum_{\vec{\sigma}} \mathrm{e}^{-\beta\sum_{i=1}^{\infty}\sigma_i(\epsilon_i-\mu)}\Bigg] \\
&= -k_B T \ln \Bigg[\prod_{i=1}^{\infty}\sum_{\sigma_i=0}^{1} \mathrm{e}^{-\beta\sigma_i(\epsilon_i-\mu)}\Bigg] \\
&= -k_B T \ln \Bigg[\prod_{i=1}^{\infty}\Big(1+\mathrm{e}^{-\beta(\epsilon_i-\mu)}\Big)\Bigg] \\
&= -k_B T \sum_{i=1}^{\infty} \ln\Big(1+\mathrm{e}^{-\beta(\epsilon_i-\mu)}\Big)
\end{aligned} \tag{4.56}$$

The interesting observation is that the thermodynamic potential of non-interacting particles, here electrons, is a sum over one-particle states.

4.5.4 Density of States

The grand potential turns out to be a simple sum over one-electron states with an integrand that depends on the one-particle energy ϵ_i, the chemical potential μ and the temperature T. This implies that the thermodynamics quantities, that are derivatives of the grand potential can be described as a sum over one-particle states as well. Therefore it is convenient to split off the spectrum of the one particle states. We can represent the **one-particle spectrum** by the **density of states**

DENSITY OF STATES

The one-particle density of states is defined as

$$D(\epsilon) \stackrel{\text{def}}{=} \sum_n \delta(\epsilon - \epsilon_n) \tag{4.57}$$

where $\delta(x)$ is Dirac's δ-function and the one-particle energies are denoted by ϵ_n.

Note that the density of state as defined here is a macroscopic quantity. In practice one often uses the density of state per volume or per unit cell.

With the density of states, we can separate the thermodynamic information from the system-specific information. The density of states is specific for the system, while the Fermi-distribution is independent of the system and contains the thermodynamic information.

Projected Density of States: The total density of states discussed above is a special case of the density of states operator

$$\hat{D}(\epsilon) = \sum_n |\psi_n\rangle\delta(\epsilon - \epsilon_n)\langle\psi_n| \tag{4.58}$$

which allows one to represent any thermal expectation value of one-particle operators by

$$\langle A\rangle = \sum_n f(\epsilon_n)\langle\psi_n|\hat{A}|\psi_n\rangle = \int d\epsilon\; f(\epsilon)\;\mathrm{Tr}\left[\hat{D}(\epsilon)\hat{A}\right] \tag{4.59}$$

Grand potential expressed by the Density of States:

$$\begin{aligned}\Omega(T,\mu) &= -k_BT\sum_{n=1}^{\infty}\ln\left(1+\mathrm{e}^{-\beta(\epsilon_n-\mu)}\right)\\ &= \int_{-\infty}^{\infty} d\epsilon\; D(\epsilon)\Big[-k_BT\ln\left(1+\mathrm{e}^{-\beta(\epsilon-\mu)}\right)\Big]\end{aligned} \tag{4.60}$$

4.5.5 Fermi distribution

The mean occupations f_i are given by the **Fermi-distribution function** evaluated for the one-particle energy ϵ_i of the state.(see *"ΦSX: Statistical Physics"*)

$$f_{T,\mu}(\epsilon) \stackrel{\text{def}}{=} \frac{1}{1+\mathrm{e}^{\beta(\epsilon-\mu)}} \tag{4.61}$$

where $\beta = 1/(k_BT)$. The Fermi function is shown in Fig. 4.7. In the absence of temperature effects we set the temperature to zero, so that the Fermi distribution function is turned into a step function.

In the following we will also need the derivatives of the Fermi distribution with respect to temperature and chemical potential. They can be expressed by the energy derivative of the Fermi function.

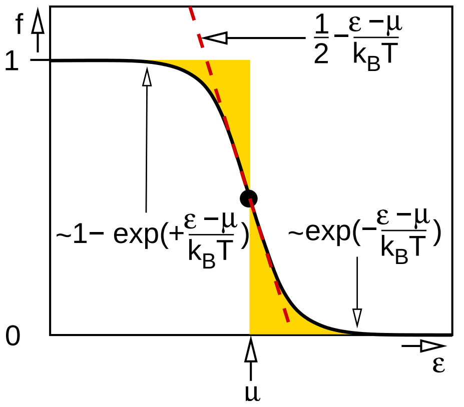


Fig. 4.7: Fermi distribution $f_{T,\mu}(\epsilon)=(1+\mathrm{e}^{\frac{1}{k_BT}(\epsilon-\mu)})^{-1}$.

[8]

$$\partial_T f_{T,\mu}(\epsilon) = -\frac{\epsilon-\mu}{T}\partial_\epsilon f_{T,\mu}(\epsilon) \tag{4.63}$$

$$\partial_\mu f_{T,\mu}(\epsilon) = -\partial_\epsilon f_{T,\mu}(\epsilon) \tag{4.64}$$

In the low-temperature limit, the energy derivative of the Fermi function approaches the negative delta function at the Fermi level.

$$\lim_{T\to 0}\partial_\epsilon f_{T,\mu}(\epsilon) = -\delta(\epsilon-\mu) \tag{4.65}$$

Thermodynamics of non-interacting electrons

As shown in section 8.5 of *"ΦSX: Statistical Physics"*[31], we can express the thermodynamic quantities of non-interacting Fermions by the density of states

Grand potential: For our case the free energy is the grand potential $\Omega(T,\mu)$, which is obtained as

$$\begin{aligned}\Omega(T,\mu) &= -k_BT\sum_{n=1}^{\infty}\ln\left(1+\mathrm{e}^{-\beta(\epsilon_n-\mu)}\right)\\ &= \int_{-\infty}^{\infty}d\epsilon\; D(\epsilon)\Big[-k_BT\ln\left(1+\mathrm{e}^{-\beta(\epsilon-\mu)}\right)\Big]\end{aligned} \tag{4.66}$$

The function in the integral has the form.

$$-k_BT\ln\left(1+\mathrm{e}^{-\beta(\epsilon-\mu)}\right)\approx\begin{cases}\epsilon-\mu & \text{for } \epsilon-\mu\ll -k_BT\\ -k_BT\ln(2)+\frac{1}{2}\beta(\epsilon-\mu) & \text{for } |\epsilon-\mu|\ll k_BT\\ 0 & \text{for } (\epsilon-\mu\gg k_BT\end{cases} \tag{4.67}$$

---

[8]

$$\begin{aligned}\partial_\epsilon\left[1+\mathrm{e}^{\beta(\epsilon-\mu)}\right]^{-1} &= \beta\left[\frac{d}{dx}\bigg|_{x=\beta(\epsilon-\mu)}\left[1+\mathrm{e}^{\beta(\epsilon-\mu)}\right]^{-1}\right]\\ \partial_\mu\left[1+\mathrm{e}^{\beta(\epsilon-\mu)}\right]^{-1} &= -\beta\left[\frac{d}{dx}\bigg|_{x=\beta(\epsilon-\mu)}\left[1+\mathrm{e}^{\beta(\epsilon-\mu)}\right]^{-1}\right] = -\partial_\epsilon f_{T,\mu}(\epsilon)\\ \partial_T\left[1+\mathrm{e}^{\beta(\epsilon-\mu)}\right]^{-1} &= \left(-\frac{1}{T}\beta(\epsilon-\mu)\right)\left[\frac{d}{dx}\bigg|_{x=\beta(\epsilon-\mu)}\left[1+\mathrm{e}^{\beta(\epsilon-\mu)}\right]^{-1}\right] = -\frac{\epsilon-\mu}{T}\partial_\epsilon f_{T,\mu}(\epsilon)\end{aligned} \tag{4.62}$$

In the following we express many quantities in terms of the occupations. Expressing the grand potential by occupations is, however, not a good idea, because it depends extremely sensitive on the exact occupation if the orbital is nearly filled.

Particle number and chemical potential: The particle number $N(T,\mu)$ is obtained as

$$\begin{aligned}
N(T,\mu) &= -\frac{\partial\Omega(T,\mu)}{\partial\mu} && (4.68)\\
&= \int_{-\infty}^{\infty} d\epsilon\; D(\epsilon)\Big[-\partial_\mu\Big(-k_BT\ln\Big(1+\mathrm{e}^{-\beta(\epsilon-\mu)}\Big)\Big)\Big] && (4.69)\\
&= \int_{-\infty}^{\infty} d\epsilon\; D(\epsilon)f_{T,\mu}(\epsilon) && (4.70)
\end{aligned}$$

Usually, the particle number is provided instead of the chemical potential. The particle number follows typically from the charge neutrality condition. In order to extract the chemical potential on needs to evaluate The partiale number as function of chemical potential and find that value for which $N(T,\mu)=N$.

Entropy:

$$\begin{aligned}
S(T,\mu) &= -\frac{\partial\Omega(T,\mu)}{\partial T}\\
&= \int_{-\infty}^{\infty} d\epsilon\; D(\epsilon)\Big[-\partial_T\Big(-k_BT\ln\Big(1+\mathrm{e}^{-\beta(\epsilon-\mu)}\Big)\Big)\Big]\\
&= \int_{-\infty}^{\infty} d\epsilon\; D(\epsilon)\Big[\Big(k_B\ln\underbrace{\Big(1+\mathrm{e}^{-\beta(\epsilon-\mu)}\Big)}_{1/(1-f)}\Big)+\Big(k_BT\underbrace{\frac{1}{1+\mathrm{e}^{-\beta(\epsilon-\mu)}}}_{f}\underbrace{\Big(\frac{1}{k_BT^2}(\epsilon-\mu)\Big)}_{\frac{1}{T}\ln(f^{-1}-1)}\Big)\Big]\\
&= \int_{-\infty}^{\infty} d\epsilon\; D(\epsilon)\Big[k_B\Big(-\ln(1-f)+f\ln(\frac{1-f}{f})\Big)\Big]\\
&= \int_{-\infty}^{\infty} d\epsilon\; D(\epsilon)\Big[-k_B\Big(f\ln(f)+(1-f)\ln(1-f)\Big)\Big] && (4.71)
\end{aligned}$$

The entropy can be obtained from the occupations $f=f_{\beta,\mu}(\epsilon)$.

Eq. 4.71 tells us completely filled or completely empty orbitals do not contribute to the entropy of a system. A half occupied orbital contributes $k_B\ln(2)$ to the entropy. This is exactly the entropy of a yes-no decision, because the orbital can be either filled or empty with equal probability.

Internal energy: The internal energy U is the expectation value of the energy of the system.

Before we continue, we express the grand potential by the occupations, because that will simplify the equations.

$$\begin{aligned}
\Omega(T,\mu) &= \int_{-\infty}^{\infty} d\epsilon\; D(\epsilon)\Big[-k_BT\ln\Big(1+\mathrm{e}^{-\beta(\epsilon-\mu)}\Big)\Big]\\
&= \int_{-\infty}^{\infty} d\epsilon\; D(\epsilon)\Big[k_BT\ln\left(1-f_{T,\mu}(\epsilon)\right)\Big] && (4.72)
\end{aligned}$$

$$\begin{aligned}
U(T,\mu)&=\Omega(T,\mu)+TS(T,\mu)+\mu N(T,\mu)\\
&=\int_{-\infty}^{\infty} d\epsilon\; D(\epsilon)\Big[\underbrace{k_BT\ln(1-f)}_{\leftarrow\Omega}\underbrace{-k_BT\Big(f\ln(f)+(1-f)\ln(1-f)\Big)}_{\leftarrow TS}+\mu f\Big]\\
&=\int_{-\infty}^{\infty} d\epsilon\; D(\epsilon)\Big[f\Big(k_BT\underbrace{\ln(\frac{1-f}{f})}_{\beta(\epsilon-\mu)}+\mu\Big]\\
&=\int_{-\infty}^{\infty} d\epsilon\; D(\epsilon)\,f_{\beta,\mu}(\epsilon)\,\epsilon
\end{aligned}\tag{4.73}$$

where, again, $f=f_{\beta,\mu}(\epsilon)$. The internal energy has a simply interpretation: it is the sum of the energies of the occupied states.

Pressure

The pressure is defined as the derivative of the total energy upon contraction.

$$p=-\frac{\partial U(S,V,N)}{\partial V}\tag{4.74}$$

Note, that it is not allowed to calculate the pressure as derivative of $U(T,V,\mu)$! The pressure is the derivative of the total energy with fixed occupations.

However, we can calculate the pressure from the grand potential

$$\begin{aligned}
p&=-\frac{\partial\Omega(T,V,\mu)}{\partial V}\\
&=\int_{-\infty}^{\infty} d\epsilon\;\frac{dD(\epsilon)}{dV}\Big[+k_BT\ln\Big(1+\mathrm{e}^{-\beta(\epsilon-\mu)}\Big)\Big]
\end{aligned}\tag{4.75}$$

Why does this give the same result as the derivative of $U(S,V,N)$: The grand potential describes the energy of the system and the reservoirs. If the volume expansion changes the entropy or the particle number of the system, the corresponding energy change is exactly balanced to first order by an opposite energy change of the reservoirs. Thus including the energies of the reservoirs produces the same energy change as keeping entropy and particle number fixed.

Bulk modulus: The **bulk modulus** is the measure of how a material can resist compression.

$$B_0=-V\frac{\partial P}{\partial V}=V\frac{\partial^2 U(S,V,N)}{\partial V^2}\tag{4.76}$$

Thus it is a second derivative of the internal energy.

Heat capacity The heat capacity is the amount of energy required to raise the temperature by one unit.

$$C_V=\frac{\partial U(T,V,N)}{\partial T}\tag{4.77}$$

When calculating the heat capacity from $U(T,V,\mu)$ we need to account for the shift of the chemical displacement with temperature

$$
\begin{aligned}
c_V &= \frac{\partial U(T,V,\mu)}{\partial T} + \frac{\partial U(T,V,\mu)}{\partial \mu}\frac{d\mu(T,V,N)}{dT} \\
&= \int_{-\infty}^{\infty} d\epsilon\ D(\epsilon)\frac{\partial f_{T,\mu}(\epsilon)}{\partial T}\epsilon + \int_{-\infty}^{\infty} d\epsilon\ D(\epsilon)\frac{\partial f_{T,\mu}(\epsilon)}{\partial \mu}\epsilon\frac{d\mu(T,V,N)}{dT} \\
&= \int_{-\infty}^{\infty} d\epsilon\ D(\epsilon)\frac{\epsilon-\mu}{T}\frac{\partial f_{T,\mu}(\epsilon)}{\partial \mu}\epsilon + \int_{-\infty}^{\infty} d\epsilon\ D(\epsilon)\frac{\partial f_{T,\mu}(\epsilon)}{\partial \mu}\epsilon\frac{d\mu(T,V,N)}{dT} \\
&= \frac{1}{T}\int_{-\infty}^{\infty} d\epsilon\ D(\epsilon)\epsilon\frac{\partial f_{T,\mu}(\epsilon)}{\partial \mu}\Big[\epsilon-\mu+T\frac{d\mu(T,V,N)}{dT}\Big] \\
&= \frac{1}{T}\int_{-\infty}^{\infty} d\epsilon\ D(\epsilon)\frac{\partial f_{T,\mu}(\epsilon)}{\partial \mu}\Big[\epsilon-\mu+T\frac{d\mu(T,V,N)}{dT}\Big]^2 \\
&+ \frac{1}{T}\Big[\mu-T\frac{d\mu(T,V,N)}{dT}\Big]\int_{-\infty}^{\infty} d\epsilon\ D(\epsilon)\frac{\partial f_{T,\mu}(\epsilon)}{\partial \mu}\Big[\epsilon-\mu+T\frac{d\mu(T,V,N)}{dT}\Big]
\end{aligned}
\tag{4.78}
$$

The particle number conservation provides us with $d\mu/dT$.

$$
\begin{aligned}
0 = \frac{dN}{dT} &= \int d\epsilon D(\epsilon)\Big[\frac{\partial f_{T,\mu}(\epsilon)}{\partial T} + \frac{\partial f_{T,\mu}(\epsilon)}{\partial \mu}\frac{d\mu}{dT}\Big] \\
&= \int d\epsilon D(\epsilon)\Big[\frac{(\epsilon-\mu)}{T}\frac{\partial f_{T,\mu}(\epsilon)}{\partial \mu} + \frac{\partial f_{T,\mu}(\epsilon)}{\partial \mu}\frac{d\mu}{dT}\Big] \\
&= \frac{1}{T}\int d\epsilon D(\epsilon)\frac{\partial f_{T,\mu}(\epsilon)}{\partial \mu}\Big[\epsilon-\mu+T\frac{d\mu}{dT}\Big]
\end{aligned}
\tag{4.79}
$$

Thus the second term in the specific heat vanishes completely so that

$$
c_V = \frac{1}{T}\int_{-\infty}^{\infty} d\epsilon\ D(\epsilon)\frac{\partial f_{T,\mu}(\epsilon)}{\partial \mu}\Big[\epsilon-\mu+T\frac{d\mu(T,V,N)}{dT}\Big]^2 \tag{4.80}
$$

where

$$
T\frac{d\mu}{dT} = -\frac{\int d\epsilon D(\epsilon)\frac{\partial f_{T,\mu}(\epsilon)}{\partial \mu}\Big(\epsilon-\mu\Big)}{\int d\epsilon D(\epsilon)\frac{\partial f_{T,\mu}(\epsilon)}{\partial \mu}} \tag{4.81}
$$

The function used to obtain the specific heat from an integration over the density of states is shown in Fig. 4.8.

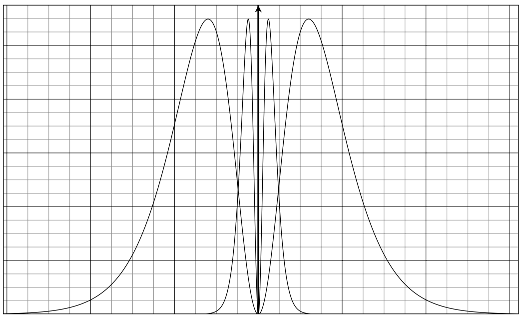


Fig. 4.8: Weight function $g(\epsilon) = \beta\frac{\partial f}{\partial \mu}(\epsilon-\mu)^2$ which yields the specific heat $c_V = k_B\int d\epsilon\ D(\epsilon)g(\epsilon)$ without a thermal shift of the chemical potential.

4.6 Jellium model

We use model systems to understand the qualitative behavior of a system before we attempt to describe a real system in all its complexity. For the electron gas, the model system is the **free electron gas** or the **jellium model**. The jellium model consists of electrons and a spatially constant, positive charge background, which ensures overall charge neutrality.

4.6.1 Dispersion relation

The **dispersion relation** of free electrons forms a parabola

$$\epsilon_{\vec{k},\sigma} = V_0 + \frac{(\hbar\vec{k})^2}{2m_e} \tag{4.82}$$

where V_0 is the value of the constant potential. The dispersion relation of electrons in a solid is often called the **band structure**.

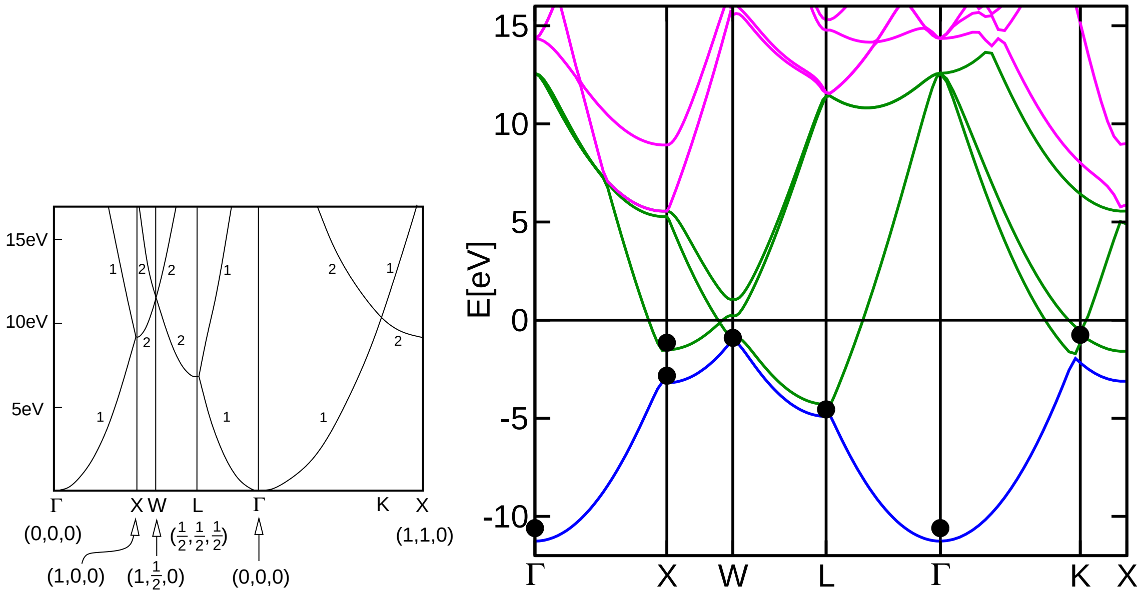


Fig. 4.9: Band structure of free, non-interacting electrons. The lattice is an fcc-cell with a lattice constant of 4.05 Å corresponding to aluminum. The high symmetry points are given in units of $\frac{2\pi}{a_{lat}}$. The numbers indicate the degeneracy beyond spin-degeneracy. Below the band structure of aluminum calculated by density functional theory is shown in comparison.

In Fig. 4.9, the band structure of free and non-interacting electron gas is compared to that of a real material, namely aluminium. We observe that the free electron gas already provides a fairly good description of some realistic systems.

The jellium model approximates real metals surprisingly well, even though the potential of the nuclei is far from being a constant. In a hand-waving manner we can say that the valence electrons are expelled from the nuclear region by the Pauli repulsion of the core electrons, so that they move around in a region with fairly constant potential.

4.6.2 Boundary conditions

To describe the wave functions in a solid one usually chooses periodic boundary conditions. Periodic boundary conditions are sometimes considered abstract and at times the connection to large but finite systems is not immediately evident. This is the reason for contrasting periodic boundary conditions with those of a finite system.

Particles in a box

Consider N electrons in a box with side-lengths L_x, L_y, L_z in the three spatial directions.

$$\left[\frac{\hat{\vec{p}}^2}{2m_e}+V(\hat{\vec{x}})-\epsilon\right]|\phi\rangle=0 \tag{4.83}$$

where the potential

$$V(\vec{r})=\begin{cases}0 & \text{for } \vec{r}\in\Omega\\ \infty & \text{else}\end{cases} \tag{4.84}$$

vanishes inside the box Ω, while the box is enclosed by impenetrable walls. The box represents a macroscopic crystal, which contains many unit cells of the material. The boundaries of the box are the surfaces.

The one-particle eigenstates of the Hamiltonian are

$$\phi_{\vec{k},\sigma}(\vec{r},\sigma')=\langle\vec{r},\sigma'|\phi_{\vec{k},\sigma}\rangle=\begin{cases}\delta_{\sigma,\sigma'}\sqrt{\frac{8}{\Omega}}\sin(k_xx)\sin(k_yy)\sin(k_zz) & \text{for } \vec{r}\in\Omega\\ 0 & \text{else}\end{cases} \tag{4.85}$$

where

$$\vec{k}_{i_1,i_2,i_3}=\left(\frac{\pi}{L_x}i_1,\frac{\pi}{L_y}i_2,\frac{\pi}{L_z}i_3\right)\text{ with } i_1,i_2,i_3=1,2,\ldots \tag{4.86}$$

We denote the size $\Omega=L_xL_yL_z$ of the box with the same symbol as the box itself.

The energy of the state is its kinetic energy

$$\epsilon_{\sigma,\vec{k}}=\frac{\hbar^2\vec{k}^2}{2m_e} \tag{4.87}$$

Note here, that the lowest energy eigenvalue does not vanish. It is $\epsilon_{\sigma,1,1,1}=\frac{\hbar^2\pi^2}{2m_e}\left(\frac{1}{L_x^2}+\frac{1}{L_y^2}+\frac{1}{L_z^2}\right)$. This is a consequence of Heisenberg's uncertainty relation[9] This is noteworthy, because using periodic boundary conditions, the wave functions are extended and the lowest state has energy zero. This becomes relevant for the Bose-Einstein condensation, where the thermodynamic limit $L_i\to\infty$ and the zero-temperature limit are taken simultenously. If the thermodynamic limit is taken first, the result is paradoxical.

For a given spin, each state occupies a volume in reciprocal space, namely

$$\Delta k_x\Delta k_y\Delta k_Z=\frac{\pi^3}{L_xL_yL_z} \tag{4.88}$$

In the limit of large boxes, it is convenient to replace the sum over states by an integral. Let us

---

[9] Heisenberg's uncertainty relation says that position and momentum cannot be certain at the same time. A confined space, such as the box, therefore leads to a spread of momenta, which implies that the expectation value of the momentum cannot vanish. This implies that the kinetic energy does not vanish. The wave functions of the particle in the box are no momentum eigenstates, because they are not extended over all space.

evaluate the expectation value of a one-particle operator $\hat{A}$. It is given

$$\begin{aligned}
\langle A\rangle &= \sum_\sigma \sum_{i_x,i_y,i_z=1}^{\infty} f_{\sigma,k_{ix,iy,iz}}\langle\phi_{\sigma,\vec{k}_{ix,iy,iz}}|\hat{A}|\phi_{\sigma,\vec{k}_{ix,iy,iz}}\rangle \\
&= \sum_\sigma \frac{\Omega}{\pi^3} \underbrace{\sum_{i_x,i_y,i_z=1}^{\infty} \Delta k_x \Delta k_y \Delta k_z}_{\rightarrow\int_{k_i>0} d^3k}\ f_{\sigma,k_{ix,iy,iz}}\langle\phi_{\sigma,\vec{k}_{ix,iy,iz}}|\hat{A}|\phi_{\sigma,\vec{k}_{ix,iy,iz}}\rangle \\
&\xrightarrow{L_x,L_y,L_z\rightarrow\infty} \sum_\sigma \frac{\Omega}{\pi^3}\int_{k_i>0} d^3k\ f_{\vec{k},\sigma}\langle\phi_{\vec{k},\sigma}|\hat{A}|\phi_{\vec{k},\sigma}\rangle
\end{aligned} \tag{4.89}$$

Here the wave vectors $\vec{k}$ cover only the all-positive section of k-space. That is, only one-eighth of the k-space is occupied. if a component of the wave vector is reversed, the wave function reverses its sign, but it is still the same quantum state. Thus we may formally extend the integration over all space, but we divide the integrand by eight. That is

$$\langle A\rangle \xrightarrow{L_x,L_y,L_z\rightarrow\infty} \sum_\sigma \int \Omega\frac{d^3k}{(2\pi)^3}\ f_{\vec{k},\sigma}\langle\phi_{\vec{k},\sigma}|\hat{A}|\phi_{\vec{k},\sigma}\rangle \tag{4.90}$$

This construction appears artificial at first, but it will become more natural, when we study infinite solids.

Periodic boundary conditions

For a solid in its pure sense the description of particles in a box is inconvenient, because the surfaces distract from the essence of the bulk material. One of the inconveniencies is that the eigenstates are standing waves and not propagating waves. The ability to act as medium for currents is however essential for the description.[10]

Therefore, one has introduced periodic boundary conditions. This implies that we connect the right side of the box with the left side, the back with the front and the bottom with the top. In other words, we consider solutions that extend over all space, but they repeat each other after a certain distance L in each direction. This implies

======================

Because of the translational symmetry, the potential is spatially constant. Thus, the one-particle states are simply plane waves.

$$\langle\vec{r},\sigma'|\phi_{\vec{k},\sigma}\rangle = \phi_{\vec{k},\sigma}(\vec{r},\sigma') = \frac{1}{\sqrt{\Omega}}\mathrm{e}^{i\vec{k}\vec{r}}\delta_{\sigma,\sigma'} \tag{4.91}$$

We consider here states in a very large, but finite, volume Ω such as the universe. The states are normalized to one within this volume.

This is an unusual notation for an ideally infinite system. Using this notation we avoid having to distinguish discrete and continuous sections of the spectrum and different normalization conditions for localized and extended states. A disadvantage is that extended states have a nearly vanishing amplitude, and that the spacing in k-space is extremely small. For a finite volume, only discrete values for $\vec{k}$ are allowed. We will later see that in the final expression the factors Ω from the normalization can be translated into a volume element in k-space, so that the sums can be converted into integrals.

Now we face the problem, that the energy spectrum for the infinite electron gas is no more discrete, but continuous. Therefore, let us study how to make the transition to an infinite system. There are several different ways to do this transition. One is to enclose the electron gas into a

---

[10] Such a description is also possible using particles in the box, but requires extra measures such as alternating and spatially varying currents.

box and then form the limes for an infinite box. The other procedure is to use **periodic boundary conditions**[11]. In the periodic boundary condition we impose that the wave functions are periodic[12] with repeat units L_x, L_y, L_z for the three spatial directions $\vec{e}_x, \vec{e}_y, \vec{e}_z$. $\vec{e}_i$ are the unit vectors in the three spatial directions. The periodic boundary conditions imply

$$\phi(\vec{r}+\vec{e}_xL_x)=\phi(\vec{r}) \qquad \phi(\vec{r}+\vec{e}_yL_y)=\phi(\vec{r}) \qquad \phi(\vec{r}+\vec{e}_zL_z)=\phi(\vec{r})$$

As a consequence of the periodic boundary conditions, only certain wave vectors are allowed, namely those fulfilling $k_iL_i=2\pi n_i$, where the n_i are arbitrary integers. Thus, there are only states with $\vec{k}=\left(\frac{2\pi}{L_x}i,\frac{2\pi}{L_y}j,\frac{2\pi}{L_z}k\right)$ and the density of states can be written as for a finite system. Then, we can make the transition $L_i\to\infty$. We will use the volume of the repeat box $\Omega\overset{\text{def}}{=}L_xL_yL_z$.

$$\begin{aligned}
D(\epsilon_0)&=\sum_\sigma\sum_{i,j,k=-\infty}^{\infty}\delta\Big(\epsilon(\vec{k}_{i,j,k},\sigma)-\epsilon_0\Big)\\
&=\sum_\sigma\sum_{i,j,k=-\infty}^{\infty}\underbrace{\overbrace{\frac{(2\pi)^3}{L_xL_yL_z}}^{d^3k}\frac{\Omega}{(2\pi)^3}}_{=1}\delta\Big(\epsilon(\vec{k}_{i,j,k},\sigma)-\epsilon_0\Big)\\
&=\frac{\Omega}{(2\pi)^3}\sum_\sigma\int d^3k\;\delta\Big(\epsilon(\vec{k},\sigma)-\epsilon_0\Big)
\end{aligned}\tag{4.92}$$

Thus, we replaced the sum over states by an integral over k-space.

Note, that the density of states is defined as an extensive quantity[13], that is, it is proportional to the volume Ω. This is understandable, because the energy levels come closer together as the system size is increased until the discrete spectrum is converted at infinite volume into a continuous spectrum. A little inconvenient is that the density of states is multiplied with an, in principle, infinite volume. This can be avoided by defining the **density of states per volume** $g(\epsilon)$

$$g(\epsilon)\overset{\text{def}}{=}\frac{1}{\Omega}D(\epsilon)\tag{4.93}$$

Now we can calculate the density of states for the free, non-interacting electron gas

$$\begin{aligned}
D(\epsilon_0)&\overset{Eq.\ 4.92}{=}\frac{\Omega}{(2\pi)^3}\underbrace{\sum_\sigma}_{2}\int d^3k\;\delta\left(\frac{(\hbar\vec{k})^2}{2m_e}-\epsilon_0\right)\\
&=\frac{2\Omega}{(2\pi)^3}\frac{d}{d\epsilon_0}\int d^3k\;\theta\left(\epsilon_0-\frac{(\hbar\vec{k})^2}{2m_e}\right)\\
&=\frac{2\Omega}{(2\pi)^3}\frac{d}{d\epsilon_0}\frac{4\pi}{3}\left(\frac{\sqrt{2m_e\epsilon_0}}{\hbar}\right)^3=\frac{2\Omega}{(2\pi)^3}\frac{4\pi}{3}\left(\frac{\sqrt{2m_e}}{\hbar}\right)^3\frac{3}{2}\epsilon_0^{\frac{1}{2}}\\
&=\frac{4\pi\Omega\,(2m_e)^{\frac{3}{2}}}{(2\pi\hbar)^3}\epsilon_0^{\frac{1}{2}}
\end{aligned}$$

When comparing with similar results, note, that this result already contains the spin degeneracy.

Next we wish to determine the Fermi energy as function of the electron density. We will use the

---

[11] For periodic boundary conditions see also ΦSX: Statistical Physics, Chapter: Non-interacting Particles.

[12] Here, the periodicity is not related to the lattice constant of the crystal, which in our case, is the jellium.

[13] The notion of extensive and intensive quantities originates from thermodynamics. An extensive quantity scales with the size of the system, while an intensive does not.

Fermi momentum defined by radius of the sphere of occupied states in reciprocal space.

$$\begin{aligned}
n &\overset{Eq.\,4.93}{=} \frac{1}{\Omega}\int_{-\infty}^{\epsilon_F} d\epsilon\, D(\epsilon) \\
&\overset{Eq.\,4.92}{=} \frac{1}{\Omega}\int_{-\infty}^{\epsilon_F} d\epsilon \frac{\Omega}{(2\pi)^3}\sum_\sigma \int d^3k\, \delta(\frac{(\hbar k)^2}{2m_e}-\epsilon) \\
&= \frac{2}{(2\pi)^3}\int d^3k \int_{-\infty}^{\epsilon_F} d\epsilon\, \delta\left(\frac{(\hbar k)^2}{2m_e}-\epsilon\right) \\
&= \frac{2}{(2\pi)^3}\int d^3k\, \theta\left(\epsilon_F-\frac{(\hbar k)^2}{2m_e}\right) \\
&= 2\cdot \underbrace{\int_{|k|<k_F}\frac{d^3k}{(2\pi)^3}}_{\frac{1}{(2\pi)^3}\frac{4\pi}{3}k_F^3} \\
\Rightarrow n &= 2\cdot\frac{1}{(2\pi)^3}\frac{4\pi}{3}k_F^3 && (4.94) \\
\Rightarrow k_F &= \sqrt[3]{3\pi^2 n} && (4.95)
\end{aligned}$$

Note that this expression for the density already includes the factor 2 from the spin degeneracy.

A short excursion: It is worth remembering the form of the **Brillouin-zone integral**

$$\langle\hat{A}\rangle = \Omega\sum_\sigma\int_{\epsilon(\vec{k})<\epsilon_F}\frac{d^3k}{(2\pi)^3}A_\sigma(\vec{k}) \tag{4.96}$$

where $A_\sigma(\vec{k}) := \langle\psi_\sigma(\vec{k})|\hat{A}|\psi_\sigma(\vec{k})\rangle$ is the expectation value of a one-particle operator with a spin state for a given Bloch vector.

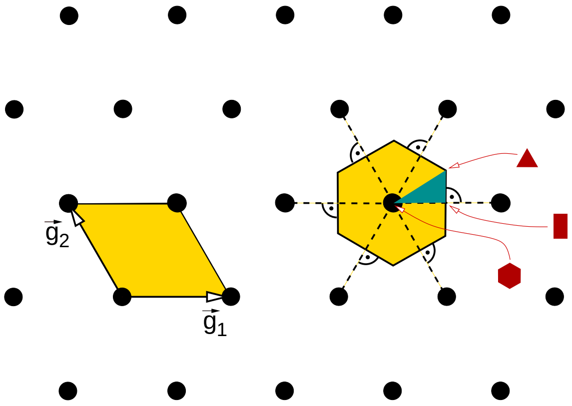


Fig. 4.10: Illustration of the Brillouin zone and high-symmetry points for a plane hexagonal lattice. On the left the unit cell of the reciprocal lattice is shown. On the right the Brillouin zone is shown, which consists of all points that are closer to one lattice point than to any other. In the Brillouin zone there is the irreducible zone, which contains all points that are symmetry inequivalent. The corners of the irreducible zone are high-symmetry points.

With the Fermi momentum from Eq. 4.95, we can determine the Fermi energy

$$\epsilon_F = \frac{(\hbar k_F)^2}{2m_e} = \frac{(3\pi^2\hbar^3)^{\frac{2}{3}}}{2m_e}n^{\frac{2}{3}}$$

While the Fermi energy of the free electron gas depends on the mass, the potential and the electron-electron interaction, the Fermi-momentum of a free electron gas only depends on the density. The reason is that the dispersion relation of a free electron gas is isotropic, and the volume of the volume inside the Fermi surface is determined by the number of states.

The density of the electron gas is often represented by a r_s, which is the radius of a sphere which occupies the mean volume per electron. If the electron density is $n=\frac{N}{\Omega}$, we obtain

$$n=\frac{1}{\frac{4\pi}{3}r_s^3} \qquad \Rightarrow \qquad r_s=\sqrt[3]{\frac{3}{4\pi n}} \tag{4.97}$$

The **electron-gas parameter** or **Seitz radius**[14] r_s is in common use.

Thus, we can express the Fermi momentum by r_s

$$k_F=\sqrt[3]{\frac{3\pi^2}{\frac{4\pi}{3}r_s^3}}=\sqrt[3]{\frac{9\pi}{4}}\frac{1}{r_s}$$

4.7 Properties of the electron gas

4.7.1 Fermi velocity

4.8 Tour trough band structures and densities of states of real materials

4.8.1 Free-electron like metals

The free electron like metals are the main group elements with a valence of one, two and three. This includes the **alkali metals**, the **alkaline earth** metals and tri-valent metals such als Aluminium.

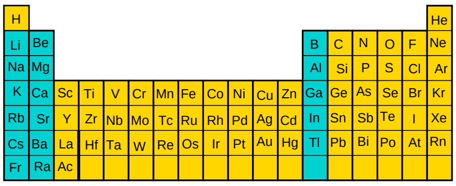


Fig. 4.11: Location of free-electron like metals in the periodic table marked in green.

The band structure of these materials is very similar to that of the free electron gas. The main difference is the lifted degeneracies at the zone boundaries, where the free-electron bands cross each other.

The valence of the element determines the number of electrons per unit cell, and thus the position of the Fermilevel.

The density of states shows the typical square-root-like behavior of the free electron gas.

---

[14] I found the name Seitz radius in [32].

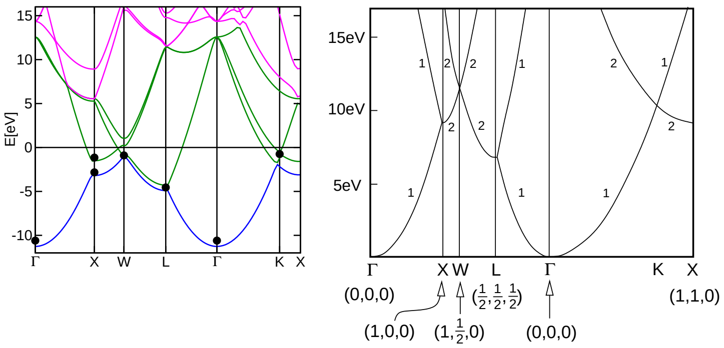
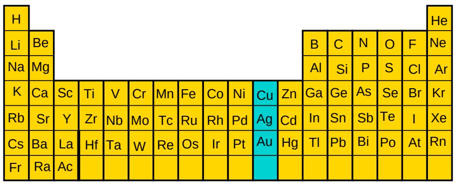


Fig. 4.12: Band structure of the free-electron-like metal aluminium (left) and , for comparison, the free-electron gas in an fcc-cell with a lattice constant of 4.05 Å corresponding to aluminum. The high symmetry points are given in units of $\frac{2\pi}{a_{lat}}$. The numbers indicate the degeneracy beyond spin-degeneracy. Below the band structure of aluminum is shown in comparison. The band structure of Al has been calculated with density-functional theory using the CP-PAW code. This figure is identical to4.9.

4.8.2 Noble metals



H																	He
Li	Be											B	C	N	O	F	Ne
Na	Mg											Al	Si	P	S	Cl	Ar
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
Fr	Ra	Ac															

Fig. 4.13: Location of noble metals in the periodic table marked in green.

The noble metals are free-electron like, but their valence band structure is interrupted by a narrow band of the d-electrons.

4.8.3 Transition metals

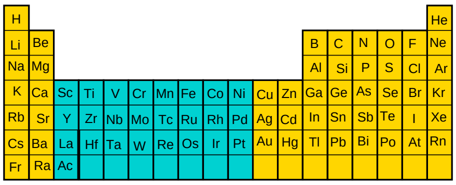


Fig. 4.14: Location of noble metals in the periodic table marked in green.

The band structure of the transition metals is determined by the narrow d-bands. Below and above the the d-states, one finds the free-electron like s-p bands. At the Fermi level, these free-electron bands are not important.

The transition metals form either closed-packed lattices such as fcc and hcp[15]. However in the middle of the transition metal series the bcc-lattice is dominant. The origin of this behavior can be understood from the density of states. The d-denisty of states in the closed packed lattices is block-like. In contrast, the density of states of d-electrons in the body-centered lattice consists of an upper and a lower parts with high density of states, which are separated by a region of low density of states. This quasi-band gap results in a lowering of the lower half of the states. As a consequence this structure is favorable when the Fermi-level is in the middle, the quasi band gap.

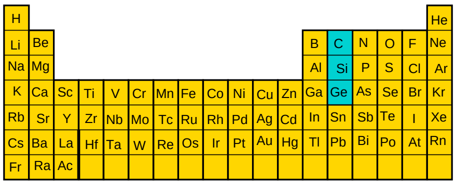
2) iron, bcc vs fcc, ferrite vs austenite, magnetism

---

[15] fcc denotes the face-centered cubic lattice, hcp denotes the hexagonal closed-packed lattice, and bcc stands for body-centered lattice.

4.8.4 Diamond, Zinc-blende and wurtzite-type semiconductors

The typical semiconductor materials have the diamond structure with four-fold coordinated atoms in a tetrahedral arrangement. This is the structure of the elemental semiconductors diamond, silicon and germanium.



H																	He
Li	Be											B	C	N	O	F	Ne
Na	Mg											Al	Si	P	S	Cl	Ar
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
Fr	Ra	Ac															

Fig. 4.15: Location of noble metals in the periodic table marked in green.

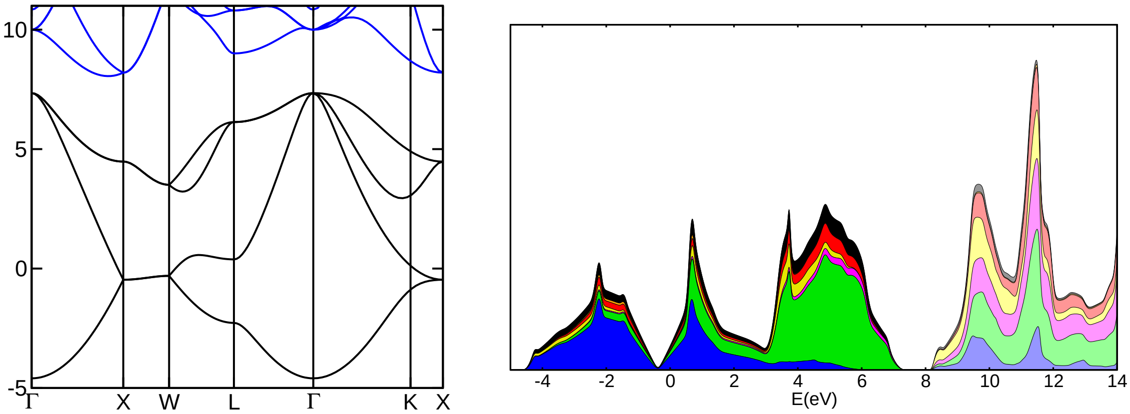


Fig. 4.16: Left: band structure of silicon. The valence bands (occupied states) are drawn in black and the conduction bands (unoccupied states) are drawn in blue. Right: density of states of siliucon. the colored areas correspond to the weight of s (blue), p (green) and d (red) electrons. The empty states (above +8-eV) are drawn in a lighter color. Based on density functional calculations using the CP-PAW code.

Besides the elemental semiconductors which are four-valent, there exists a variety of compound semiconductors such as GaAs, which consist, for example out of tri-valent and penta-valent ions. These compounds are named III-V compounds. There are also II-VI compounds such as ZnS and ZnSe.

Compound semiconductors such as GaAs distribute the ions onto the lattice sites of he diamond structure, which results in the zinc-blende structure.

However there is also a hexagonal variation of the diamond structure, the wurtzite lattice. Also in the Wurtzite lattice, the atoms are tetrahedrally coordinated. The difference to the diamond lattice becomes evident only in the second nearest neighbor shell. Therefore the properties in the cubic and the hexagonal variants are very similar.

4.8.5 Simple oxides

Oxides are a tremendously important class of materials. The are the major corrosion endproduct in our atmosphere, and thus the major component of rocks. They also include a wide range of functional materials such as the high-Tc super conductors. You look at them when you stare out of the window, you lie on them on the beach, because glass and sand are both the oxide of silicon. The most simple oxides are, as the name says, **simple oxides**. They are the oxides of the most electropositive elements, for which the cations do not contain transition metals, and for which the valence electrons are used up completely to for O^{-2} anions.

Simple oxides form typically a closed backed lattice of oxygen ions, with cations in the voids, between the anion spheres.

These materials are a good starting point for understanding the more complex oxides. One can simplify them even further, by making the same trick as for the jellium model: when we smear out the cation cores to a homogenous positive charge background and leave only the electrons to saturate the oxygen valence shell, we obtain, what I call **jellium oxide** or **jelly-O**. The oxygen ions form a closed packed lattice. The electronic structure of Jelly-O is almost indistinguishable to those of the simple oxides.

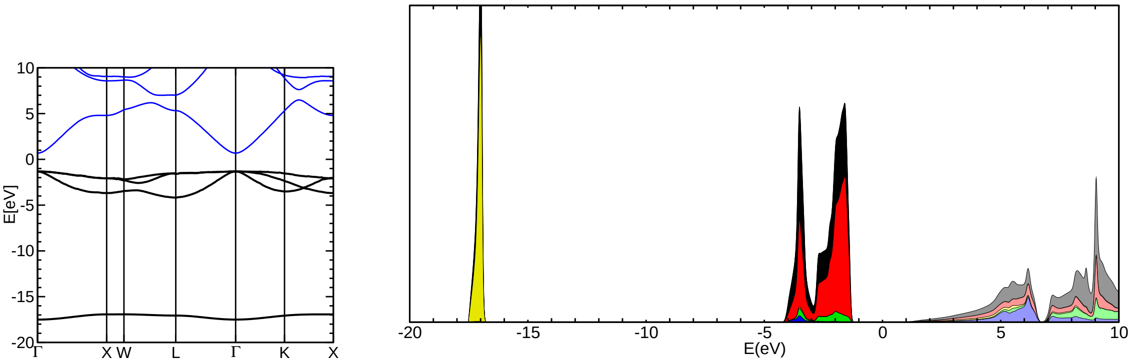


Fig. 4.17: Electronic structure of jellium oxide an fcc array of O^{2-} ions with a homogeneous positive charge background. Left: band structure. The filled bands are black, the empty ones are blue. Right: Density of states. Based on density functional calculations using the CP-PAW code.

The band structure is characterized by a low lying band of oxygen s-character. Above, and well separated from the s-band one finds the oxygen p-band which forms the valence shell. In contrast to the semiconductors, the s and p-bands are now well separated. This is, on the one hand, due to the large s-p splitting of the atom, and on the other hand it is due to the smaller band width of the oxygen p-band caused by the larger distance between oxygen ions.

Very characteristic for the oxygen valence band is the double-peaked structure. It is caused entirely by the hopping between the oxygen ions and is not to be mistaken by a bond to the cations.

The empty conduction band is very similar to a free-electron like band.

4.8.6 PCMO an example for a complex oxide

4.9 What band structures are good for

From ΦSX:Quantum mechanics of the chemical bond[30]

4.9.1 Chemical stability

4.9.2 Dispersion relations

The band structure tells us a lot about how particles move. the band structure is nothing but the dispersion relation for a particle. The dispersion relation is the relation of energy with momentum for a particle.

Using Hamilton's equation

$$\begin{aligned}\vec{v} &= \partial_t \vec{r} = \vec{\nabla}_p E(\vec{p})\\ \vec{F} &= \partial_t \vec{p} = -\vec{\nabla}_r E(\vec{p})\end{aligned}$$

- velocity of a particle is the slope of $E(p)$.
- momentum changes with time proportional to the force $\vec{F} = \partial_t \vec{p}$, which is Newton's law

Thus, a particle experiencing a constant force F, will have momentum

$$\begin{aligned}\vec{p}(t) &= \vec{p}(0) + \vec{F}t\\ \vec{v}(t) &= \vec{\nabla}_p E(\vec{p}(0) + \vec{F}t)\end{aligned}$$

Dynamics of wave packets

Abrupt interfaces

While a dispersion relation is strictly defined only for translationally invariant systems, where the energy eigenstates are also momentum eigenstates, we can also understand what happens at the interface between materials. As a particle passes from one material to the other, it may gain or loose momentum perpendicular to the interface, but we can use energy conservation and conservation of the momentum parallel to the interface to obtain some information on the interface crossing.

Slow spatial variation

Lifetime broadening

Sometimes the material has some randomness or other interactions that break translational symmetry. If these perturbations of translational symmetry are small, the eigenstates of the Hamiltonian are no more momentym eigenstates. If we still project the energy eigenstates onto momentum eigenstates, each momentum will obtain contribution from several energies. Instrad of sharp energies for each momentym we have a broadened distribution. The width of the energy distribution is a measure for the life time of a momentum: Scattering processes will redirect the particle.

4.9.3 Boltzmann equation

Demonstration of the principles in real space

Consider a particle conservation law of the form

$$\partial_t \rho(\vec{r},t) + \vec{\nabla}\vec{j}(\vec{r},t)$$

where $rho(\vec{r},t)$ is the particle density and $\vec{j}$ is the particle current.

The current of particles that follow a given velocity field $\vec{v}(\vec{r},t)$ is given by

$$\vec{j} = \rho(\vec{r},t)\vec{v}(\vec{r},t)$$

which brings the particle conservation law into the form

$$\partial_t \rho(\vec{r},t) + \vec{v}\vec{\nabla}\rho + \rho\left(\vec{\nabla}\vec{v}\right) = 0$$

If we add another assumption, namely the flow is incompressible, that is $\vec{\nabla}\vec{v} = 0$ we obtain

$$\partial_t \rho(\vec{r},t) + \vec{v}\vec{\nabla}\rho = 0 \tag{4.98}$$

From coordinate space to phase space

Now we translate these steps into phase space. That is we have particles characterized by position and momentum. The velocity field in momentym space is given by the Hamilton function

$$\begin{pmatrix} \vec{r} \\ \vec{p} \end{pmatrix} = \begin{pmatrix} \vec{\nabla}_p H(\vec{p},\vec{r}) \\ -\vec{\nabla}_r H(\vec{p},\vec{r}) \end{pmatrix}$$

Lioville's theorem says that the flow in phase space is incompressible.

$$\begin{pmatrix} \vec{\nabla}_r \\ \vec{\nabla}_p \end{pmatrix} \begin{pmatrix} \vec{\nabla}_p H(\vec{p},\vec{r}) \\ -\vec{\nabla}_r H(\vec{p},\vec{r}) \end{pmatrix} = \vec{\nabla}_r\vec{\nabla}_p H - \vec{\nabla}_p\vec{\nabla}_r H = 0$$

This allows to translate Eq. 4.98 into phase space, where it has the form

$$\partial_t f(\vec{r},\vec{p},t) + \underbrace{\vec{v}}_{\partial_t\vec{r}=\vec{\nabla}_p H} \vec{\nabla}_r f + \underbrace{\vec{F}}_{\partial_t\vec{p}=-\vec{\nabla}_r H} \vec{\nabla}_p f = 0$$

Thus, we have a description for the dynamics of a distribution of non-interacting particles.

Scattering: Boltzmann equation

Finally we may introduce the interaction between particles in the form of collisions yielding the Boltzmann equation

$$\partial_t f(\vec{r},\vec{p},t) + \underbrace{\vec{v}}_{\partial_t\vec{r}=\vec{\nabla}_p H} \vec{\nabla}_r f + \underbrace{\vec{F}}_{\partial_t\vec{p}=-\vec{\nabla}_r H} \vec{\nabla}_p f = \left(\frac{\partial f}{\partial t}\right)_{scatt} \tag{4.99}$$

The form of the scattering term depends on the physical interaction. Usually they are considered to be local: In a scattering event, two particles may collide at a given position in space.

Relaxation time approximation

The information on the scattering events is not contained in the band structure. A common and simple approximation is the relaxation time approximation, which is based on the following form of the scattering term

$$\left(\frac{\partial f}{\partial t}\right)_{scatt} = -\frac{1}{\tau}\left(f(\vec{r}, \vec{p}, t) - f^{eq}(\vec{r}, \vec{p})\right)$$

where τ is the relaxation time, and f^{eq} is the equilibrium distribution in the form of a Boltmann, Fermi or Bose distribution.

Chapter 5

Magnetism

5.1 Charged particles in a magnetic field

5.2 Dirac equation and magnetic moment

Spin-orbit coupling responsible for magnetic anisotropy.

- Non-relativistic magnetic moment $\vec{m}$ of a charge rotating about a center

$$\vec{m}_{cl} = \frac{q}{2m}\vec{L} \tag{5.1}$$

- Dirac equation defines the gyromagnetic ratio of $g = 2$.
- Quantum field theory leads to $g = 2.00232\ldots$

5.3 Magnetism of the free electron gas

Landau levels

5.4 Magnetic order

Stoner criterion, Pauli paramagnetism,

ferromagnetism, ferrimagnetism, antiferromagnetism, Ising model, Heisenberg model, Weiss regions.

Chapter 6

The Hartree-Fock approximation and exchange

The **Hartree-Fock method**[28, 33, 34] is an electronic structure method that is the work-horse of quantum chemistry. Today it plays an important role as a starting point of more accurate and involved methods. The Hartree-Fock method had a predecessor, the **Hartree method**[35, 36], which, however, does not play an important role in practice.

The basic idea of the Hartree-Fock method is to restrict the wave functions to single Slater determinants[28]. For this Slater determinant, the energy is determined as expectation value of the true many-particle Hamiltonian. That is, the Hartree-Fock approximation takes the complete electron-electron interaction into account. Furthermore the one-particle orbitals are optimized to yield the lowest energy.

One can easily show that the Hartree-Fock energy always provides an upper bound for the energy. The Hamilton operator is, up to a constant, positive definite and the ground state is the state with the lowest energy. As we restricted our wave function to single Slater determinants, the wave function is, normally, not the ground state and therefore higher in energy than the ground-state energy.

After having defined Slater determinants in the previous section, in this section we will gain some familiarity with them. We will also see why systems of non-interacting many-particle problems can be described by a set of one-particle wave functions alone, something that we implicitely assumed in the section about non-interacting electrons. Furthermore, we will become familiar with the most important contributions of the electron correlation, namely Hartree energy and exchange.

6.1 One-electron and two-particle operators

The Born-Oppenheimer Hamiltonian Eq. 2.4 for an N-electron system has the form

$$E = E_{NN} + \langle\Psi|\sum_{i=1}^{N}\hat{h}_i + \frac{1}{2}\sum_{i\neq j}\hat{W}_{i,j}|\Psi\rangle$$

Where

$$E_{NN}(\vec{R}_1,\dots,\vec{R}_M)\stackrel{\text{def}}{=}\frac{1}{2}\sum_{i\neq j}^{M}\frac{e^2Z_iZ_j}{4\pi\epsilon_0|\vec{R}_i-\vec{R}_j|} \tag{6.1}$$

is the electrostatic repulsion between the nuclei, and the operators acting on the electrons can be divided into one-particle terms $\hat{h}_i$ and two-particle terms $\hat{W}_{i,j}$.

The one-particle terms describe the kinetic and potential energy of the electrons in an external potential $v_{ext}(\vec{r})$

$$\hat{h}_i = \int d^4x_1 \cdots \int d^4x_N \, |\vec{x}_1, \ldots, \vec{x}_N\rangle \left[\frac{-\hbar^2}{2m_e}\vec{\nabla}^2_{\vec{r}_i} + v_{ext}(\vec{r}_i)\right] \langle\vec{x}_1, \ldots, \vec{x}_N| \tag{6.2}$$

The external potential describes the electrostatic attraction between electrons and nuclei.

$$v_{ext}(\vec{r}; \vec{R}_1, \ldots, \vec{R}_M) = -\sum_{j=1}^{M} \frac{e^2 Z_j}{4\pi\epsilon_0|\vec{r} - \vec{R}_j|} \tag{6.3}$$

The two-particle operator describes the interaction between the electrons

$$\hat{W}_{i,j} = \int d^4x_1 \cdots \int d^4x_N \, |\vec{x}_1, \ldots, \vec{x}_N\rangle \frac{e^2}{4\pi\epsilon_0|\vec{r}_i - \vec{r}_j|} \langle\vec{x}_1, \ldots, \vec{x}_N| \tag{6.4}$$

The indices on $\hat{h}_i$ and $\hat{W}_{i,j}$ determine, onto which electron coordinates the operator acts.

The **one-particle operators** act on the coordinates of one particle at a time. The kinetic energy and the external potential are examples for one-particle operators, because they can be determined for each particle individually and then be summed up. The interaction energy Eq. 6.4 on the other hand depends on the coordinates of two particles simultaneously and is therefore a true **two-particle operator**. It is the two-particle term that is the cause for the dazzling complexity of many-particle physics.

Let us consider a general operator in real space

$$\begin{aligned}
\hat{A} &= \underbrace{\int d^4x_1 \cdots \int d^4x_N |\vec{x}_1, \ldots, \vec{x}_N\rangle\langle\vec{x}_1, \ldots, \vec{x}_N|}_{\hat{1}} \\
&\qquad \cdot\hat{A} \underbrace{\int d^4x'_1 \cdots \int d^4x'_N \, |\vec{x'}_1, \ldots, \vec{x'}_N\rangle\langle\vec{x'}_1, \ldots, \vec{x'}_N|}_{\hat{1}} \\
&= \int d^4x_1 \cdots \int d^4x_N \int d^4x'_1 \cdots \int d^4x'_N \, |\vec{x}_1, \ldots, \vec{x}_N\rangle \\
&\qquad \cdot\langle\vec{x}_1, \ldots, \vec{x}_N|\hat{A}|\vec{x'}_1, \ldots, \vec{x'}_N\rangle\langle\vec{x'}_1, \ldots, \vec{x'}_N|
\end{aligned}$$

A general matrix element has therefore $2N$ arguments.

- If the matrix element of an operator has the special form

  $$\langle\vec{x}_1, \ldots, \vec{x}_N|\hat{A}|\vec{x'}_1, \ldots, \vec{x'}_N\rangle = \sum_{i=1}^{N} A(\vec{x}_i, \vec{x}'_i) \prod_{j\neq i} \delta(\vec{x}_j - \vec{x'}_j)$$

  we call the operator a **one-particle operator**.

- If furthermore the function $A(\vec{x}, \vec{x'})$ has the special form that it is diagonal in the primed and unprimed arguments, i.e.

  $$A(\vec{x}, \vec{x'}) = a(\vec{x})\delta(\vec{x} - \vec{x'})$$

  we call the one-particle operator **local**.

- If the matrix element of an operator has the special form

  $$\langle\vec{x}_1, \ldots, \vec{x}_N|\hat{A}|\vec{x'}_1, \ldots, \vec{x'}_N\rangle = \frac{1}{2}\sum_{i,j=1}^{N} A(\vec{x}_i, \vec{x}_j, \vec{x}'_i, \vec{x'}_j) \prod_{k\neq\{i,j\}} \delta(\vec{x}_k - \vec{x'}_k)$$

  we call the operator a **two-particle operator**.

- If furthermore the function $A(\vec{x}_1,\vec{x}_2,\vec{x'}_1,\vec{x'}_2)$ has the special form that it is diagonal in the primed and unprimed arguments, i.e.

  $$A(\vec{x}_1,\vec{x}_2,\vec{x'}_1,\vec{x'}_2)=a(\vec{x}_1,\vec{x}_2)\delta(\vec{x}_1-\vec{x'}_1)\delta(\vec{x}_2-\vec{x'}_2)$$

  we call the two-particle operator **local**.

6.2 Expectation values of Slater determinants

Let us now work out the expectation values for the one-particle and two-particle operators for a Slater determinant.

6.2.1 Expectation value of a one-particle operator

Here, we will work out the expectation value of a one-particle operator for a Slater determinant. An example for such a one-particle operator is the non-interacting part of the Hamiltonian.

Explicit example for the two-particle wave function

The two-particle Slater determinant of two one-particle orbitals $|\phi_a\rangle$ and $|\phi_b\rangle$ has the form

$$\underbrace{\langle\vec{r}_1\vec{r}_2|\Psi\rangle}_{\Psi(\vec{r}_1,\vec{r}_2)}=\frac{1}{\sqrt{2}}\underbrace{\Big[\langle\vec{r}_1|\phi_a\rangle\langle\vec{r}_2|\phi_b\rangle-\langle\vec{r}_1|\phi_b\rangle\langle\vec{r}_2|\phi_a\rangle\Big]}_{\phi_a(\vec{r}_1)\phi_b(\vec{r}_2)-\phi_b(\vec{r}_1)\phi_a(\vec{r}_2)} \tag{6.5}$$

To avoid unnecessary complication we drop the spin-coordinates, that is we consider spin-less fermions.

As an specific example for a one-particle operator, we consider the particle-density operator $\hat{n}(\vec{r})$ defined as

$$\hat{n}(\vec{r})=\int d^3r_1\int d^3r_2\,|\vec{r}_1,\vec{r}_2\rangle\Big[\delta(\vec{r}-\vec{r}_1)+\delta(\vec{r}-\vec{r}_2)\Big]\langle\vec{r}_1,\vec{r}_2|$$

The electron density is

$$
\begin{aligned}
n(\vec r) &= \langle\Psi|\hat n(\vec r)|\Psi\rangle \\
&= \int d^3r_1 \int d^3r_2\ \Psi^*(\vec r_1,\vec r_2)\Big[\delta(\vec r-\vec r_1)+\delta(\vec r-\vec r_2)\Big]\Psi(\vec r_1,\vec r_2) \\
&\overset{\Psi(\vec r_1,\vec r_2)=-\Psi(\vec r_2,\vec r_1)}{=} \int d^3r_1 \int d^3r_2\ \Psi^*(\vec r_1,\vec r_2)\delta(\vec r-\vec r_1)\Psi(\vec r_1,\vec r_2) \\
&\qquad + \int d^3r_1 \int d^3r_2\ \Psi^*(\vec r_2,\vec r_1)\delta(\vec r-\vec r_2)\Psi(\vec r_2,\vec r_1) \\
&\overset{\vec r_1\leftrightarrow\vec r_2}{=} 2\int d^3r_1 \int d^3r_2\ \Psi^*(\vec r_1,\vec r_2)\delta(\vec r_1-\vec r)\Psi(\vec r_1,\vec r_2) \\
&\overset{\text{Eq. 6.5}}{=} 2\int d^3r_1 \int d^3r_2\ \frac{1}{\sqrt{2}}\Big[\phi_a(\vec r_1)\phi_b(\vec r_2)-\phi_a(\vec r_2)\phi_b(\vec r_1)\Big]^* \\
&\qquad \cdot\,\delta(\vec r-\vec r_1)\frac{1}{\sqrt{2}}\Big[\phi_a(\vec r_1)\phi_b(\vec r_2)-\phi_a(\vec r_2)\phi_b(\vec r_1)\Big] \\
&= \int d^3r_1 \int d^3r_2\ \phi_a^*(\vec r_1)\phi_b^*(\vec r_2)\delta(\vec r-\vec r_1)\phi_a(\vec r_1)\phi_b(\vec r_2) \\
&- \int d^3r_1 \int d^3r_2\ \phi_a^*(\vec r_1)\phi_b^*(\vec r_2)\delta(\vec r-\vec r_1)\phi_a(\vec r_2)\phi_b(\vec r_1) \\
&- \int d^3r_1 \int d^3r_2\ \phi_a^*(\vec r_2)\phi_b^*(\vec r_1)\delta(\vec r-\vec r_1)\phi_a(\vec r_1)\phi_b(\vec r_2) \\
&+ \int d^3r_1 \int d^3r_2\ \phi_a^*(\vec r_2)\phi_b^*(\vec r_1)\delta(\vec r-\vec r_1)\phi_a(\vec r_2)\phi_b(\vec r_1) \\
&= \langle\phi_a|\vec r\rangle\langle\vec r|\phi_a\rangle\underbrace{\langle\phi_b|\phi_b\rangle}_{=1}-\langle\phi_a|\vec r\rangle\langle\vec r|\phi_b\rangle\underbrace{\langle\phi_b|\phi_a\rangle}_{=0} \\
&\qquad -\langle\phi_b|\vec r\rangle\langle\vec r|\phi_a\rangle\underbrace{\langle\phi_a|\phi_b\rangle}_{=0}+\langle\phi_b|\vec r\rangle\langle\vec r|\phi_b\rangle\underbrace{\langle\phi_a|\phi_a\rangle}_{=1} \\
&= \phi_a^*(\vec r)\phi_a(\vec r)+\phi_b^*(\vec r)\phi_b(\vec r)
\end{aligned}
$$

Let us mention a few observations:

- The number of terms is drastically reduced just because we have used an orthonormal set of one-particle wave functions. This is the sole reason for using orthonormal basissets in many-particle physics.
- The sum over particles is turned into a sum over orbitals. This is true for a Slater determinant, but not for a general many particle states.
- The antisymmetry of the wave function ensures that the same expectation value is obtained if we work out a property of the first or the second electron.[1]

General derivation

Now, we will determine the same result in its full generality: First, we evaluate only the matrix element of the one-particle operator $\hat h_1$, that acts only on the coordinates of only one particle, namely the first. This result is then generalized for the other particles and then summed up. Thus, we obtain the expectation value of the non-interacting part of the Hamiltonian.

---

[1] The wave function is, up to a sign, the same when the coordinates of the coordinates of the first and the second electron are interchanged. The sign drops out for expectation values, because the wave function enters twice. Example: Using the antisymmetry of the wave function $\psi(x,x')=-\psi(x',x)$ one can convert the density ρ_1 of the first particle $\rho_1(x)=\int dx'\ \psi^*(x,x')\psi(x,x')=(-1)^2\int dx'\ \psi^*(x',x)\psi(x',x)=\rho_2(x)$ into the density $\rho_2(x)$ of the second.

To be concise, we write here the detailed expressions for the matrix elements of the one-particle Hamilton operator, we have in mind

$$\begin{aligned}\langle\phi|\hat{h}|\phi\rangle &= \overbrace{\int d^4x \int d^4x'\, \langle\phi|\vec{x}\rangle\, \overbrace{\langle\vec{x}|\hat{h}|\vec{x}'\rangle}\, \langle\vec{x}'|\phi\rangle}^{\delta(\vec{r}-\vec{r}')\delta_{\sigma,\sigma'}\left(-\frac{\hbar^2}{2m_e}\vec{\nabla}^2+v(\vec{r})\right)} \\ &= \sum_\sigma \int d^3r\; \phi^*(\vec{r},\sigma)\left[\frac{-\hbar^2}{2m_e}\vec{\nabla}^2+v_{ext}(\vec{r})\right]\phi(\vec{r},\sigma) \qquad (6.6)\end{aligned}$$

The one-particle Hamiltonian $\hat{h}_j$ acting on the j-th particle in a many particle wave function $\Psi\rangle$ yields the expectation value

$$\langle\Psi|\hat{h}_j|\Psi\rangle = \int d^4x_1 \cdots \int d^4x_N\; \Psi^*(\vec{x}_1,\ldots,\vec{x}_N)\left(\frac{-\hbar^2}{2m_e}\vec{\nabla}_j^2+v_{ext}(\vec{r}_j)\right)\Psi(\vec{x}_1,\ldots,\vec{x}_N)$$

so that the matrix element for two product wave functions is

$$\langle\phi_1,\ldots,\phi_N|\hat{h}_j|\psi_1,\ldots,\psi_N\rangle = \langle\phi_1|\psi_1\rangle\cdots\langle\phi_j|\hat{h}|\psi_j\rangle\cdots\langle\phi_N|\psi_N\rangle \qquad (6.7)$$

The Slater determinant, given in Eq. 3.21, has the form

$$\langle\vec{x}_1,\ldots,\vec{x}_N|\Psi\rangle = \frac{1}{\sqrt{N!}}\sum_{i_1,\ldots,i_N=1}^{N}\epsilon_{i_1,\ldots,i_N}\langle\vec{x}_1|\phi_{i_1}\rangle\cdots\langle\vec{x}_N|\phi_{i_N}\rangle \qquad (6.8)$$

Thus, the matrix element of a one-particle operator Eq. 6.2 is

$$\begin{aligned}\langle\Psi|\sum_{i=1}^{N}\hat{h}_i|\Psi\rangle &= N\langle\Psi|\hat{h}_1|\Psi\rangle \\ &\overset{Eq.\ 6.8}{=} N\cdot\frac{1}{N!}\sum_{i_1,i_2,\ldots,i_N}^{N}\sum_{j_1,j_2,\ldots,j_N}^{N}\epsilon_{i_1,i_2,\ldots,i_N}\langle\phi_{i_1},\ldots,\phi_{i_N}|\hat{h}_1|\phi_{j_1},\ldots,\phi_{j_N}\rangle\epsilon_{j_1,j_2,\ldots,j_N} \\ &\overset{\text{Eq. 6.7}}{=} N\frac{1}{N!}\sum_{i_1,i_2,\ldots,i_N}^{N}\sum_{j_1,j_2,\ldots,j_N}^{N}\epsilon_{i_1,i_2,\ldots,i_N}\epsilon_{j_1,j_2,\ldots,j_N}\langle\phi_{i_1}|\hat{h}_1|\phi_{j_1}\rangle\underbrace{\langle\phi_{i_2}|\phi_{j_2}\rangle}_{\delta_{i_2,j_2}}\cdots\underbrace{\langle\phi_{i_N}|\phi_{j_N}\rangle}_{\delta_{i_N,j_N}} \\ &= N\frac{1}{N!}\sum_{i_1,j_1=1}^{N}\langle\phi_{i_1}|\hat{h}_1|\phi_{j_1}\rangle\underbrace{\sum_{i_2,\ldots,i_N}^{N}\epsilon_{i_1,i_2,\ldots,i_N}\epsilon_{j_1,i_2,\ldots,i_N}}_{\delta_{i_1,j_1}(N-1)!} \\ &= \sum_{j=1}^{N}\langle\phi_j|\hat{h}_1|\phi_j\rangle\end{aligned}$$

In the last step we exploited, that the sum over j_1 only contributes when $j_1 = i_1$. There is only one orbital left if all the orbitals but the first are determined. Thus, the first orbital must be the same for both Levi-Civita Symbols.

Thus, we obtain the expectation value of the one-particle Hamiltonian for a Slater determinant $|\Psi\rangle$ as

EXPECTATION VALUE OF A ONE-PARTICLE OPERATOR WITH A SLATER DETERMINANT

$$\langle\Psi|\sum_{i=1}^{N}\hat{h}_i|\Psi\rangle=\sum_{i=1}^{N}\langle\phi_i|\hat{h}|\phi_i\rangle \tag{6.9}$$

Here, $|\Psi\rangle$ is a N-particle Slater determinant built from the one-particle orbitals $|\phi_i\rangle$, and $\sum_{i=1}^{N}\hat{h}_i$ is a one-particle operator.

Expectation values of common one-particle operators with Slater determinants

Analogously, the expectation value of any one-particle operator with a Slater determinant can be represented by a sum over one-particle orbitals. Thus, we obtain the expression for the density

$$n(\vec{r})=\langle\Psi|\sum_{i=1}^{N}\sum_{\sigma}\Big(|\vec{r},\sigma\rangle\langle\vec{r},\sigma|\Big)_i|\Psi\rangle=\sum_{i=1}^{N}\sum_{\sigma}\phi_i^*(\vec{r},\sigma)\phi_i(\vec{r},\sigma) \tag{6.10}$$

where

$$\Big(|\vec{r}\rangle\langle\vec{r}|\Big)_i\stackrel{\text{def}}{=}\int d^4x_1\cdots\int d^4x_N\;|\vec{x}_1,\ldots,\vec{x}_N\rangle\delta(\vec{r}-\vec{r}_i)\langle\vec{x}_1,\ldots,\vec{x}_N|$$

The expectation value of the kinetic energy for a Slater determinant has the form

$$E_{kin}=\langle\Psi|\sum_{i=1}^{N}\frac{\hat{\vec{p}}_i^2}{2m_e}|\Psi\rangle=\sum_{i=1}^{N}\sum_{\sigma}\int d^3r\;\phi_i^*(\vec{r},\sigma)\frac{-\hbar^2}{2m_e}\vec{\nabla}^2\phi_i(\vec{r},\sigma) \tag{6.11}$$

6.2.2 Expectation value of a two-particle operator

Let us now turn to the two-particle term. We introduce the symbol

$$W(\vec{x},\vec{x}')=\frac{e^2}{4\pi\epsilon_0|\vec{r}-\vec{r}'|}$$

The interaction is independent of the spin index. Furthermore it is local in the two coordinates.[2]

We proceed as for the expectation value of a one-particle operator and relate the interaction energy between all particles to the interaction between the first two electrons. This is allowed because of the indistinguishability of the electrons: The interaction between any two electrons is identical to any

---

[2] A general two-particle operator has the form

$$\begin{aligned}\hat{W}_{1,2}&=\int dx_1\cdots\int dx_N\int dx'_1\cdots\int dx'_N|\vec{x}_1,\ldots,\vec{x}_N\rangle\langle\vec{x}_1,\ldots,\vec{x}_N|\hat{W}_{1,2}|\vec{x'}_1,\ldots,\vec{x'}_N\rangle\langle\vec{x'}_1,\ldots,\vec{x'}_N|\\&=\int dx_1\cdots\int dx_N\int dx'_1\int dx'_2|\vec{x}_1,\ldots,\vec{x}_N\rangle\langle\vec{x}_1,\ldots,\vec{x}_N|\hat{W}_{1,2}|\vec{x'}_1,\vec{x'}_2,\vec{x}_3,\ldots,\vec{x}_N\rangle\langle\vec{x'}_1,\vec{x'}_2,\vec{x}_3,\ldots,\vec{x}_N|\\&=\int dx_1\cdots\int dx_N\int dx'_1\int dx'_2|\vec{x}_1,\ldots,\vec{x}_N\rangle\langle\vec{x}_1,\vec{x}_2|\hat{W}_{1,2}|\vec{x'}_1,\vec{x'}_2\rangle\langle\vec{x'}_1,\vec{x'}_2,\vec{x}_3,\ldots,\vec{x}_N|\end{aligned}$$

Thus, a general two-particle operator is nonlocal in the two particle coordinates and therefore its matrix elements depends on four arguments.

other pair. There are $\frac{N(N-1)}{2}$ ordered pairs and therefore the interaction energy is

$$
\begin{aligned}
E_W &= \langle\Psi|\frac{1}{2}\sum_{i\neq j}^{N}\hat{W}_{i,j}|\Psi\rangle = \frac{N(N-1)}{2}\langle\Psi|\hat{W}_{1,2}|\Psi\rangle \\
&= \frac{N(N-1)}{2}\frac{1}{N!}\sum_{i_1,i_2,\dots,i_N}^{N}\sum_{j_1,j_2,\dots,j_N}^{N}\epsilon_{i_1,i_2,\dots,i_N}\epsilon_{j_1,j_2,\dots,j_N}\langle\phi_{i_1},\dots,\phi_{i_N}|\hat{W}_{12}|\phi_{j_1},\dots,\phi_{j_N}\rangle \\
&= \frac{1}{2\cdot(N-2)!}\sum_{i_1,i_2,\dots,i_N}^{N}\sum_{j_1,j_2,\dots,j_N}^{N}\epsilon_{i_1,i_2,\dots,i_N}\epsilon_{j_1,j_2,\dots,j_N}\langle\phi_{i_1},\phi_{i_2}|\hat{W}_{12}|\phi_{j_1},\phi_{j_2}\rangle\underbrace{\langle\phi_{i_3}|\phi_{j_3}\rangle}_{\delta_{i_3,j_3}}\cdots\underbrace{\langle\phi_{i_N}|\phi_{j_N}\rangle}_{\delta_{i_N,j_N}} \\
&= \frac{1}{2\cdot(N-2)!}\sum_{i_1,i_2,j_1,j_2}^{N}\langle\phi_{i_1},\phi_{i_2}|\hat{W}_{12}|\phi_{j_1},\phi_{j_2}\rangle\underbrace{\sum_{i_3,\dots,i_N}^{N}\epsilon_{i_1,i_2,i_3\dots,i_N}\epsilon_{j_1,j_2,i_3\dots,i_N}}_{(N-2)!\epsilon_{i_1,i_2}\epsilon_{j_1,j_2}\text{for } i_1,i_2\in\{j_1,j_2\}} \\
&= \frac{1}{2}\sum_{i_1,i_2}^{N}\sum_{j_1,j_2\in\{i_1,i_2\}}\langle\phi_{i_1},\phi_{i_2}|\hat{W}_{12}|\phi_{j_1},\phi_{j_2}\rangle\underbrace{\epsilon_{i_1,i_2}\epsilon_{j_1,j_2}}_{\delta_{i_1,j_1}\delta_{i_2,j_2}-\delta_{i_1,j_2}\delta_{i_1,j_2}} \\
&= \frac{1}{2}\sum_{i_1,i_2}^{N}\Big(\langle\phi_{i_1},\phi_{i_2}|\hat{W}_{12}|\phi_{i_1},\phi_{i_2}\rangle - \langle\phi_{i_1},\phi_{i_2}|\hat{W}_{12}|\phi_{i_2},\phi_{i_1}\rangle\Big)
\end{aligned}
$$

Note, that we exploited in the last step that the two interaction terms cancel when the indices i_1 and i_2 are identical. Therefore they need not be excluded from the sum. The latter formula is more convenient to discuss the physical implications, while the former is commonly used in the quantum chemical literature.

Thus, we obtain

$$
\langle\Psi|\frac{1}{2}\sum_{i\neq j}^{N}\hat{W}_{i,j}|\Psi\rangle = \frac{1}{2}\sum_{i,j=1}^{N}\Big[\langle\phi_i,\phi_j|\hat{W}|\phi_i,\phi_j\rangle - \langle\phi_i,\phi_j|\hat{W}|\phi_j,\phi_i\rangle\Big] \qquad (6.12)
$$

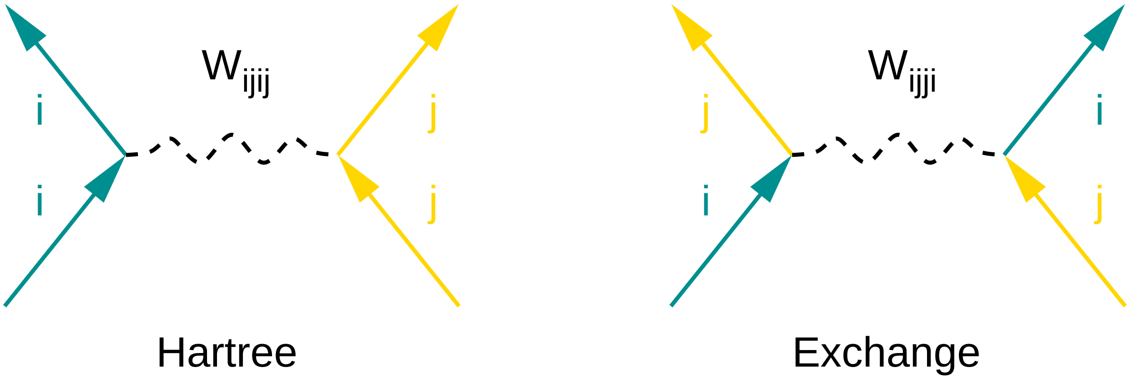


Fig. 6.1: The left diagram describes that two particles are scattered by the Coulomb interaction. The right diagram describes the same process, but the two electrons are exchanged. The second process is possible, because the two electrons are indistinguishable so that we cannot detect if the two electrons are still the same or not.

To be concise, we write here the detailed expressions for the matrix elements used in the above equation Eq. 6.12

$$\langle\phi_a,\phi_b|\hat{W}|\phi_c,\phi_d\rangle=\sum_{\sigma,\sigma'}\int d^3r\int d^3r'\;\phi^*_a(\vec{r},\sigma)\phi^*_b(\vec{r'},\sigma')\frac{e^2}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\phi_c(\vec{r},\sigma)\phi_d(\vec{r'},\sigma') \tag{6.13}$$

Homework: Work out the expectation value of the Coulomb interaction between two electrons for a two-electron Slater determinant. Consider once a Slater determinant from non-orthonormal one-particle orbitals and once for orthonormal one-particle orbitals. Investigate the role of the spin indices.

6.3 Hartree energy

The surprising fact of Eq. 6.12 is the appearance of two terms. Therefore let us try to give some physical meaning to the two contributions.

The first interaction term in Eq. 6.12 is the so-called **Hartree energy**. The Hartree energy turns out to be the classical electrostatic interaction of the electron density.

$$\begin{aligned}E_H&=\frac{1}{2}\sum_{i,j}\sum_{\sigma,\sigma'}\int d^3r\int d^3r'\;\phi^*_i(\vec{r},\sigma)\phi^*_j(\vec{r'},\sigma')\frac{e^2}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\phi_i(\vec{r},\sigma)\phi_j(\vec{r'},\sigma')\\&=\frac{1}{2}\int d^3r\int d^3r'\;\underbrace{\left[\sum_{i,\sigma}\phi^*_i(\vec{r},\sigma)\phi_i(\vec{r},\sigma)\right]}_{n(\vec{r})}\frac{e^2}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\underbrace{\left[\sum_{j,\sigma'}\phi^*_j(\vec{r'},\sigma')\phi_j(\vec{r'},\sigma')\right]}_{n(\vec{r'})}\end{aligned} \tag{6.14}$$

The equation can be simplified using the definition of the charge density $\rho(\vec{r})$, or the electron density $n(\vec{r})$, as

$$\rho(\vec{r})=-en(\vec{r})=-e\underbrace{\sum_{i=1}^{N}\sum_{\sigma}\phi^*_i(\vec{r},\sigma)\phi_i(\vec{r},\sigma)}_{n(\vec{r})}$$

Thus, we obtain the final expression for the Hartree energy

HARTREE ENERGY

$$E_H=\frac{1}{2}\int d^3r\int d^3r'\;\frac{\rho(\vec{r})\rho(\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|} \tag{6.15}$$

6.4 Exchange energy

Exchange energy

The Hartree energy is clearly not the correct electrostatic energy, as it describes the interaction of N electrons with N electrons. However, each electron can only interact with $N-1$ electrons. Thus, the Hartree term also includes, incorrectly, the interaction of each electron with itself.

This so-called **self interaction** is subtracted by the so-called **exchange energy**, the second interaction term in Eq. 6.12

$$E_X = -\frac{1}{2}\sum_{i,j}\sum_{\sigma,\sigma'}\int d^3r \int d^3r'\ \phi_i^*(\vec{r},\sigma)\phi_j^*(\vec{r'},\sigma')\frac{e^2}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\phi_j(\vec{r},\sigma)\phi_i(\vec{r'},\sigma') \tag{6.16}$$

The exchange term consists of the electrostatic interaction of densities[3]

$$n_{i,j}(\vec{r})\stackrel{\text{def}}{=}\sum_\sigma \phi_i(\vec{r},\sigma)\phi_j^*(\vec{r},\sigma) \tag{6.17}$$

The exchange energy can be written as

$$E_X = -\frac{1}{2}\sum_{i,j}\int d^3r \int d^3r'\ \frac{e^2 n_{j,i}(\vec{r})n_{i,j}(\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|} \tag{6.18}$$

while the Hartree term has the form

$$E_H = \frac{1}{2}\sum_{i,j}\int d^3r \int d^3r'\ \frac{e^2 n_{i,i}(\vec{r})n_{j,j}(\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\ . \tag{6.19}$$

Self interaction

All pair densities with different indices integrate to zero and the one with identical indices integrate to one, that is

$$\int d^3r\ n_{i,j}(\vec{r}) \stackrel{\text{Eq. 6.17}}{=} \langle\phi_j|\phi_i\rangle = \delta_{i,j}$$

Thus, we can expect that the dominant contribution to the exchange energy results from the diagonal terms. This so-called **self-interaction energy**

$$\begin{aligned}E_{SI} &= -\frac{1}{2}\sum_{i}\sum_{\sigma,\sigma'}\int d^3r \int d^3r'\ \phi_i^*(\vec{r},\sigma)\phi_i^*(\vec{r'},\sigma')\frac{e^2}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\phi_i(\vec{r},\sigma)\phi_i(\vec{r'},\sigma')\\ &= -\frac{1}{2}\sum_{i}\int d^3r \int d^3r'\ \frac{e^2 n_{ii}(\vec{r})n_{ii}(\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\end{aligned}$$

describes the electrostatic interaction of each electron with itself, that has been included in the Hartree correction. There is a well-known electronic structure method, based on density-functional theory, which uses the same trick to get rid of artificial self-interaction terms.[37] This theory subtracts the self energy from the Hartree term and thus includes the so-called self-interaction correction $E_{SIC} = -E_{SI}$.

Exchange and spin alignment

Only electron pairs with the same spin contribute to the exchange energy. This can be shown when one considers one-particle spin-orbitals with defined spin along the z-axis. That is, each one-particle spin-orbital has only one spin component, the other being zero.

---

[3] In the expression for $n_{i,j}$ the complex conjugate of the right orbital is taken on purpose. This choice is analogous to that of the one-particle density matrix.

- If the two orbitals have the same spin, pointing for example into the $\vec{e}_z$ direction, they can be written as

$$\phi_i(\vec{x}) = \begin{pmatrix} \phi_i(\vec{r},\uparrow) \\ 0 \end{pmatrix} \qquad \text{and} \qquad \phi_j(\vec{x}) = \begin{pmatrix} \phi_j(\vec{r},\uparrow) \\ 0 \end{pmatrix}$$

  The resulting exchange-energy would be formed from the pair densities

$$n_{i,j}(\vec{r}) \overset{\text{Eq. 6.17}}{=} \sum_\sigma \phi_i(\vec{r},\sigma)\phi_j^*(\vec{r},\sigma) = \phi_i(\vec{r},\uparrow)\phi_j^*(\vec{r},\uparrow) \qquad (6.20)$$

  so that its contribution to the exchange energy has the form

$$\begin{aligned} \Delta E_X &\overset{\text{Eq. 6.18}}{=} -\frac{1}{2}\sum_{i,j}\int d^3r \int d^3r' \, \frac{e^2 n_{j,i}(\vec{r}) n_{i,j}(\vec{r}\,')}{4\pi\epsilon_0|\vec{r}-\vec{r}\,'|} \\ &\overset{\text{Eq. 6.20}}{=} -\frac{1}{2}\int d^3r \int d^3r' \, \frac{e^2 \phi_i^*(\vec{r},\uparrow)\phi_j(\vec{r},\uparrow)\phi_j^*(\vec{r}\,',\uparrow)\phi_i(\vec{r}\,',\uparrow)}{4\pi\epsilon_0|\vec{r}-\vec{r}\,'|} \end{aligned}$$

- If the two orbitals have the opposite spin, they can be written as

$$\phi_i(\vec{x}) = \begin{pmatrix} \phi_{i,\uparrow}(\vec{r}) \\ 0 \end{pmatrix} \qquad \text{and} \qquad \phi_j(\vec{x}) = \begin{pmatrix} 0 \\ \phi_{j,\downarrow}(\vec{r}) \end{pmatrix}$$

  The contribution to the exchange energy vanishes, because the pair densities vanish

$$n_{i,j}(\vec{r}) \overset{\text{Eq. 6.17}}{=} \sum_\sigma \phi_i(\vec{r},\sigma)\phi_j(\vec{r},\sigma) = \phi_i(\vec{r},\uparrow)\cdot 0 + 0\cdot\phi_j(\vec{r},\downarrow) = 0$$

Note that this behavior of the exchange term is different from the one of the Hartree term: There is a Coulomb interaction between electrons of opposite spin, but no exchange term, because spin-up and spin-down electrons are distinguishable.

HUND'S RULE AND FERROMAGNETISM

Exchange acts only between electrons pairs with the same spin. Because exchange counteracts the Coulomb repulsion due to its changed sign, it favors if many electrons have the same spin.
A result is **Hund's 1st rule**: *For a given electron configuration, the term with maximum multiplicity has the lowest energy. The multiplicity is equal to , where is the total spin angular momentum for all electrons. The term with lowest energy is also the term with maximum .*
The same effect is responsible for the tendency of some solids to become **ferromagnetic**. Due to the finite band width of metals, spin-alignment leads to some loss of kinetic energy. The balance of kinetic energy and exchange determines whether ferromagnetism wins.

6.5 Hartree-Fock equations

The Hartree-Fock energy is an upper bound

The Hartree-Fock method yields an upper estimate E_{HF} for the ground-state energy E_{GS}.

The ground state of an N-electron system is that state that minimizes the total energy for a given Hamiltonian, i.e.

$$E_{GS} = \min_{|\Phi\rangle,E}\Big[\langle\Phi|\hat{H}|\Phi\rangle - E\Big(\langle\Phi|\Phi\rangle - 1\Big)\Big] \qquad (6.21)$$

for all normalized many-particle wave functions $|\Phi\rangle$. The normalization constraint has been added with the method of Lagrange multipliers and E is the corresponding Lagrange multiplier.[4]

The Hartree-Fock method restricts this search for the minimum to a subset $\mathbb{S}$ of wave functions, namely those that can be expressed in the form of a single Slater determinant.

$$E_{HF} = \min_{|\Phi\rangle\in\mathbb{S},E'} \Big[\langle\Phi|\hat{H}|\Phi\rangle - E'\Big(\langle\Phi|\Phi\rangle - 1\Big)\Big] \tag{6.22}$$

Because the search for the minimum is limited in the Hartree-Fock, the Hartree-Fock energy is always above the ground-state energy.

$$E_{GS} \le E_{HF} \tag{6.23}$$

Hartree-Fock equations

Because every Slater determinant $|\Phi\rangle \in \mathbb{S}$ can be expressed by a set of one-particle orbitals $|\phi_j\rangle$, the Hartree-Fock energy can be obtained by a minimum condition for the latter.

$$E_{HF} = \min_{|\phi_1\rangle,\ldots,|\phi_n\rangle,\Lambda} \Big[E_{NN} + \sum_{i=1}^{N}\langle\phi_i|\hat{h}|\phi_i\rangle + \frac{1}{2}\sum_{i,j=1}^{N}\Big[\langle\phi_i,\phi_j|\hat{W}|\phi_i,\phi_j\rangle - \langle\phi_i,\phi_j|\hat{W}|\phi_j,\phi_i\rangle\Big]$$
$$-\sum_{i,j}\Lambda_{i,j}\left[\langle\phi_i|\phi_j\rangle - \delta_{i,j}\right]\Big]$$

The last term ensures the orthonormality of the one-particle orbitals using the method of Lagrange multipliers. The translation of the matrix elements between Slater determinants into those of the one-particle orbitals has been given in section 6.2. The matrix elements of one-particle Hamiltonian and interaction are defined in Eq. 6.6 and Eq. 6.13. The energy E_{NN} is the Coulomb repulsion of the nuclei as defined in Eq. 6.1.

The minimum condition yields the **Hartree-Fock equations**. The derivation of the Hartree-Fock equations is analogous to the derivation of the Euler-Lagrange equations. Special is that here we work with fields, the one-particle wave functions, and that the variational parameters are complex. It is advantageous to use the **Wirtinger derivatives**, which are described briefly in ΦSX: Klassische Mechanik.

Wirtinger derivatives The essence of Wirtinger derivatives is that we can write the identity

$$df(\mathrm{Re}[c],\mathrm{Im}[c]) = \frac{df}{d\mathrm{Re}[c]}d\mathrm{Re}[c] + \frac{df}{d\mathrm{Im}[c]}d\mathrm{Im}[c] = \frac{df}{dc}dc + \frac{df}{dc^*}dc^* = df(c,c^*)$$

for the first variation of f. Thus, we can form the variations of a function of complex arguments, as if the variable and its complex conjugate were completely independent variables.
For the functional derivatives we use

$$\frac{d\Psi_n^*(\vec{r},\sigma)}{d\Psi_k^*(\vec{r}_0,\sigma_0)} = \delta(\vec{r}-\vec{r}_0)\delta_{\sigma,\sigma_0}\delta_{n,k}$$
$$\frac{d\Psi_n(\vec{r},\sigma)}{d\Psi_k^*(\vec{r}_0,\sigma_0)} = 0$$

---

[4] The expression is written a little sloppy, because the expression need not be be a minimum with respect to the Lagrange multiplier. The requirement is only that the gradient vanishes.

$$
\begin{aligned}
&\left[\frac{-\hbar^2}{2m_e}\vec{\nabla}^2+\sum_{j=1}^{M}\frac{-e^2Z_j}{4\pi\epsilon_0|\vec{r}-\vec{R}_j|}+\int d^3r'\;\frac{e^2n(\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\right]\phi_i(\vec{r},\sigma)\\
&-\sum_{j=1}^{N}\sum_{\sigma'}\int d^3r'\;\frac{e^2\phi_j^*(\vec{r'},\sigma')\phi_i(\vec{r'},\sigma')}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\phi_j(\vec{r},\sigma)-\sum_{j=1}^{N}\phi_j(\vec{r},\sigma)\Lambda_{j,i}=0
\end{aligned}
$$

Note that this is not a simple Schrödinger equation, because it mixes all orbitals.

However, we can rewrite the equations to obtain a true Hamiltonian, which acts on all wave functions in the same way:

$$
\begin{aligned}
&\left[\frac{-\hbar^2}{2m_e}\vec{\nabla}^2+\sum_{j=1}^{M}\frac{-e^2Z_j}{4\pi\epsilon_0|\vec{r}-\vec{R}_j|}+\int d^3r'\;\frac{e^2n(\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\right]\phi_i(\vec{r},\sigma)\\
&-\sum_{\sigma'}\int d^3r'\;\underbrace{\left[\sum_{j=1}^{N}\frac{e^2\phi_j(\vec{r},\sigma)\phi_j^*(\vec{r'},\sigma')}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\right]}_{-V_{X,\sigma,\sigma'}(\vec{r},\vec{r'})}\phi_i(\vec{r'},\sigma')-\sum_{j}\phi_j(\vec{r},\sigma)\Lambda_{j,i}=0
\end{aligned}
\tag{6.24}
$$

This form actually shows that the one-particle orbitals can be superimposed so that the matrix Λ becomes diagonal. It can be shown that the system of equations is invariant under a unitary transformation of the one-particle orbitals. Thus, one can use the matrix that diagonalizes Λ, to bring the equations into a diagonal form. Hence, we can write the Hartree-Fock equations as true eigenvalue equations.

The following observation provides some physical insight: The nonlocal exchange potential

$$
V_{X,\sigma,\sigma'}(\vec{r},\vec{r'})=\left[-\sum_{j=1}^{N}\frac{e^2\phi_j(\vec{r},\sigma)\phi_j^*(\vec{r'},\sigma')}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\right]
\tag{6.25}
$$

contains a spin and angular momentum (ℓ,m) projection. Thus, every occupied atomic orbital is shifted towards lower energies, while empty atomic orbitals do not experience this exchange potential. The more compact the orbital, the larger the denominator and the stronger is the stabilization.

Non-local potential

However, the price to pay for the eigenvalue equation Eq. 6.24 is that we have to deal with a truly non-local potential in the exchange term. What is a **non-local potential**? Consider an arbitrary operator $\hat{A}$

$$
\begin{aligned}
\langle\psi|\hat{A}|\phi\rangle&=\langle\psi|\underbrace{\left[\int d^3r\;|\vec{r}\rangle\langle\vec{r}|\right]}_{\hat{1}}\hat{A}\underbrace{\left[\int d^3r'\;|\vec{r'}\rangle\langle\vec{r'}|\right]}_{\hat{1}}|\phi\rangle\\
&=\int d^3r\int d^3r'\;\underbrace{\langle\psi|\vec{r}\rangle}_{\psi^*(\vec{r})}\underbrace{\langle\vec{r}|\hat{A}|\vec{r'}\rangle}_{A(\vec{r},\vec{r'})}\underbrace{\langle\vec{r'}|\phi\rangle}_{\phi(\vec{r'})}\\
&=\int d^3r\int d^3r'\;\psi^*(\vec{r})A(\vec{r},\vec{r'})\phi(\vec{r'})
\end{aligned}
$$

Thus, a general one-particle operator has always two sets of coordinates in its real space representation. We call them non-local because they act on two positions at the same time. It is a special property of most potentials, that they are local. For a **local operator** or **local potential** $\hat{V}$ the real space matrix elements are diagonal in real space, that is

$$V(\vec{r},\vec{r'})\overset{\text{def}}{=}\langle\vec{r}|\hat{V}|\vec{r'}\rangle = V(\vec{r})\delta(\vec{r}-\vec{r'})$$

so that its matrix elements can be represented by a single real space integral.

Self-consistency cycle

The Hartree-Fock equations are **self-consistent equations**. This means that the Hamiltonian depends itself on the wave functions. Thus, the problem has to be solved iteratively.

- We start out with trial wave functions $|\phi_i\rangle$ to determine the Hamiltonian.
- Then, we obtain new wave functions from the eigenvalue problem Eq. 6.24.
- New and old wave functions are mixed, which is often necessary to obtain convergence of the cycle.
- The mixed wave functions are used as new trial functions to repeat the cycle.

If the new and old wave functions or the new and old Hamiltonian are identical, the wave functions are consistent with the Hamiltonian and the Hartree-Fock equations are solved.

In the quantum chemical literature the term self-consistent equations is used synonymous with Hartree-Fock equations. This is misleading because there are many other schemes that also require a self-consistent cycle.

Fock operator

We can now write the Hamilton operator of the Hartree-Fock equation, namely the so-called **Fock operator**

FOCK OPERATOR

$$\begin{aligned}
\hat{H}_{Fock} = &\sum_{\sigma}\int d^3r\;|\vec{r},\sigma\rangle\left(\frac{-\hbar^2}{2m_e}\vec{\nabla}^2-\sum_{j=1}^{M}\frac{Z_je^2}{4\pi\epsilon_0|\vec{r}-\vec{R}_j|}+\int d^3r'\;\frac{n(\vec{r'})e^2}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\right)\langle\vec{r},\sigma|\\
&-\sum_{i,j=1}^{n}\sum_{\sigma,\sigma'}\int d^3r\int d^3r'\;|\vec{r},\sigma\rangle\frac{e^2\phi^*_{j,\sigma}(\vec{r})\phi_{j,\sigma'}(\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\langle\vec{r'},\sigma'|\\
&+\left[\hat{1}-\sum_{j=1}^{n}|\phi_j\rangle\langle\phi_j|\right]\hat{C}\left[\hat{1}-\sum_{j=1}^{n}|\phi_j\rangle\langle\phi_j|\right]
\end{aligned}$$

The operator $\hat{C}$ in the last term is completely arbitrary. Because of the projection operator on both sides, it only acts on the "unoccupied" one-particle wave functions. Because this Hamilton operator is never applied to the unoccupied states, it is also irrelevant. We included this term here to show that the Hartree-Fock operator contains a lot of arbitrariness and that furthermore the action of the operator on the unoccupied states is arbitrary.

Note, that the Hartree-Fock total energy is not identical to the sum of one-particle energies as it is the case for non-interacting particles.

6.6 Hartree-Fock of the free-electron gas

Let us now study the free-electron gas in the Hartree-Fock approximation. Our goal is to determine the changes of the dispersion relation due to the interaction.

First we determine non-local potential as defined in Eq. 6.24.

$$V_x(\vec{x},\vec{x}') \overset{Eq.\ 6.24}{=} -\left[\sum_j \frac{e^2\phi_j^*(\vec{x})\phi_j(\vec{x}')}{4\pi\epsilon_0|\vec{r}-\vec{r}'|}\right] \qquad (6.26)$$

$$\overset{Eq.\ N.3}{=} \frac{e^2\delta_{\sigma,\sigma'}}{4\pi\epsilon_0 s}\frac{1}{(2\pi)^3}\frac{4\pi}{3}k_F^3\left[3\frac{(k_Fs)\cos(k_Fs)-\sin(k_Fs)}{(k_Fs)^3}\right] \qquad (6.27)$$

where $s=|\vec{r}-\vec{r}'|$ and k_F is the Fermi momentum.

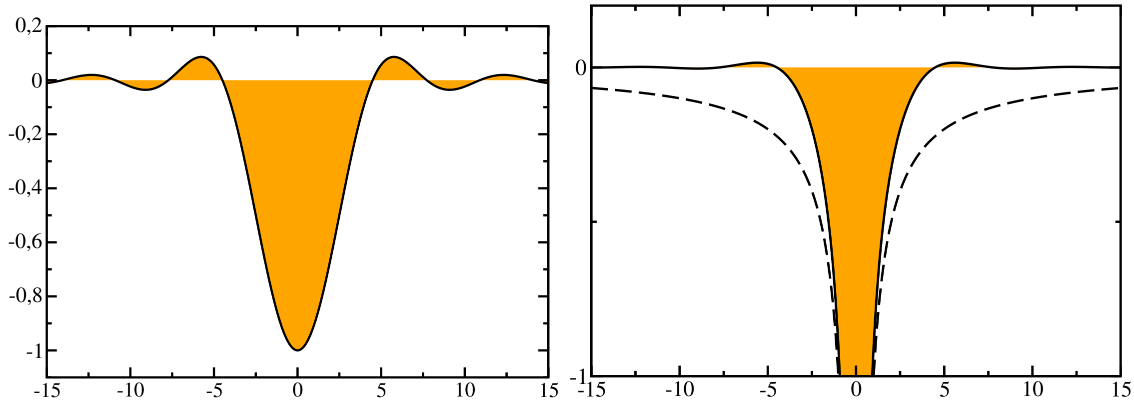


Fig. 6.2: The shape of the non-local exchange potential (right) for a free electron gas as calculated in the Hartree-Fock method. The dashed line corresponds to a Coulomb interaction. The function $3\frac{x\cos(x)-\sin(x)}{x^3}$ is shown on the left-hand side.

Now we need to evaluate the expectation values of this potential in order to obtain the energy shifts:

$$d\epsilon_{\vec{k},\sigma} = \frac{1}{\Omega}\int d^3r\int d^3r'\ V_x(|\vec{r}-\vec{r}'|,\sigma,\sigma')e^{i\vec{k}(\vec{r}-\vec{r}')}$$
$$= -\frac{e^2}{4\pi\epsilon_0}\frac{2k_F}{\pi}\left[\frac{1}{2}+\frac{1-a^2}{4a}\ln\left|\frac{1+a}{1-a}\right|\right]$$

Here, $a=|\vec{k}|/k_F$. The function in parenthesis is shown in Fig. 6.3

We find, as shown in Fig. 6.4, that the occupied band width is larger in Hartree-Fock than for the non-interacting free electrons. However, the experimental band width of real free-electron-like materials is smaller. For potassium the band width of the free electron gas is 2 eV. The Hartree-Fock result is 5.3 eV, while the band width of real potassium is in the range 1.5-1.6 eV.[1]

A second problem[1] of the Hartree-Fock theory is that the density of states at the Fermi level vanishes. This is again in contradiction with experiment, which is described pretty well with $\sqrt{\epsilon}$-behavior of the density of states. The experimental result is closer to the free electron gas. The discrepancy is related to the fact that the electron can not respond to the potential describing the interaction. Note that the wave functions are still plane waves.[5]

---

[5] This argument as such is not conclusive. The band structure is obtained for a fixed potential. The potential depends explicitly on the number of electrons. In an experiment the Fermi level is measured for example by extracting an electron, which changes the electron number and therefore the interaction potential of Hartree-Fock.

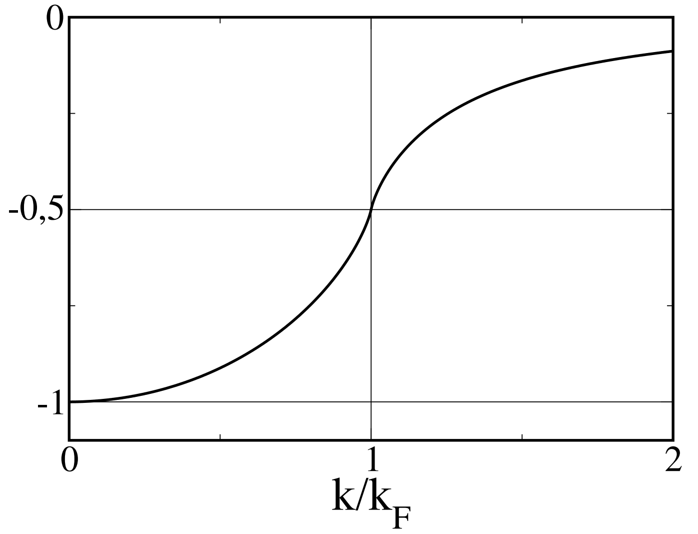


Fig. 6.3: The function $f(a) = -\left(\frac{1}{2} + \frac{1-a^2}{4a}\ln\left|\frac{1+a}{1-a}\right|\right)$ as function of $a = k/k_F$. Note that the slope for $k = k_F$ is infinite. For $k >> k_F$ the function approaches zero.

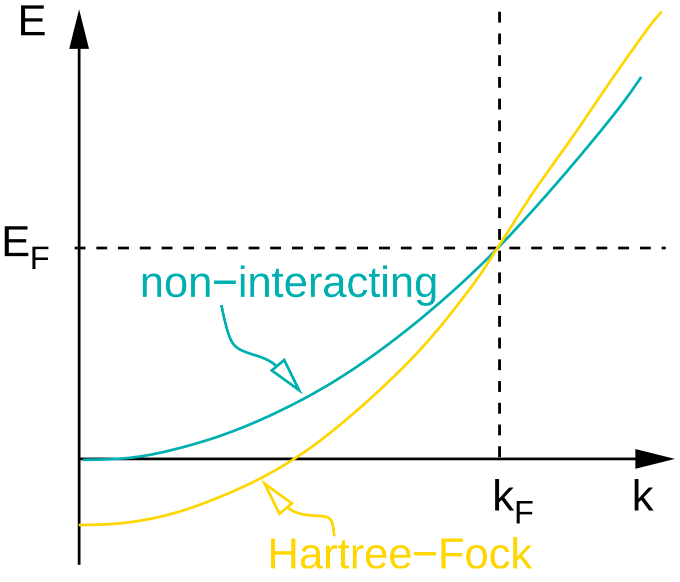


Fig. 6.4: Dispersion relation of the free electron gas as calculated with Hartree-Fock and without interactions. The two dispersion relations are shifted so that their Fermi levels agree.

6.7 Beyond the Hartree-Fock Theory: Correlations

The result obtained by the Hartree-Fock approximation always lies above the correct result. The energy difference is called electron correlation. The term was coined by Per-Olov Löwdin. [P.-O. Löwdin, "Correlation Problem in Many-Electron quantum mechanics", Adv. Chem. Phys. 2, 207 (1959)]

In this article I will point out a number of failures of Hartree-Fock theory, in order to introduce into the terminology of correlations.

6.7.1 Left-right and in-out correlation

The notation of left-right correlation and in-out correlation has been termed by Kolos and Roothaan[38], who performed accurate calculations for the hydrogen molecule.

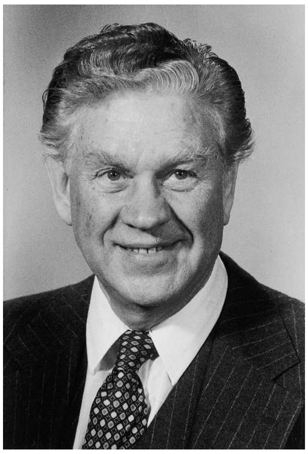


Fig. 6.5: Per-Olov Löwdin, Swedish physicist and theoretical chemist 1916-2000. Courtesy of the American Philosophical Society.

Left-right correlation describes that the two electrons in one bond are most likely on the opposite sides of the bond. This principle is violated by a Slater determinant build from bonding orbitals.

Let us consider a hydrogen molecule. We start from a one-particle basis $|L,\uparrow\rangle, |L,\downarrow\rangle, |R,\uparrow\rangle, |R,\downarrow\rangle$, where L refers to the left atom and R refers to the right atom. First we transform the basis onto bonding and antibonding orbitals.

$$
\begin{aligned}
|B,\uparrow\rangle &\stackrel{\text{def}}{=} \frac{1}{\sqrt{2}}\Big(|L,\uparrow\rangle + |R,\uparrow\rangle\Big) && (6.28)\\
|B,\downarrow\rangle &\stackrel{\text{def}}{=} \frac{1}{\sqrt{2}}\Big(|L,\downarrow\rangle + |R,\downarrow\rangle\Big) && (6.29)\\
|A,\uparrow\rangle &\stackrel{\text{def}}{=} \frac{1}{\sqrt{2}}\Big(|L,\uparrow\rangle - |R,\uparrow\rangle\Big) && (6.30)\\
|A,\downarrow\rangle &\stackrel{\text{def}}{=} \frac{1}{\sqrt{2}}\Big(|L,\downarrow\rangle - |R,\downarrow\rangle\Big) && (6.31)
\end{aligned}
$$

Now we form a Slater determinant from the two bonding orbitals

$$
\begin{aligned}
\Psi_g(\vec{r}_1,\sigma_1,\vec{r}_2,\sigma 2) &= \frac{1}{\sqrt{2}}\Big(\langle\vec{r}_1,\sigma_1|B,\uparrow\rangle\langle\vec{r}_2,\sigma_2|B,\downarrow\rangle - \langle\vec{r}_1,\sigma_1|B,\downarrow\rangle\langle\vec{r}_2,\sigma_2|B,\uparrow\rangle\Big)\\
&= \frac{1}{2}\Big\{\langle\vec{r}_1,\sigma_1|\Big(|L,\uparrow\rangle + |R,\uparrow\rangle\Big)\langle\vec{r}_2,\sigma_2|\Big(|L,\downarrow\rangle + |R,\downarrow\rangle\Big)\\
&\quad -\langle\vec{r}_1,\sigma_1|\Big(|L,\downarrow\rangle + |R,\downarrow\rangle\Big)\langle\vec{r}_2,\sigma_2|\Big(|L,\uparrow\rangle + |R,\uparrow\rangle\Big)\\
&= \frac{1}{2}\Big\{\Big(\langle\vec{r}_1,\sigma_1||L,\uparrow\rangle\langle\vec{r}_1,\sigma_1||R,\uparrow\rangle\Big)\Big(\langle\vec{r}_2,\sigma_2|L,\downarrow\rangle + \langle\vec{r}_2,\sigma_2|R,\downarrow\rangle\Big)\\
&- \Big(\langle\vec{r}_1,\sigma_1|L,\downarrow\rangle\langle\vec{r}_1,\sigma_1|R,\downarrow\rangle\Big)\Big(\langle\vec{r}_2,\sigma_2|L,\uparrow\rangle + \langle\vec{r}_2,\sigma_2|R,\uparrow\rangle\Big)\\
&= \frac{1}{2}\Big\{\Big(\langle\vec{r}_1,\sigma_1|L,\uparrow\rangle\langle\vec{r}_2,\sigma_2|L,\downarrow\rangle - \langle\vec{r}_1,\sigma_1|L,\downarrow\rangle\langle\vec{r}_2,\sigma_2|L,\uparrow\rangle\Big)\\
&\quad +\Big(\langle\vec{r}_1,\sigma_1|L,\uparrow\rangle\langle\vec{r}_2,\sigma_2|R,\downarrow\rangle - \langle\vec{r}_1,\sigma_1|R,\downarrow\rangle\langle\vec{r}_2,\sigma_2|L,\uparrow\rangle\Big)\\
&\quad +\Big(\langle\vec{r}_1,\sigma_1|R,\uparrow\rangle\langle\vec{r}_2,\sigma_2|L,\downarrow\rangle - \langle\vec{r}_1,\sigma_1|L,\downarrow\rangle\langle\vec{r}_2,\sigma_2|R,\uparrow\rangle\Big)\\
&\quad +\Big(\langle\vec{r}_1,\sigma_1|R,\uparrow\rangle\langle\vec{r}_2,\sigma_2|R,\downarrow\rangle - \langle\vec{r}_1,\sigma_1|R,\downarrow\rangle\langle\vec{r}_2,\sigma_2|R,\uparrow\rangle\Big)\\
&= \frac{1}{\sqrt{2}}\left[(\uparrow\downarrow,0) + (\downarrow,\uparrow) + (\uparrow,\downarrow) + (0,\uparrow\downarrow)\right]
\end{aligned}
$$

Thus, the Slater determinant of the bonding orbitals contains determinants where both electrons are either both localized to the right or with both electrons localized on the left side. These ionized configuration lie high in energy of the bond is long. This is the explanation that the Hartree-Fock theory in this form results in a too large dissociation energy.

In order to obtain the correct limit, namely $\frac{1}{\sqrt{2}}(\uparrow,\downarrow) + (\downarrow,\uparrow)$ we need to mix in the Slater determinant with the antibonding orbitals.

$$
\Psi_u(\vec{r}_1,\sigma_1,\vec{r}_2,\sigma 2) = \frac{1}{\sqrt{2}}\left[(\uparrow\downarrow,0) - (\downarrow,\uparrow) - (\uparrow,\downarrow) + (0,\uparrow\downarrow)\right]
$$

6.7.2 In-out correlation

- “in-out” correlation
- left-right correlation
- angular correlation
- radial correlation
- dynamic
- nondynamic
- near-degeneracy
- long-range
- shortrange

6.7.3 Spin contamination

6.7.4 Dynamic and static correlation

In quantum chemistry the definition of dynamic and static correlation is opposite to that used by physicists. In Quantum chemistry, static correlation is the effect of very few Slater determinants that contribute to the ground state with approximately equal weight. That static correlation is present if the dominant Slater determinant has a weight less than about 95% of the complete wave function.

Dynamic correlation on the other hand are the small contribution of many Slater determinants, lowering the energy. Dynamic correlation is important to describe the cusp condition: The Coulomb hole has a cusp in the density at the site of the observer electron. Because all one-particle orbitals are smooth, many Slater determinants are required to build up this cusp as superposition of smooth wave functions.

Chapter 7

Density-functional theory

Density-functional theory[39, 40] is an extremely powerful technique. It is based on an exact theorem that the ground-state energy is determined solely by the density alone.

The theorem says

- that all ground-state properties are unique functionals of the electron density and
- that the electron density can be obtained from a Schrödinger equation in an effective potential.

density-functional theory maps the interacting electron system onto a system of non-interacting electrons in an effective potential. The effective potential depends in turn on the electron distribution and describes the interacting between the electrons in an effective way.

density-functional theory provides a total energy functional of the form

$$E[n(\vec{r})] = \underbrace{\sum_n f_n \langle \phi_n | \frac{\hat{p}^2}{2m_e} | \phi_n \rangle}_{T_s[n(\vec{r})]} + \int d^3r \; n(\vec{r}) v_{ext}(\vec{r})$$

$$+ \underbrace{\frac{1}{2} \int d^3r \int d^3r' \; \frac{e^2 n(\vec{r}) n(\vec{r'})}{4\pi\epsilon_0 |\vec{r} - \vec{r'}|}}_{E_{Hartree}} + E_{xc}[n(\vec{r})]$$

T_s is not the kinetic energy of the many-electron system, but the one of the non-interacting reference system with density $n(\vec{r})$. We will see later that it can be written as functional of the density. E_{xc} is another functional of the density, called the exchange and correlation energy functional. This functional is not known in practice. Therefore, one uses approximate functionals, that are derived from known analytical properties of the functional and/or even experimental data.

As a result of the minimum principle we obtain the following Kohn-Sham equation for the one-particle orbitals

$$\left[\frac{\hat{p}^2}{2m_e} + v_{eff}(\hat{r}) - \epsilon_n \right] |\phi_n\rangle = 0$$

where the effective potential is defined as

$$v_{eff}(\vec{r}) = v_{ext}(\vec{r}) + \frac{e^2}{4\pi\epsilon_0} \int d^3r' \; \frac{n(\vec{r'})}{|\vec{r} - \vec{r'}|} + \underbrace{\frac{\delta E_{xc}}{\delta n(\vec{r})}}_{\mu_{xc}(\vec{r})}$$

The proofs will be given later. Let us start out with the language used in the field of density-functional theory and the physical pictures that emerge.

7.1 One-particle and two-particle densities

In order to get used with the terminology, let me introduce a number of definitions related to density matrices[41].

7.1.1 N-particle density matrix

Let us start with the **N-particle density matrix** of an N-particle system

$$\rho^{(N)}(\vec{x}_1,\ldots,\vec{x}_N;\vec{x'}_1,\ldots,\vec{x'}_N)\overset{\text{def}}{=}\Psi(\vec{x}_1,\ldots,\vec{x}_N)\Psi^*(\vec{x'}_1,\ldots,\vec{x'}_N)$$

The N-particle density matrix is formed from the matrix elements

$$\rho^{(N)}(\vec{x}_1,\ldots,\vec{x}_N;\vec{x'}_1,\ldots,\vec{x'}_N)=\langle\vec{x}_1,\ldots,\vec{x}_N|\hat{\rho}^{(N)}|\vec{x'}_1,\ldots,\vec{x'}_N\rangle$$

of the **density operator**

$$\hat{\rho}^{(N)}\overset{\text{def}}{=}|\Psi\rangle\langle\Psi| \tag{7.1}$$

This is the same density operator, that has been used in statistical mechanics a lot. In Eq. 7.1, however, we restricted ourselves to the density matrix of a pure state. A pure state can be described by a single many-particle wave function, whereas a statistical mixture is an ensemble of many-particle states, of which each contributes with a certain probability.

The N-particle density matrix contains all the information of the wave function itself, because one can determine the expectation value of any operator from it:

We use the representation of the unity operator in a N-particle Hilbert space of spin-orbitals

$$\hat{1}=\int d^4x_1\cdots\int d^4x_N\;|\vec{x}_1,\ldots,\vec{x}_N\rangle\langle\vec{x}_1,\ldots,\vec{x}_N|$$

and obtain

$$\begin{aligned}
\langle\Psi|\hat{A}|\Psi\rangle &= \langle\Psi|\hat{A}|\Psi\rangle=\langle\Psi|\hat{1}\hat{A}\hat{1}|\Psi\rangle\\
&=\int d^4x_1\cdots\int d^4x_N\int d^4x'_1\cdots\int d^4x'_N\\
&\quad\cdot\underbrace{\langle\vec{x}_1\ldots,\vec{x}_N|\Psi\rangle\langle\Psi|\vec{x'}_1\ldots,\vec{x'}_N\rangle}_{\rho^{(N)}}\langle\vec{x'}_1\ldots,\vec{x'}_N|\hat{A}|\vec{x}_1\ldots,\vec{x}_N\rangle\\
&=\int d^4x_1\cdots\int d^4x_N\int d^4x'_1\cdots\int d^4x'_N\\
&\quad\cdot\rho^{(N)}(\vec{x}_1,\ldots,\vec{x}_N;\vec{x'}_1,\ldots,\vec{x'}_N)\langle\vec{x'}_1\ldots,\vec{x'}_N|\hat{A}|\vec{x}_1\ldots,\vec{x}_N\rangle
\end{aligned} \tag{7.2}$$

For us the expectation value of the Hamilton operator plays a special role. The energy expectation value can be constructed from the N-particle density matrix. However, the special form of the Hamilton operator relevant to us allows to use simpler, so-called reduced density matrices. We will define those in the following.

Fertig: WS06/07 8 Doppelstunde 27.Nov.06

7.1.2 One-particle reduced density matrix

From the N-particle density matrix we can form several **contractions** that are physically important, because they allow to represent expectation values of one-particle operators and the Coulomb interaction in an elegant fashion.

For example, we do not need the full N-particle density matrix to determine the expectation value of a one-particle operator, if we know the so-called one-particle density matrix.

A general one-particle operator has the matrix elements

$$\langle \vec{x}_1,\dots,\vec{x}_N|\hat{A}|\vec{x}'_1,\dots,\vec{x}'_N\rangle = \sum_{i=1}^{N} A(\vec{x}_i,\vec{x}'_i) \prod_{k(k\neq i)}^{N} \delta(\vec{x}_k-\vec{x}'_k) \tag{7.3}$$

This allows us to simplify the expression for the expectation value as follows

$$\begin{aligned}
\langle\Psi|\hat{A}|\Psi\rangle \overset{Eq.\ 7.2,7.3}{=} & \int d^4x_1 \cdots \int d^4x_N \int d^4x'_1 \cdots \int d^4x'_N \\
& \rho^{(N)}(\vec{x}_1,\dots,\vec{x}_N;\vec{x}'_1,\dots,\vec{x}'_N) \sum_{i=1}^{N} A(\vec{x}_i,\vec{x}'_i) \prod_{k(k\neq i)}^{N} \delta(\vec{x}_k-\vec{x}'_k) \\
\overset{(1)}{=} & \int d^4x_1 \cdots \int d^4x_N \int d^4x'_1 \cdots \int d^4x'_N \\
& \rho^{(N)}(\vec{x}_1,\dots,\vec{x}_N;\vec{x}'_1,\dots,\vec{x}'_N) N A(\vec{x}_1,\vec{x}'_1) \prod_{k=2}^{N} \delta(\vec{x}_k-\vec{x}'_k) \\
= & \int d^4x_1 \int d^4x'_1\, A(\vec{x}_1,\vec{x}'_1) \\
& \cdot N \int d^4x_2 \cdots \int d^4x_N \int d^4x'_2 \cdots \int d^4x'_N \\
& \cdot \rho^{(N)}(\vec{x}_1,\dots,\vec{x}_N;\vec{x}'_1,\dots,\vec{x}'_N) \prod_{k=2}^{N} \delta(\vec{x}_k-\vec{x}'_k) \\
= & \int d^4x_1 \int d^4x'_1\, A(\vec{x}_1,\vec{x}'_1) \\
& \cdot \underbrace{N \int d^4x_2 \cdots \int d^4x_N\, \rho^{(N)}(\vec{x}_1,\vec{x}_2,\dots,\vec{x}_N;\vec{x}'_1,\vec{x}_2\dots,\vec{x}_N)}_{\rho^{(1)}(\vec{x}_1,\vec{x}'_1)}
\end{aligned} \tag{7.4}$$

In (1), we exploited that the density matrix describes identical particles. Thus, it is invariant under permutation of two particle indices. (Simultaneously on both sides.)

Thus, if we define the **one-particle reduced density matrix**[1] $\rho^{(1)}(\vec{x},\vec{x}')$, as[2]

$$\begin{aligned}
\rho^{(1)}(\vec{x},\vec{x}') &\overset{\text{def}}{=} N \int d^4x_2 \cdots \int d^4x_N\, \rho^{(N)}(\vec{x},\vec{x}_2,\dots,\vec{x}_N;\vec{x}',\vec{x}_2,\dots,\vec{x}_N) \\
&= N \int d^4x_2 \cdots \int d^4x_N\, \Psi(\vec{x},\vec{x}_2,\dots,\vec{x}_N)\Psi^*(\vec{x}',\vec{x}_2,\dots,\vec{x}_N)
\end{aligned} \tag{7.5}$$

we can determine the expectation values of any one-particle operator as

$$\langle\Psi|\hat{A}|\Psi\rangle \overset{\text{Eq. } 7.4}{=} \int d^4x \int d^4x'\, \rho^{(1)}(\vec{x},\vec{x}')A(\vec{x}',\vec{x}) \tag{7.6}$$

The integral in the first equation of the definition Eq. 7.5 provides the density of a single electron. This density is the same for all the electrons. Therefore, we obtain the total density by multiplication by the number N of electrons, which explains the factor in front of the integral.

---

[1] The word “reduced” is usually dropped. However the context must avoid confusion with the density matrix of a one-particle system.

[2] Note that this definition only holds for density matrices of indistinguishable particles.

Electron density

The operator for the electron density at a position $\vec{r}$ is

$$\hat{n}(\vec{r}) = \sum_{\sigma} |\vec{r},\sigma\rangle\langle\vec{r},\sigma| \tag{7.7}$$

It has the matrix elements

$$\langle\vec{r},\sigma|\hat{n}(\vec{r}_0)|\vec{r'},\sigma'\rangle = \delta(\vec{r}-\vec{r}_0)\delta(\vec{r'}-\vec{r}_0)\delta_{\sigma,\sigma'} \tag{7.8}$$

in a one-particle, real-space and spin representation.

The matrix elements of the operator in a N-particle, real-space and spin representation are

$$\langle\vec{x}_1,\ldots,\vec{x}_N|\hat{n}(\vec{r}_0)|\vec{x'}_1,\ldots,\vec{x'}_N\rangle \overset{\text{Eq. 7.3}}{=} \sum_{i=1}^{N} \delta(\vec{r}_i-\vec{r}_0)\delta(\vec{r'}_i-\vec{r}_0)\delta_{\sigma_i,\sigma'_i} \prod_{k(k\neq i)} \delta(\vec{x}_k-\vec{x'}_k)$$

Thus, we obtain the electron density as

$$\begin{aligned} n(\vec{r}_0) = \langle\Psi|\hat{n}(\vec{r}_0)|\Psi\rangle &\overset{\text{Eq. 7.6}}{=} \int d^4x \int d^4x'\ \hat{\rho}^{(1)}(\vec{x},\vec{x'})\delta(\vec{r}-\vec{r}_0)\delta(\vec{r'}-\vec{r}_0)\delta_{\sigma,\sigma'} \\ &= \sum_{\sigma} \hat{\rho}^{(1)}(\vec{r}_0,\sigma,\vec{r}_0,\sigma) \end{aligned} \tag{7.9}$$

Magnetization

The magnetization operator has the form

$$\hat{\vec{m}} = \frac{q}{m_e}\hat{\vec{S}} = \frac{\hbar q}{2m_e}\hat{\vec{\sigma}} \tag{7.10}$$

Its matrix elements in a one-particle real-space representation are

$$\langle\vec{r},\sigma|\hat{m}_j|\vec{r'},\sigma'\rangle = \frac{\hbar q}{2m_e}\delta(\vec{r}-\vec{r'})\sigma_{j,\sigma,\sigma'} \tag{7.11}$$

where $\hat{\vec{S}}$ is the spin operator, q is the electron charge and m_e is its mass. With $\boldsymbol{\sigma}_j$ we denote the Pauli matrices defined in Eq. 3.2 for $j \in \{x,y,z\}$. Unfortunately, it is a little difficult not to get confused by the many σ's.

The matrix elements in the N-particle Hilbert space are

$$\langle\vec{x}_1,\ldots,\vec{x}_N|\hat{m}_n(\vec{r})|\vec{x'}_1,\ldots,\vec{x'}_N\rangle \overset{\text{Eq. 7.3}}{=} \sum_{i=1}^{N} \underbrace{\delta(\vec{r}_i-\vec{r'}_i)\frac{\hbar q}{2m_e}\sigma_{j,\sigma,\sigma'}}_{\langle\vec{r}_i,\sigma_i|\hat{m}_j|\vec{r'}_i,\sigma'_i\rangle} \prod_{k(k\neq i)} \delta(\vec{x}_k-\vec{x'}_k)$$

Thus, we obtain the magnetization as

$$\begin{aligned} \vec{m}(\vec{r}_0) = \langle\Psi|\hat{\vec{m}}(\vec{r}_0)|\Psi\rangle &= \int d^4x \int d^4x'\ \hat{\rho}^{(1)}(\vec{x},\vec{x'})\delta(\vec{r}_0-\vec{r})\vec{\sigma}_{\sigma,\sigma'} \\ &= \sum_{\sigma,\sigma'} \hat{\rho}^{(1)}(\vec{r}_0,\sigma,\vec{r}_0,\sigma')\vec{\sigma}_{\sigma',\sigma} \end{aligned}$$

Kinetic energy

The matrix elements for the operator that provides the kinetic energy in a given direction at a given point are

$$\langle\vec{x}_1,\dots,\vec{x}_N|\hat{T}|\vec{x'}_1,\dots,\vec{x'}_N\rangle=\sum_{i=1}^{N}\delta(\vec{x}_i-\vec{x'}_i)\frac{-\hbar^2}{2m_e}\vec{\nabla}_i'^2\prod_{k(k\neq i)}\delta(\vec{x}_k-\vec{x'}_k)$$

This expression is weird: The matrix element is a differential operator that points no-where! Not really. Let us investigate the meaning of this expression:

Consider a one-particle state $|f\rangle$ and let us define a second one as $|g\rangle\overset{\text{def}}{=}\frac{\hat{\vec{p}}^2}{2m_e}|f\rangle$. Then we find

$$g(\vec{r})=\langle\vec{r}|g\rangle=\langle\vec{r}|\frac{\hat{\vec{p}}^2}{2m_e}|f\rangle=\langle\vec{r}|\frac{\hat{\vec{p}}^2}{2m_e}\underbrace{\int d^3r'\;|\vec{r'}\rangle\langle\vec{r'}|}_{\hat{1}}|f\rangle=\int d^3r'\;\langle\vec{r}|\frac{\hat{\vec{p}}^2}{2m_e}|\vec{r'}\rangle\underbrace{\langle\vec{r'}|f\rangle}_{f(\vec{r'})}$$

On the other hand, we know that $g(\vec{r})=\frac{-\hbar^2}{2m_e}\vec{\nabla}^2f(\vec{r})$, so that we can formally identify the kinetic energy operator in a one-particle Hilbert space as

$$\langle\vec{r}|\frac{\hat{\vec{p}}}{2m_e}|\vec{r'}\rangle=\delta(\vec{r}-\vec{r'})\frac{-\hbar^2}{2m_e}\vec{\nabla}'^2$$

It is, however, important that the gradient stands in front of the function depending on $\vec{r'}$. I do not recommend to use tricks like this, because one has to be awfully careful.

The gradient $\vec{\nabla}_i'$ acts exclusively on the i-th primed position $\vec{r'}_i$. Note that the gradient cannot be interchanged with the δ-function preceding it, because a gradient does not commutate with a function of the corresponding coordinate.

Thus, we obtain the kinetic energy T as

$$\begin{aligned}T&\overset{\text{def}}{=}\langle\Psi|\hat{T}|\Psi\rangle=\int d^4x\int d^4x'\;\delta(\vec{x}-\vec{x'})\frac{-\hbar^2}{2m_e}\left(\vec{\nabla}'\right)^2\hat{\rho}^{(1)}(\vec{x'},\vec{x})\\&=\sum_{\sigma}\int d^3r\;\lim_{\vec{r}\to\vec{r'}}\frac{-\hbar^2}{2m_e}\vec{\nabla}'^2\hat{\rho}^{(1)}(\vec{r'},\sigma,\vec{r},\sigma)\end{aligned}$$

7.1.3 Two-particle density

In order to evaluate the Coulomb interaction energy between electrons, we need to know the two-particle density.

The matrix elements of the interaction in a real space representation are

$$\langle\vec{x}_1,\dots,\vec{x}_N|\hat{W}|\vec{x'}_1,\dots,\vec{x'}_N\rangle=\frac{1}{2}\sum_{i\neq j}w(\vec{x}_i,\vec{x}_j)\prod_{i=1}^{N}\delta(\vec{x}_i-\vec{x'}_i)$$

where $w(x,x')=\frac{e^2}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}$.

$$
\begin{aligned}
\langle\Psi|\hat{W}|\Psi\rangle &= \int d^4x_1\cdots\int d^4x_N\int d^4x'_1\cdots\int d^4x'_N \\
&\quad \rho^{(N)}(\vec{x}_1,\ldots,\vec{x}_N;\vec{x'}_1,\ldots,\vec{x'}_N)\langle\vec{x'}_1,\ldots,\vec{x'}_N|\hat{W}|\vec{x}_1,\ldots,\vec{x}_N\rangle \\
&= \int d^4x_1\cdots\int d^4x_N\int d^4x'_1\cdots\int d^4x'_N \\
&\quad \rho^{(N)}(\vec{x}_1,\ldots,\vec{x}_N;\vec{x'}_1,\ldots,\vec{x'}_N)\frac{1}{2}\sum_{i\neq j}w(\vec{x}_i,\vec{x}_j)\prod_{i=1}^{N}\delta(\vec{x}_i-\vec{x'}_i) \\
&= \int d^4x_1\cdots\int d^4x_N\;\rho^{(N)}(\vec{x}_1,\ldots,\vec{x}_N;\vec{x}_1,\ldots,\vec{x}_N)\frac{1}{2}\sum_{i\neq j}w(\vec{x}_i,\vec{x}_j) \\
&= \int d^4x_1\cdots\int d^4x_N\;\rho^{(N)}(\vec{x}_1,\ldots,\vec{x}_N;\vec{x}_1,\ldots,\vec{x}_N)\frac{N(N-1)}{2}w(\vec{x}_1,\vec{x}_2) \\
&= \frac{1}{2}\int d^4x\int d^4x'\;w(\vec{x},\vec{x'}) \\
&\quad \underbrace{N(N-1)\int d^4x_1\cdots\int d^4x_N\;\rho^{(N)}(\vec{x}_1,\ldots,\vec{x}_N;\vec{x}_1,\ldots,\vec{x}_N)\delta(\vec{x}-\vec{x}_1)\delta(\vec{x'}-\vec{x}_2)}_{n^{(2)}(\vec{x'},\vec{x})} \\
&= \frac{1}{2}\int d^4x\int d^4x'\;n^{(2)}(\vec{x},\vec{x'})w(\vec{x},\vec{x'})
\end{aligned}
$$

Thus, we define the **two-particle density**[3]. $n^{(2)}(\vec{x},\vec{x'})$ as

$$
\begin{aligned}
n^{(2)}(\vec{x},\vec{x'}) &\overset{\text{def}}{=} N(N-1)\int d^4x_1\cdots\int d^4x_N \\
&\qquad \rho^{(N)}(\vec{x}_1,\ldots,\vec{x}_N;\vec{x}_1,\ldots,\vec{x}_N)\delta(\vec{x}-\vec{x}_1)\delta(\vec{x'}-\vec{x}_2)
\end{aligned}
\tag{7.12}
$$

so that we can determine the interaction energy as

$$
\langle\Psi|\hat{W}|\Psi\rangle = \frac{1}{2}\int d^4x\int d^4x'\;n^{(2)}(\vec{x},\vec{x'})w(\vec{x},\vec{x'})
$$

The two-particle density can also be expressed as expectation value as follows

$$
\begin{aligned}
n^{(2)}(\vec{x},\vec{x'}) &\overset{Eq.\ 7.12}{=} \int d^4x_1\cdots\int d^4x_N\;\Psi(\vec{x}_1,\ldots,\vec{x}_N)\sum_{i\neq j}\delta(\vec{x}-\vec{x}_i)\delta(\vec{x'}-\vec{x}_j)\Psi^*(\vec{x}_1,\ldots,\vec{x}_N) \\
&= \langle\Psi|\sum_{i\neq j}\delta(\vec{x}-\vec{x}_i)\delta(\vec{x'}-\vec{x}_j)|\Psi\rangle
\end{aligned}
\tag{7.13}
$$

The two-particle density describes the probability density that there is one electron at $\vec{r}$ and another one at $\vec{r'}$. It integrates up to the number of ordered pairs, namely $N(N-1)$.

Note that the two-particle density is often confused with the one-particle density matrix. Both have the same number of arguments, but they are fundamentally different quantities.

The q-particle densities contain the information only about the probability to find a q-tuple of electrons with a given set of coordinates. It is, however, not possible to extract for example the momentum or the kinetic energy from a given q-particle density. In contrast, the q-electron density

---

[3] There are different conventions in use, which differ by a factor 1/2. We follow here the choice of McWeeny[41], who does not have the factor $\frac{1}{2}$. Here, the factor $\frac{1}{2}$, which compensates for the double counting of two pairs at $\vec{x}$ and $\vec{x'}$ (one obtained by exchanging the two particles from the other) must be done explicitly. On the other hand the two-particle density as defined by McWeeny has the simple physical interpretation of a density for one particle at $\vec{x}$ and another one at $\vec{x'}$, irrespective of their order.

matrix contains the complete information on q-tuple of electrons. For a q-particle system, the q-particle density matrix contains the complete information on the system, that is the same information contained in the wave function. Already from the one-particle density matrix we can determine the momentum or the kinetic energy.

In the following we list various densities and density-matrices in comparison, to show the difference

- one-particle density matrix

$$\rho^{(1)}(\vec{x}_1,\vec{x'}_1)=N\int d^4x_2\cdots\int d^4x_N\;\rho^{(N)}(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_N;\vec{x'}_1,\vec{x}_2,\ldots,\vec{x}_N)$$

- two-particle density matrix

$$\rho^{(2)}(\vec{x}_1,\vec{x}_2,\vec{x'}_1\vec{x'}_2)=N(N-1)\int d^4x_3\cdots\int d^4x_N\;\rho^{(N)}(\vec{x}_1,\vec{x}_2,\vec{x}_3,\ldots,\vec{x}_N;\vec{x'}_1,\vec{x'}_2,\vec{x}_3,\ldots,\vec{x}_N)$$

- one-particle density

$$n^{(1)}(\vec{x}_1)=N\int d^4x_2\cdots\int d^4x_N\;\rho^{(N)}(\vec{x}_1,\ldots,\vec{x}_N;\vec{x}_1,\ldots,\vec{x}_N)$$

- two-particle density

$$n^{(2)}(\vec{x}_1,\vec{x}_2)=N(N-1)\int d^4x_3\cdots\int d^4x_N\;\rho^{(N)}(\vec{x}_1,\ldots,\vec{x}_N;\vec{x}_1,\ldots,\vec{x}_N)$$

- N-particle density

$$n^{(N)}(\vec{x}_1,\ldots,\vec{x}_N)=\rho^{(N)}(\vec{x}_1,\ldots,\vec{x}_N;\vec{x}_1,\ldots,\vec{x}_N)$$

The q-particle density can directly be obtained from the q-particle density matrix

$$n^{(q)}(\vec{x}_1,\ldots,\vec{x}_q)=\rho^{(q)}(\vec{x}_1,\ldots,\vec{x}_q,\vec{x}_1,\ldots,\vec{x}_q)$$

The density matrix has always two arguments per particle while the density has only one per particle.

The two-particle density is so important for many-particle physics, because we can directly express the interaction energy by the two-particle density. The Coulomb interaction between electrons can be written as

$$E_H+E_X=\langle\Psi|\frac{1}{2}\sum_{i\neq j}\frac{e^2}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}|\Psi\rangle=\frac{1}{2}\sum_{\sigma,\sigma'}\int d^3r\int d^3r\;\frac{e^2n^{(2)}(\vec{x},\vec{x'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\qquad(7.14)$$

In chapter 6 we derived the expression for the expectation value of a two-particle operator with a Slater determinant. According to Eq. 7.13, the two-particle density can be expressed as an expectation value. For a single Slater determinant, the two-particle density would have the form

$$\begin{aligned}
n^{(2)}(\vec{r},\vec{r'}) &\overset{Eqs.\ 6.12,7.13}{=} \sum_{i,j}^{N}\int d^4x_1\int d^4x_2\;\phi_i^*(\vec{x}_1)\phi_j^*(\vec{x}_2)\delta(\vec{r}-\vec{r}_1)\delta(\vec{r}-\vec{r}_2)\phi_i(\vec{x}_1)\phi_j(\vec{x}_2)\\
&- \sum_{i,j}^{N}\int d^4x_1\int d^4x_2\;\phi_i^*(\vec{x}_1)\phi_j^*(\vec{x}_2)\delta(\vec{r}-\vec{r}_1)\delta(\vec{r}-\vec{r}_2)\phi_j(\vec{x}_1)\phi_i(\vec{x}_2)\\
&= n(\vec{r})n(\vec{r'})-\sum_{i,j}^{N}\sum_{\sigma,\sigma'}\phi_i^*(\vec{r},\sigma)\phi_j(\vec{r},\sigma)\phi_j^*(\vec{r'},\sigma')\phi_i(\vec{r'},\sigma')
\end{aligned}\qquad(7.15)$$

7.1.4 Two-particle density of the free electron gas

Let us work out the two-particle density of the free electron gas in the Hartree-Fock approximation:

For the free electron gas the one-particle orbitals are, according to Eq. 4.91,

$$\phi_{\vec{k},\sigma}(\vec{r},\sigma') = \frac{1}{\sqrt{\Omega}}\mathrm{e}^{i\vec{k}\vec{r}}\delta_{\sigma,\sigma'}$$

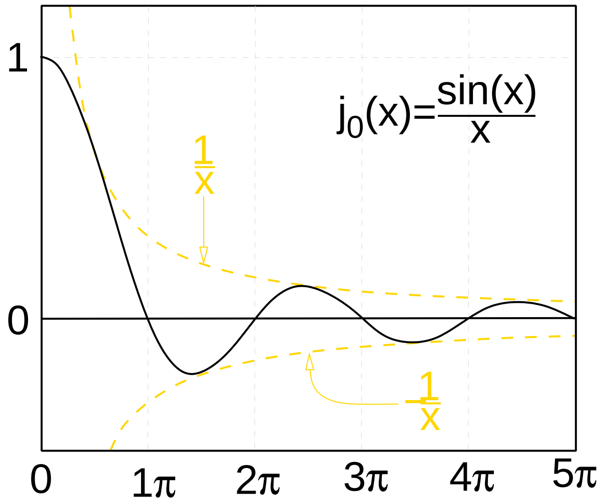
Now we need to determine the two-particle density according to Eq. 7.15. The first term is trivial because the density is constant. The exchange term is more complicated. let us therefore work out the expression $C(\vec{x},\vec{x'})$ which occurs in the exchange term of Eq. 7.15.

$$\begin{aligned}
C(\vec{x},\vec{x'}) &\stackrel{\text{def}}{=} \sum_{i,j}^{N}\phi_i^*(\vec{x})\phi_j(\vec{x})\phi_j^*(\vec{x'})\phi_i(\vec{x'}) \\
&\stackrel{Eq.\ 4.96}{=} \sum_{\sigma_1,\sigma_2}\Omega^2\int_{|\vec{k}_1|<k_F}\frac{d^3k_1}{(2\pi)^3}\int_{|\vec{k}_2|<k_F}\frac{d^3k_2}{(2\pi)^3}\ \phi^*_{\vec{k}_1,\sigma_1}(\vec{r},\sigma)\phi_{\vec{k}_2,\sigma_2}(\vec{r},\sigma)\phi^*_{\vec{k}_2,\sigma_2}(\vec{r'},\sigma')\phi_{\vec{k}_1,\sigma_1}(\vec{r'},\sigma') \\
&\stackrel{Eq.\ 4.91}{=} \int_{|\vec{k}_1|<k_F}\frac{d^3k_1}{(2\pi)^3}\int_{|\vec{k}_2|<k_F}\frac{d^3k_2}{(2\pi)^3}\ \mathrm{e}^{i(\vec{k}_2-\vec{k}_1)(\vec{r}-\vec{r'})}\underbrace{\sum_{\sigma_1,\sigma_2}\delta_{\sigma,\sigma_1}\delta_{\sigma,\sigma_2}\delta_{\sigma',\sigma_2}\delta_{\sigma',\sigma_1}}_{=\delta_{\sigma,\sigma'}\delta_{\sigma',\sigma}=\delta_{\sigma,\sigma'}} \\
&= \delta_{\sigma,\sigma'}\left(\int_{|\vec{k}_1|<k_F}\frac{d^3k_1}{(2\pi)^3}\mathrm{e}^{-i\vec{k}_1(\vec{r}-\vec{r'})}\right)\left(\int_{|\vec{k}_2|<k_F}\frac{d^3k_2}{(2\pi)^3}\mathrm{e}^{i\vec{k}_2(\vec{r}-\vec{r'})}\right) \\
&\stackrel{\vec{k}_1\to\vec{k}}{=} \delta_{\sigma,\sigma'}\left|\int_{|\vec{k}|<k_F}\frac{d^3k}{(2\pi)^3}\mathrm{e}^{i\vec{k}(\vec{r}-\vec{r'})}\right|^2
\end{aligned}$$

Now we use a tremendously useful theorem, that allows to decompose plane waves into radial functions and spherical harmonics $Y_{\ell,m}$

$$\mathrm{e}^{i\vec{k}\vec{r}} = 4\pi\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} i^\ell j_\ell(|\vec{k}||\vec{r}|)Y_{\ell,m}(\vec{r})Y^*_{\ell,m}(\vec{k}) \tag{7.16}$$

where the $j_\ell(z)$ are the **spherical Bessel functions** , also named **Bessel functions of half-integer order**. The Bessel function $j_0(x)$ is shown in the following figure



In the last result we see that it depends only on the relative positions for the two electrons. Therefore we introduce

$$\vec{u}\stackrel{\text{def}}{=}\vec{r}-\vec{r'}$$

Since we integrate over a spherical volume in reciprocal space, only the terms with $\ell = 0$ contribute a finite value, because all other spherical harmonics integrate to zero.

$$
\begin{aligned}
C(\vec{x},\vec{x}') &\overset{\vec{r}-\vec{r}'\to\vec{u}}{=} \delta_{\sigma,\sigma'}\left|\int_{|\vec{k}<k_F}\frac{d^3k}{(2\pi)^3}\mathrm{e}^{i\vec{k}\vec{u}}\right|^2 \\
&\overset{Eq.\ 7.16}{=} \delta_{\sigma,\sigma'}\left|\int_{|\vec{k}<k_F}\frac{d^3k}{(2\pi)^3}4\pi\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} i^\ell j_\ell(|\vec{k}||\vec{u}|)Y_{\ell,m}(\vec{u})Y^*_{\ell,m}(\vec{k})\right|^2 \\
&= \delta_{\sigma,\sigma'}\left|\frac{1}{(2\pi)^3}\int_{k<k_F}dk\ k^2 4\pi j_0(|\vec{k}||\vec{u}|)\quad\underbrace{\frac{1}{4\pi}}_{Y_{0,0}(\vec{u})Y^*_{0,0}(\vec{k})}\right|^2 \\
&\overset{|\vec{k}|\to k}{=} \delta_{\sigma,\sigma'}\frac{1}{(2\pi)^6}\left(4\pi\int_{k<k_F}dk\ k^2\frac{\sin(k|\vec{u}|)}{k|\vec{u}|}\right)^2 \\
&\overset{k|\vec{u}|\to x}{=} \delta_{\sigma,\sigma'}\frac{1}{(2\pi)^6}\left(\frac{4\pi}{|\vec{u}|^3}\int_{x<k_F|u|}dx\ x\sin(x)\right)^2
\end{aligned}
$$

The integral can be solved using

$$
\frac{d}{dx}\left(x\cos(x)-\sin(x)\right)=\cos(x)-x\sin(x)-cos(x)=x\sin(x)
$$

Now we continue

$$
\begin{aligned}
C(\vec{x},\vec{x}') &\overset{k|\vec{u}|\to x}{=} \delta_{\sigma,\sigma'}\frac{1}{(2\pi)^6}\left|\frac{1}{|\vec{u}|^3}\left(k_F|\vec{u}|\cos(k_F|\vec{u}|)-\sin(k_F|\vec{u}|)\right)\right|^2 \\
&= \frac{\delta_{\sigma,\sigma'}k_F^6}{(2\pi)^6}\left(\frac{k_F|\vec{u}|\cos(k_F|\vec{u}|)-\sin(k_F|\vec{u}|)}{(k_F|\vec{u}|)^3}\right)^2
\end{aligned}
$$

For the sake of esthetics[4], let scale the expression in parenthesis such that it starts with value 1 at $u = 0$. (See footnote below[5])

$$
\begin{aligned}
\lim_{x\to 0}\frac{x\cos(x)-\sin(x)}{x^3} &= \lim_{x\to 0}\sum_{n=0}^{\infty}\left(\frac{(-1)^n}{2n!}x^{2n-2}-\frac{(-1)^n}{(2n+1)!}x^{2n-2}\right) \\
&= \lim_{x\to 0}\sum_{n=0}^{\infty}\left(\frac{1}{(2n)!}-\frac{1}{(2n+1)!}\right)(-1)^n x^{2n-2} \\
&= \lim_{x\to 0}\sum_{n=0}^{\infty}(1-1)\,x^{-2}-\left(\frac{1}{2}-\frac{1}{6}\right)x^0+O(x^2)=\frac{1}{3}
\end{aligned}
$$

---

[4] esthetics=Ästetik"

[5]

$$
\begin{aligned}
&\mathrm{e}^x=\sum_{n=0}^{\infty}\frac{1}{n!}x^n \\
\Rightarrow\ &\mathrm{e}^{ix}=\sum_{n=0}^{\infty}\frac{1}{n!}i^n x^n=\sum_{n=0}^{\infty}\left(\frac{1}{(2n)!}(-1)^n x^{2n}+i\frac{1}{(2n+1)!}(-1)^n x^{2n+1}\right)=\cos(x)+i\sin(x) \\
\Rightarrow\ &\cos(x)=\sum_{n=0}^{\infty}\frac{(-1)^n}{2n!}x^{2n}\qquad;\qquad\sin(x)=\sum_{n=0}^{\infty}\frac{(-1)^{n+1}}{(2n+1)!}x^{2n+1}
\end{aligned}
$$

We rescale the term in parenthesis and obtain

$$C(\vec{x},\vec{x}') = \delta_{\sigma,\sigma'} \underbrace{\left(\frac{1}{(2\pi)^3}\frac{4\pi}{3}k_F^3\right)^2}_{\frac{1}{2}n} \left(3\frac{k_F|\vec{u}|\cos(k_F|\vec{u}|)-\sin(k_F|\vec{u}|)}{(k_F|\vec{u}|)^3}\right)^2$$

With $n = 2\cdot\frac{1}{(2\pi)^3}\frac{4\pi}{3}k_F^3$ we obtain the final result

$$\begin{aligned}
n^{(2)}(\vec{r},\vec{r}') &= n^2 - \sum_{\sigma,\sigma'} C(\vec{x},\vec{x}') \\
&= \sum_{\sigma,\sigma'} \left(\frac{n_\uparrow + n_\downarrow}{2}\right)^2 \left(1 - \delta_{\sigma,\sigma'}\left(3\frac{k_F|\vec{r}-\vec{r}'|\cos(k_F|\vec{r}-\vec{r}'|)-\sin(k_F|\vec{r}-\vec{r}'|)}{(k_F|\vec{r}-\vec{r}'|)^3}\right)^2\right) \\
&= (n_\uparrow + n_\downarrow)^2 \left(1 - \frac{1}{2}\left(3\frac{k_F|\vec{r}-\vec{r}'|\cos(k_F|\vec{r}-\vec{r}'|)-\sin(k_F|\vec{r}-\vec{r}'|)}{(k_F|\vec{r}-\vec{r}'|)^3}\right)^2\right)
\end{aligned}$$

The result is shown in Fig. 7.1. We observe that the two-particle density for identical spin direction vanishes at small distances in accordance with the Pauli principle, which excludes that two identical particles can be at the same position. The two-particle density for opposite spin is that of two uncorrelated electrons, namely the product of the two spin densities. For large distances the two-particle densities for both spin directions approach the uncorrelated value.

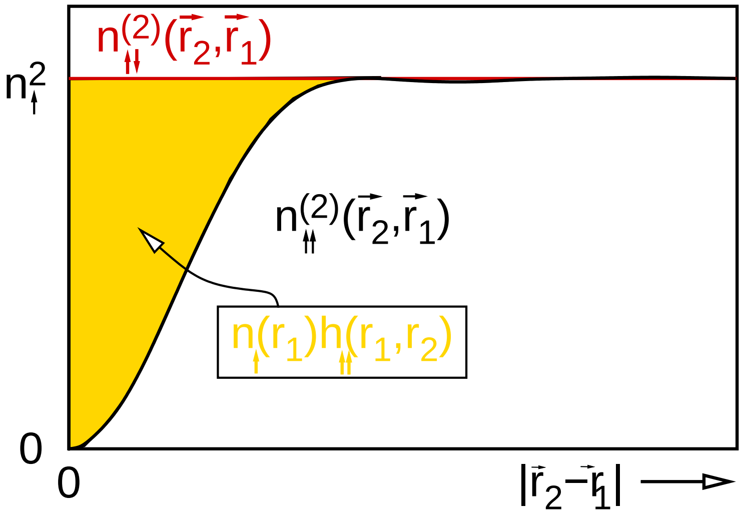


Fig. 7.1: Two-particle density of the non-interacting free electron gas.

7.1.5 Pair-correlation function and hole function

Another quantity that is useful is the **pair-correlation function**

PAIR-CORRELATION FUNCTION

$$g(\vec{x},\vec{x}_0)=\frac{n^{(2)}(\vec{x},\vec{x}_0)}{n(\vec{x}_0)}$$

The pair-correlation function describes the probability to find an electron with coordinates $\vec{x}$, given that there is one at $\vec{x}_0$. The pair-correlation function is the electron density seen by one electron that is located at $\vec{x}_0$. Consider an N-electron system. Pick out one electron and determine the density of the remaining $N-1$ electrons.

Given that an electron will see at large distances nothing but the charge density itself, it makes sense to define a more short-ranged quantity, namely the **exchange-correlation hole function** $h(\vec{x},\vec{x}_0)$

HOLE FUNCTION $H(\vec{X},\vec{X}_0)$

$$g(\vec{x},\vec{x}_0)=n(\vec{x})+h(\vec{x},\vec{x}_0)$$

The physical picture of the hole function is that each electron "sees" the total density $n(\vec{x})$ and the density $h(\vec{x},\vec{x'})$ of "one missing electron". The density $h(\vec{x},\vec{x'})$ is like a hole in the total charge density due to the fact that the remaining electrons do not come near to the electron in question. Thus, the Coulomb repulsion between the electrons is reduced relative to the Hartree energy by the attraction of the electron to its exchange-correlation hole.

Accurate calculations of the exchange-correlation hole have been obtained by quantum Monte Carlo calculations[42].

The two-particle density can now be expressed by the density and the exchange-correlation hole function as

$$n^{(2)}(\vec{x},\vec{x}_0)=n(\vec{x}_0)\left[n(\vec{x})+h(\vec{x},\vec{x}_0)\right]$$

Now we can also rewrite the interaction energy Eq. 7.14

$$E_{int}=\underbrace{\frac{1}{2}\sum_{\sigma,\sigma'}\int d^3r\int d^3r'\,\frac{e^2n(\vec{x})n(\vec{x'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}}_{E_H}+\underbrace{\sum_{\sigma}\int d^3r\;n(\vec{x})\frac{1}{2}\sum_{\sigma'}\int d^3r'\,\frac{e^2h(\vec{x'},\vec{x})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}}_{U_{xc}}$$

7.2 Self-made density functional

In order to understand, how density functionals are made, let us demonstrate the principle on a simple example. We will make our own, first-principles, density functional. It will not be terribly good, but it will not be a complete disaster either. However, it will be our own!

The typical process of creating a density functional is that we take all available information on the hole function, that may be relevant, and find an expression that reproduces this information as best as possible. As more and more information about the hole function becomes available, the functionals are continuously improved.

Apparently all the information about the interaction is contained in the hole function $h(\vec{r},\vec{r'})$. Let us collect all we know so far about the hole function:

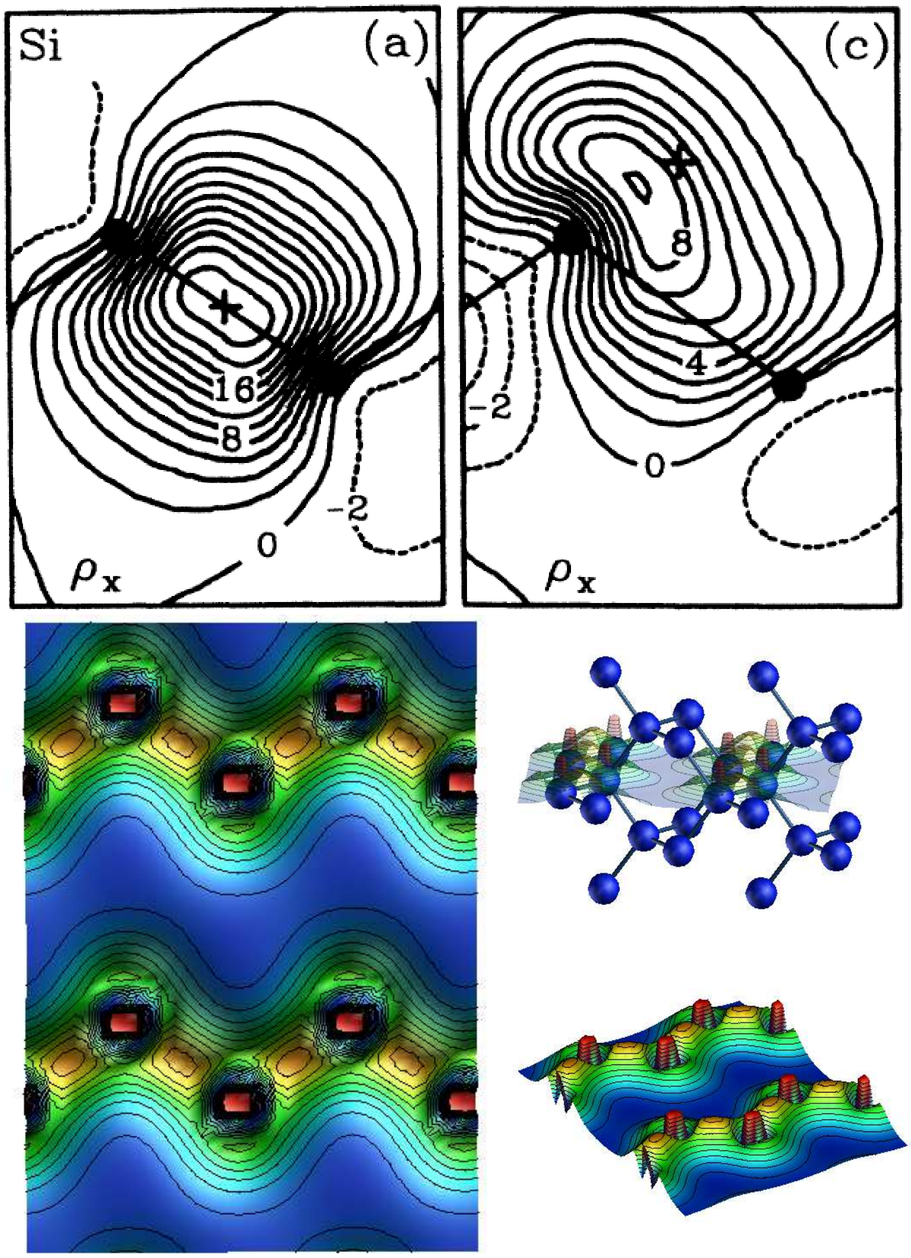


Fig. 7.2: Top: Exchange-correlation hole in the 110 plane of silicon[43]. The cross indicates the position of the reference electron. In the bond, the exchange-correlation hole is centered on the reference electron, while in the tail region it is located off-center. Bottom: Valence charge density of silicon in the (110) plane.

KNOWN PROPERTIES OF THE HOLE FUNCTION

- The hole function vanishes for large distances between the electrons

$$\lim_{|\vec{r}-\vec{r}_0|\to\infty} h(\vec{x},\vec{x}_0)=0$$

- The integral over the hole function with equal spins is equal to one, that is

$$\int d^3r\; h(\vec{x},\vec{x}_0)=-\delta_{\sigma,\sigma_0}$$

  This statement follows from the fact that an electron in an N-electron system, always sees the $N-1$ other electrons. Thus, the correlation function $\vec{g}$ integrates to $N-1$ and the hole function integrates to -1. The argument is easily extended to different spin directions.

Thus, we know the height, the width and the volume of the hole function. What is unknown is the shape. Let us assume the most simple shape, namely a spherical box centered at the reference electron, that is

$$h(\vec{x},\vec{x}_0) = -n(\vec{x}_0)\theta\left(\sqrt[3]{\frac{3}{4\pi n(\vec{x}_0)}} - |\vec{r}-\vec{r}_0|\right)\delta_{\sigma,\sigma_0} \tag{7.17}$$

where $\theta(x)$ is the Heaviside function.[6]

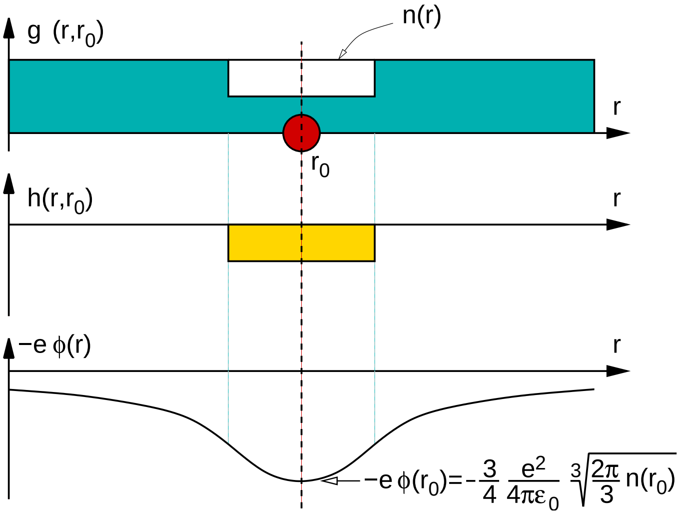


Fig. 7.3: Schematic representation of the xc-hole and the xc-potential energy. The electron density $n(\vec{r})$ is assumed to be constant. The electron shall be located at $\vec{r}_0$. The two-particle density $n(\vec{r},\vec{r}_0)$ describes the density of the $N-1$ electrons as experienced from the reference electron at $\vec{r}_0$. The two-particle density differs from the total density by the exchange hole $h(\vec{r}-\vec{r}')$. If we use the model that the exchange hole is a homogeneously charged sphere, the interaction energy of the reference electron and the exchange hole can be calculated as $v(\vec{r}_0)$.

---

[6] The Heaviside function is defined by

$$\theta(x) = \begin{cases} 0 \text{ for} & x<0 \\ \frac{1}{2} \text{ for} & x=0 \\ 1 \text{ for} & x>0 \end{cases}$$

Now we can determine the exchange energy as

$$\begin{aligned}
E_{xc} &= \int d^3r \sum_\sigma n(\vec{x})\frac{1}{2}\sum_{\sigma'}\int d^3r' \; \frac{e^2 h(\vec{x'},\vec{x})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|} \\
&= \int d^3r \sum_\sigma n(\vec{x})\frac{1}{2}\sum_{\sigma'}\int d^3r' \frac{e^2}{4\pi\epsilon_0}\frac{-n(\vec{x})\delta_{\sigma,\sigma'}\theta\left(\sqrt[3]{\frac{3}{4\pi n(\vec{x})}}-|\vec{r'}-\vec{r}|\right)}{|\vec{r'}-\vec{r}|} \\
&= \int d^3r \; \frac{1}{2}\frac{e^2}{4\pi\epsilon_0}\sum_\sigma n(\vec{x})\int d^3r' \frac{-n(\vec{x})\theta\left(\sqrt[3]{\frac{3}{4\pi n(\vec{x})}}-|\vec{r'}-\vec{r}|\right)}{|\vec{r'}-\vec{r}|} \\
&= \int d^3r \; \frac{1}{2}\frac{e^2}{4\pi\epsilon_0}\sum_\sigma n(\vec{x})\left[-\frac{3}{2\sqrt[3]{\frac{3}{4\pi n(\vec{x})}}}\right] \\
&= -\int d^3r \; \frac{3}{4}\sqrt[3]{\frac{4\pi}{3}}\frac{e^2}{4\pi\epsilon_0}\sum_\sigma n(\vec{x})^{\frac{4}{3}}
\end{aligned}$$

The integral over $\vec{r'}$ is easily solved directly. Here, we take a little detour, that may add some physical insight. The integral corresponds to the potential in the center of a homogeneously charged sphere with radius $s_0 = \sqrt[3]{\frac{3}{4\pi n}}$, with a total charge of one electron. The potential of a homogeneously charged sphere can be determined by fitting the potential outside and inside.

$$\Phi(\vec{s}) = \int d^3s' \; \frac{en\theta(s_0-|\vec{s'}|)}{4\pi\epsilon_0|\vec{s'}|} = \frac{e}{4\pi\epsilon_0}\begin{cases}\frac{1}{|\vec{s}|} & \text{for} \quad s > s_0 \\ \frac{1}{s_0}\left[\frac{3}{2}-\frac{1}{2}\left(\frac{|\vec{s}|}{s_0}\right)^2\right] & \text{for} \quad s < s_0\end{cases}$$

Thus, we obtain the exchange-correlation energy per electron within our model as

$$\epsilon_{xc}(n(\vec{r},\uparrow),n(\vec{r},\uparrow)) = -\frac{3}{4}\sqrt[3]{\frac{4\pi}{3}}\frac{e^2}{4\pi\epsilon_0}\frac{\sum_\sigma n(\vec{r},\sigma)^{\frac{4}{3}}}{\sum_\sigma n(\vec{r},\sigma)}$$

If the spin-densities of both spin directions are equal, we obtain, introducing the symbol $n_t(\vec{r})$ for the total electron density,

$$\epsilon_{xc}(\frac{1}{2}n_t(\vec{r}),\frac{1}{2}n_t(\vec{r})) = -\frac{3}{4}\sqrt[3]{\frac{4\pi}{3}}\frac{e^2}{4\pi\epsilon_0}\cdot\left(\frac{n_t(\vec{r})}{2}\right)^{\frac{1}{3}}$$

and if the all electrons have the same spin, we obtain

$$\epsilon_{xc}(n_t(\vec{r}),0) = -\frac{3}{4}\sqrt[3]{\frac{4\pi}{3}}\frac{e^2}{4\pi\epsilon_0}\cdot n_t(\vec{r})^{\frac{1}{3}}$$

The exchange-correlation energy is more favorable, if the electrons are all parallel. This is a reflection of **Hund's rule**.

In Fig. 7.4 the exchange-correlation energy per electron as function of the electron density is compared to the Hartree-Fock and the numerically correct result. Note that our model approaches the exact result for low densities. The deviation of our model is somewhat better than the Hartree-Fock result. Still, the errors are of order few eV, which is in the range of to binding energies. In an actual calculation of bond formation or bond-breaking the density changes only by a small amount and only in a local region, so that the errors will be acceptable.

Since we can now determine the interaction energy, at least in our model, we can write down a

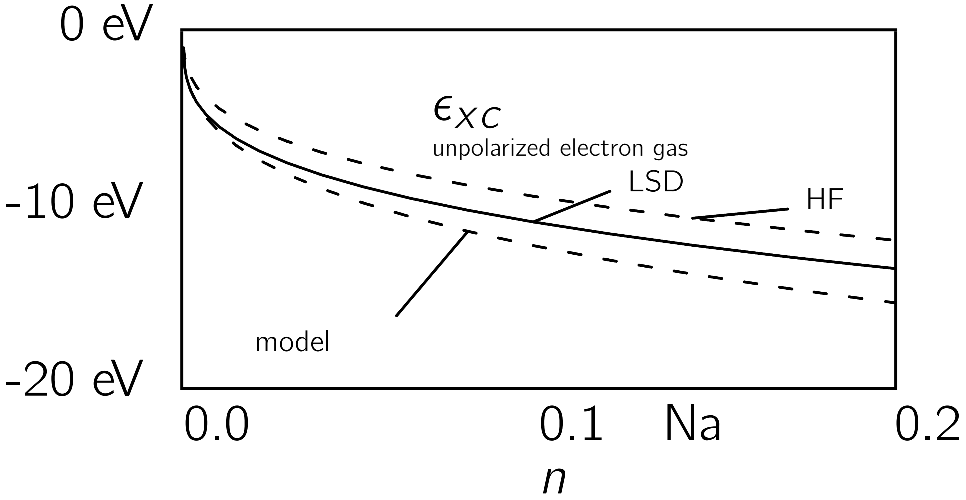


Fig. 7.4: Exchange-correlation energy per electron for a free electron gas as function of the density. The density is given in Hartree atomic units. The dashed line denoted "model" denotes the result obtained from the model of a homogeneously charged sphere as exchange-correlation hole. The dashed line denoted "HF" is the result obtained from Hartree-Fock. The correct result, which is the result from density-functional theory is denoted by "LSD". Note that the hard-sphere model describes the exchange-correlation energy per electron well in the low-density region, where the electrostatic repulsion dominates over the kinetic energy cost required to deform the exchange-correlation hole. The correct result lies in between the hard-sphere model and the Hartree-Fock result. This can be understood as follows: The exchange hole of the free electron gas (Hartree-Fock) does not yet adjust to the interaction. It corresponds to the limit where the kinetic energy determines the shape of the hole. The interaction makes the hole more compact than the true exchange-correlation hole and thus lowers the total energy. In the hard-sphere model, the hole is chosen to optimize the interaction energy. It corresponds to the limit where the kinetic energy cost to deform the hole is negligible. However, it does not consider the kinetic energy cost at all. Thus, it underestimates the true exchange-correlation energy.

total energy expression

$$
\begin{aligned}
E[n(\vec{x})] = & \underbrace{\sum_n f_n\langle\psi_n|\frac{\hat{p}^2}{2m_e}|\psi_n\rangle}_{E_{kin}} + \int d^3r\; n_t(\vec{r})v_{ext}(\vec{r}) \\
& + \underbrace{\frac{1}{2}\int d^3r \int d^3r'\; \frac{e^2 n_t(\vec{r})n_t(\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}}_{Hartree} \underbrace{- \int d^3r\; \frac{3}{4}\sqrt[3]{\frac{4\pi}{3}}\frac{e^2}{4\pi\epsilon_0}\sum_\sigma n(\vec{r},\sigma)^{\frac{4}{3}}}_{E_{xc}}
\end{aligned}
$$

The effective potential is obtained as the derivative of the potential energy

$$
v_{eff}(\vec{x}) = \frac{\delta E_{pot}}{\delta n(\vec{x})} = v_{ext}(\vec{r}) + \int d^3r'\; \frac{e^2 n_t(\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|} \underbrace{-\frac{e^2}{4\pi\epsilon_0}\sqrt[3]{\frac{4\pi}{3}}n(\vec{x})^{\frac{1}{3}}}_{\mu_{xc}(\vec{x})}
$$

Note, that the exchange-correlation potential differs from the exchange-correlation energy per electron. Both scale in the same way, but they differ by a factor.

There is a second observation. The effective potential is spin dependent. This spin dependence is responsible that we can calculate whether materials are magnetic or not. The potential has nothing to do with a magnetic interaction, even though it looks exactly like a magnetic field in the Pauli equation.

In fact the origin for the magnetization of materials is unrelated to the magnetic interactions. Its true origin is the so-called **exchange interaction**, which again roots in the Pauli principle. Note, however, that the exchange interaction is not a new physical form of interaction between particles.

Fertig: WS06/07 9 Doppelstunde 4.Dec.06

7.3 Constrained search

Density-functional theory says that all ground-state properties can be expressed, at least in principle, as functionals of the charge density alone. This is a nontrivial statement, because the wave function contains far more information than the density. It also provides a dramatic simplification as the density, in contrast to the wave function, can be handled on a computer: The density is a function in three dimensions, while the wave function is a function in 3^N dimensions.

The original proof of existence for density functionals goes back to Hohenberg and Kohn[44]. Later, Levy[45, 46] and Lieb[47] have given a recipe how the density functional can actually be constructed . This proof contains the proof of existence. For this reason, I will not present the original proof of Hohenberg and Kohn, but show the one by Levy.

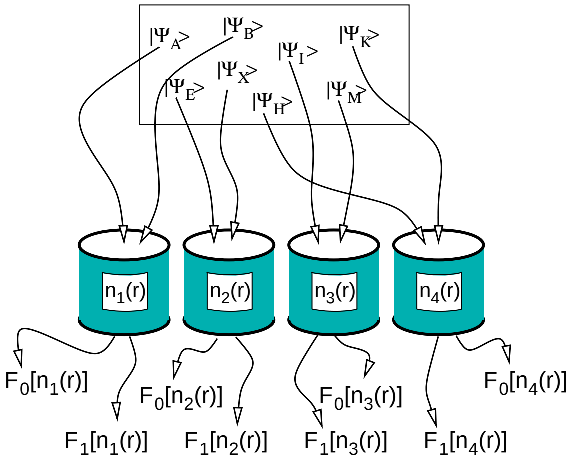


Fig. 7.5: Sketch of Levy's derivation of a density functional.

Let us start with a Hamilton operator

$$\hat{H}_\lambda = \hat{T} + \hat{V}_{ext} + \lambda\hat{W}$$

where

$$\hat{T} = \int d^4x_1 \cdots \int d^4x_N \, |\vec{x}_1, \ldots, \vec{x}_N\rangle \left(\sum_{i=1}^{N} \frac{-\hbar^2}{2m_e} \vec{\nabla}_i^2 \right) \langle \vec{x}_1, \ldots, \vec{x}_N|$$

is the kinetic energy operator,

$$\hat{V} = \int d^4x_1 \cdots \int d^4x_N \, |\vec{x}_1, \ldots, \vec{x}_N\rangle \left(\sum_{i=1}^{N} v_{ext}(\vec{r}_i) \right) \langle \vec{x}_1, \ldots, \vec{x}_N|$$

represents the external potential, and

$$\hat{W} = \int d^4x_1 \cdots \int d^4x_N \, |\vec{x}_1, \ldots, \vec{x}_N\rangle \left(\frac{1}{2} \sum_{i \neq j} \frac{e^2}{4\pi\epsilon_0 |\vec{r}_i - \vec{r}_j|} \right) \langle \vec{x}_1, \ldots, \vec{x}_N|$$

is the interaction energy operator. We have scaled the electron-electron interaction by a coupling parameter λ, which allows us later to switch the electron-electron interaction on an off. This is not required for the proof of Levy, but later on, we will often refer to Hamiltonians without interaction.

The idea of the proof is very simple:

1. Group all normalized, antisymmetric many-particle wave functions $\{|\Psi\rangle\}$ according to their density. Each such subset $\mathcal{A}[n(\vec{r})]$ contains all wave functions in the **Fock space** $\mathcal{F}$ with the same density.

$$\mathcal{A}[n(\vec{r})] = \{|\Psi\rangle \in \mathcal{F} : \langle\Psi|\hat{n}(\vec{r})|\Psi\rangle = n(\vec{r})\}$$

   The Fock space is the union[7] of all N-particle Hilbert spaces $\mathcal{H}^N$, that is

$$\mathcal{F} \stackrel{\text{def}}{=} \bigcup_{N=0}^{\infty} \mathcal{H}^N$$

   The Fock space will be introduced in more detail later. If the reader feels uncomfortable with this concept, he may replace the Fock space with the N-particle Hilbert space for a specific N.

2. For each subset $\mathcal{A}[n(\vec{r})]$, that is for each density, we determine that wave function $|\Psi^{(0)}_\lambda[n(\vec{r})]\rangle$ that provides the lowest energy $\langle\Psi|\hat{H}_\lambda|\Psi\rangle$. The minimum of the total energy defines the total energy functional $E_\lambda[n(\vec{r})]$ of the electron density.

$$E_\lambda[n(\vec{r})] = \min_{|\Psi\rangle\in\mathcal{A}[n(\vec{r})]} \langle\Psi|\hat{H}_\lambda|\Psi\rangle = \langle\Psi^{(0)}_\lambda[n(\vec{r})]|\hat{H}_\lambda|\Psi^{(0)}_\lambda[n(\vec{r})]\rangle$$

3. The energy can be divided into **universal density functional** $F_\lambda[n(\vec{r})]$, which is a general property of the electron gas, and a system-specific term, that depends on the external potential.

$$E_\lambda[n(\vec{r})] = \underbrace{F_\lambda[n(\vec{r})]}_{\langle\Psi^{(0)}_\lambda[n]|\hat{T}+\lambda\hat{W}|\Psi^{(0)}_\lambda[n]\rangle} + \underbrace{\int d^3r\; n(\vec{r})v_{ext}(\vec{r})}_{\langle\Psi^{(0)}_\lambda[n]|\hat{V}|\Psi^{(0)}_\lambda[n]\rangle}$$

   Because the energy related to the external potential only depends on the density, it is a constant within each group of a given density. Thus, we can directly determine the universal functional $F_\lambda[n(\vec{r})]$ as minimum of $\langle\Psi|\hat{T}+\lambda\hat{W}|\Psi\rangle$ for all wave functions with the specified density.

   UNIVERSAL DENSITY FUNCTIONAL

$$F_\lambda[n(\vec{r})] \stackrel{\text{def}}{=} \min_{|\Psi\rangle\in\mathcal{A}[n(\vec{r})]} \langle\Psi|\hat{T}+\lambda\hat{W}|\Psi\rangle$$

4. To find the ground-state energy $E^{(0)}_\lambda$ of an electron gas with N electrons in a given external potential, we minimize the total energy functional under the constraint of a specified total number of electrons.

$$E^{(0)}_\lambda = \min_{n(\vec{r}):\int d^3r\; n(\vec{r})=N} \left\{ F_\lambda[n(\vec{r})] + \int d^3r\; n(\vec{r})v_{ext}(\vec{r}) \right\}$$

   The Lagrange parameter μ is the chemical potential of the electrons. This minimization with respect to the density provides us with the ground-state density and the ground-state energy. If we have remembered the wave function $|\Psi^0[n(\vec{r})]\rangle$ that minimized the total energy within a group with given density $n(\vec{r})$, we can, in principle work our way back to the ground-state wave function, which contains all information about the electronic ground state.

---

[7] germ.:union=Vereinigung

This concludes the proof by Levy on the existence of a universal functional of the density, that allows to determine the ground-state energy and, in principle, all other properties of the electronic ground state.

7.4 Self-consistent equations

In the following I will discuss, how this proof can be used for a practical scheme to find the electronic ground-state energy and density under the assumption that the universal functional is known.

We will be lead to a system of equations, that must be fulfilled simultaneously for the electronic ground state. Then we will show, how this system of equations can be solved iteratively.

7.4.1 Universal Functional

Let us write down the universal functional $F_\lambda[n(\vec{r})]$ discussed in the proof of Levy more explicitly.

$$F_\lambda[n(\vec{r})] = \underset{\eta}{\mathrm{extr}}\ \underset{\gamma(\vec{r})}{\mathrm{extr}}\ \min_{|\Psi\rangle} \Big\{ \langle\Psi|\hat{T}+\lambda\hat{W}|\Psi\rangle + \int d^3r\ \gamma(\vec{r})\Big[\langle\Psi|\hat{n}(\vec{r})|\Psi\rangle - n(\vec{r})\Big] - \eta\Big[\langle\psi|\psi\rangle - 1\Big] \Big\}$$

Instead of explicitly grouping the wave functions we enforce the constraints by the method of Lagrange multipliers $\gamma(\vec{r})$ and η. Note, that we need not specify the chemical potential here, because the particle number is fixed by the specification of the density $n(\vec{r})$.

The operator $\hat{n}(\vec{r})$ yields the one-particle density of a many-particle wave function.[8] It is defined by

$$\hat{n}(\vec{r}) = \int d^4x_1 \cdots \int d^4x_N\ |\vec{x}_1,\ldots,\vec{x}_N\rangle \left(\sum_{i=1}^{N} \delta(\vec{r}-\vec{r}_i)\right) \langle\vec{x}_1,\ldots,\vec{x}_N|$$

This means that we minimize the total energy $\langle\psi|\hat{H}_\lambda|\psi\rangle$ under the constraint that the density of the wave function is identical to $n(\vec{r})$ specified beforehand, namely.

$$\langle\Psi|\hat{n}(\vec{r})|\Psi\rangle = n(\vec{r})$$

The corresponding Lagrange multiplier is $\gamma(\vec{r})$. It is position dependent, because there is a constraint for every position. In addition, we need to impose the constraint that the wave function is normalized, that is

$$\langle\Psi|\Psi\rangle = 1$$

This normalization condition is automatically fulfilled only if we restrict the search to N particle wave function and the density $n(\vec{r})$ integrates to N. However, if we allow for wave functions with variable particle number, the norm-condition must be imposed as an extra condition.

---

[8] Thus

$$\langle\Psi|\hat{n}(\vec{r})|\Psi\rangle = \int d^4x_1 \cdots \int d^4x_N\ \langle\Psi|\vec{x}_1,\ldots,\vec{x}_N\rangle \left(\sum_{i=1}^{N} \delta(\vec{r}-\vec{r}_i)\right) \langle\vec{x}_1,\ldots,\vec{x}_N|\Psi\rangle$$
$$= \sum_{i=1}^{N} \int d^4x_1 \cdots \int d^4x_N\ \Psi^*(\vec{x}_1,\ldots,\vec{x}_N)\delta(\vec{r}-\vec{r}_i)\Psi(\vec{x}_1,\ldots,\vec{x}_N)$$

The equilibrium condition results in a Schrödinger equation

$$\left[\hat{T} + \lambda\hat{W} + \int d^3r\ \gamma(\vec{r})\hat{n}(\vec{r}) - \eta\right]|\Psi\rangle = 0$$

The wave function obeys a Schrödinger equation in an external potential $\gamma(\vec{r})$. The energy in that potential is η.

7.4.2 Exchange-correlation functional

Now we split off from the functional $F_1[n]$, that is the universal density functional of the interacting electron gas, two terms that are fairly easy to evaluate. One is the Hartree energy E_H and the other is the kinetic energy of a non-interacting electron gas with the same density, which is simply $F_0[n]$. As a result, we obtain the **exchange and correlation energy**

EXCHANGE-CORRELATION FUNCTIONAL

$$E_{\rm xc}[n(\vec{r})] \overset{\rm def}{=} F_1[n(\vec{r})] - F_0[n(\vec{r})] - \underbrace{\frac{1}{2}\int d^3r \int d^3r'\ \frac{e^2 n(\vec{r})n(\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}}_{E_H}$$

Note that F_1 contains the kinetic energy of the interacting electron system and the electrostatic interaction energy, while F_0 is the kinetic energy of a non-interacting electron gas with the same density. Even though the wave functions of the interacting and the non-interacting electron gases have the same density, they are different, because they minimize different Hamiltonians.

The total energy functional of the interacting electron system can be written with the help of the exchange and correlation energy as

$$\begin{aligned} E_1[n(\vec{r})] &= F_0[n(\vec{r})] + \frac{1}{2}\int d^3r \int d^3r'\ \frac{e^2 n(\vec{r})n(\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|} \\ &\quad + \int d^3r\ n(\vec{r})v_{ext}(\vec{r}) + E_{\rm xc}[n(\vec{r})] \end{aligned} \tag{7.18}$$

It seems as if we have not gained much because now we have to deal with two functionals F_0 and E_{xc}. The advantage is that F_0 is the functional for non-interacting electrons, which we can evaluate, and that E_{xc} is a small term that can be approximated. Important is also the E_{xc} can be expressed in physically intuitive terms, so that we can make reasonable assumptions about it.

7.4.3 The self-consistency condition

In order to determine the minimum of the total energy, we need to be able to determine the density which fulfills

$$\frac{\delta E_1}{\delta n(\vec{r})} \overset{Eq.\ 7.18}{=} \frac{\delta F_0}{\delta n(\vec{r})} + \int d^3r'\ \frac{e^2 n(\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|} + v_{ext}(\vec{r}) + \mu_{xc}(\vec{r}) - \mu = 0 \tag{7.19}$$

where the exchange-correlation potential is defined as

$$\mu_{xc}(\vec{r}) \overset{\rm def}{=} \frac{\delta E_{xc}}{\delta n(\vec{r})}$$

The exchange-correlation potential μ_{xc} should not be confused with the chemical potential μ.

We can form all the functional derivatives of Eq. 7.19 with the exception of the functional derivative of F_0. In the following we will determine it.

We will make use of a **Legendre transform** of the universal density functional of the non-interacting electron gas.

As a reminder of the Legendre transform, let us take a little detour and repeat as example for a Legendre transform. We look into how the enthalpy $H(p)$, which is a function of the pressure, is constructed from the total energy $E(V)$, given as function of the volume. We start to define the enthalpy as the minimum total energy of a system connected to a volume reservoir. The latter is characterized by a pressure.

$$H(p)=\min_V E(V)+pV$$

as minimum condition we obtain

$$\frac{dE}{dV}+p=0 \qquad \Rightarrow \qquad p(V^{(0)})=-\left.\frac{dE}{dV}\right|_{V^{(0)}}$$

This equation is resolved to obtain the optimum volume $V^{(0)}(p)$ for a given pressure. Thus, we can express the enthalpy as a function of the pressure alone, namely

$$H(p)=E(V^{(0)}(p))+pV^{(0)}(p)$$

Similarly we can obtain the energy from the enthalpy by a Legendre backtransform.

$$E(V)=\max_p H(p)-pV$$

which yields the minimum condition

$$\frac{dH}{dp}-V=0$$

from which we obtain the pressure for a given volume $p^{(0)}(V)$, which we insert to obtain the energy as function of volume

$$E(V)=H(p^{(0)}(V))-p^{(0)}(V)V$$

Thus, we showed how the Legendre transform works back and forth. The Legendre transform is unique if the function is either fully concave or fully convex.

After this little detour, let us Legendre transform the universal density functional.

$$Q[u(\vec{r}]=\min_{n(\vec{r})}\left\{F_0[n]+\int d^3r\; n(\vec{r})u(\vec{r})\right\}$$

The minimum condition with respect to the density yields

$$\left.\frac{\delta F_0}{\delta n(\vec{r})}\right|_{n^{(0)}([v_{eff}],\vec{r})}+u(\vec{r})=0$$

which specifies a density $n^{(0)}([u],\vec{r})$ for each $u(\vec{r})$, so that

$$Q[u]=F_0[n^{(0)}]+\int d^3r\; n^{(0)}(\vec{r})\; u(\vec{r})$$

The functional $Q[u]$ has an intuitive interpretation and can be calculated in a straightforward manner. It is simply the ground-state total energy of a non-interacting electron gas in an external potential $v_{eff}(\vec{r})\overset{\text{def}}{=}u(\vec{r})+\mu$, where the one-particle orbitals are occupied up to the chemical potential

μ. Because the ground state of a non-interacting system is represented by a Slater determinant we can express Q by one-particle orbitals.

$$Q[u] = \sum_n \langle \phi_n | \frac{\hat{\vec{p}}^2}{2m_e} + \hat{v}_{eff} | \phi_n \rangle = \sum_n \epsilon_n$$

where the one-particle orbitals obey the Schrödinger equation

$$\left[\frac{-\hbar^2}{2m_e} \vec{\nabla}^2 + v_{eff}(\vec{r}) - \epsilon_n \right] \phi_n(\vec{r}) = 0$$

and have energies smaller than the chemical potential μ. Note that the chemical potential drops out of the equation. It has been introduced here to draw the connection to the known physical quantities.

After we know how to obtain the Legendre transform $Q[u]$, the universal density functional of the non-interacting electron gas F_0 can be obtained by a Legendre backtransform

$$F_0[n] = \max_u \left\{ Q[u] - \int d^3r \; n(\vec{r}) u(\vec{r}) \right\} \tag{7.20}$$

We see again that F_0 corresponds to the kinetic energy of a non-interacting electron gas.

The minimum condition defines an effective potential $u^{(0)}([n], \vec{r})$ as function of the density

$$\left. \frac{\delta Q}{\delta u(\vec{r})} \right|_{u^{(0)}(\vec{r})} - n(\vec{r}) = 0 \tag{7.21}$$

so that

$$F_0[n(\vec{r})] = Q[u^{(0)}([n], \vec{r})] - \int d^3r \; n(\vec{r}) u^{(0)}([n], \vec{r})$$

Now we can form the derivative with respect to the density

$$\begin{aligned} \frac{\delta F_0}{\delta n(\vec{r})} &= \int d^3r' \underbrace{\left(\frac{\delta Q}{\delta u(\vec{r'})} - n(\vec{r'}) \right)}_{=0} \frac{\delta u^{(0)}(\vec{r'})}{\delta n(\vec{r})} - u^{(0)}(\vec{r}) \\ &\overset{Eq.\ 7.21}{=} -u^{(0)}(\vec{r}) \end{aligned} \tag{7.22}$$

Thus, we found an expression for the derivative of $F_0[n]$ and a recipe for constructing it by solving a Schrödinger equation for a non-interacting electron gas.

Now, we can insert this result, namely Eq. 7.22, into the minimum condition Eq. 7.19 of the total energy

$$\frac{\delta E}{\delta n(\vec{r})} = \underbrace{-u^{(0)}([n], \vec{r})}_{\frac{\delta F_0}{\delta n(\vec{r})}} + \int d^3r' \; \frac{e^2 n(\vec{r'})}{4\pi\epsilon_0 |\vec{r} - \vec{r'}|} + v_{ext}(\vec{r}) + \mu_{xc}(\vec{r}) - \mu = 0 \tag{7.23}$$

which is an equation that specifies the effective potential $v_{eff}(\vec{r}) = u(\vec{r}) + \mu$, that yields the ground-state density, as

$$\underbrace{v_{eff}(\vec{r})}_{u(\vec{r})+\mu} = v_{ext}(\vec{r}) + \int d^3r' \; \frac{e^2 n(\vec{r'})}{4\pi\epsilon_0 |\vec{r} - \vec{r'}|} + \mu_{xc}(\vec{r}) \tag{7.24}$$

where $n(\vec{r})$ is the ground-state density of the non-interacting electron gas in the potential $v_{eff}(\vec{r})$ and a chemical potential μ.

7.4.4 Flow chart for a conventional self-consistency loop

Thus, we can now list the equation system that defines the ground-state energy

$$\left[\frac{\hat{\vec{p}}^2}{2m_e} + v_{eff}(\hat{\vec{r}}) - \epsilon_n\right]|\phi_n\rangle = 0$$

$$n(\vec{r}) = \sum_{n=1}^{N} \phi_n^*(\vec{r})\phi_n(\vec{r})$$

$$v_{eff}(\vec{r}) = v_{ext}(\vec{r}) + \int d^3r' \frac{e^2 n(\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|} + \mu_{xc}(\vec{r})$$

In Fig. 7.6 the steps required to solve this coupled system of equations is shown.

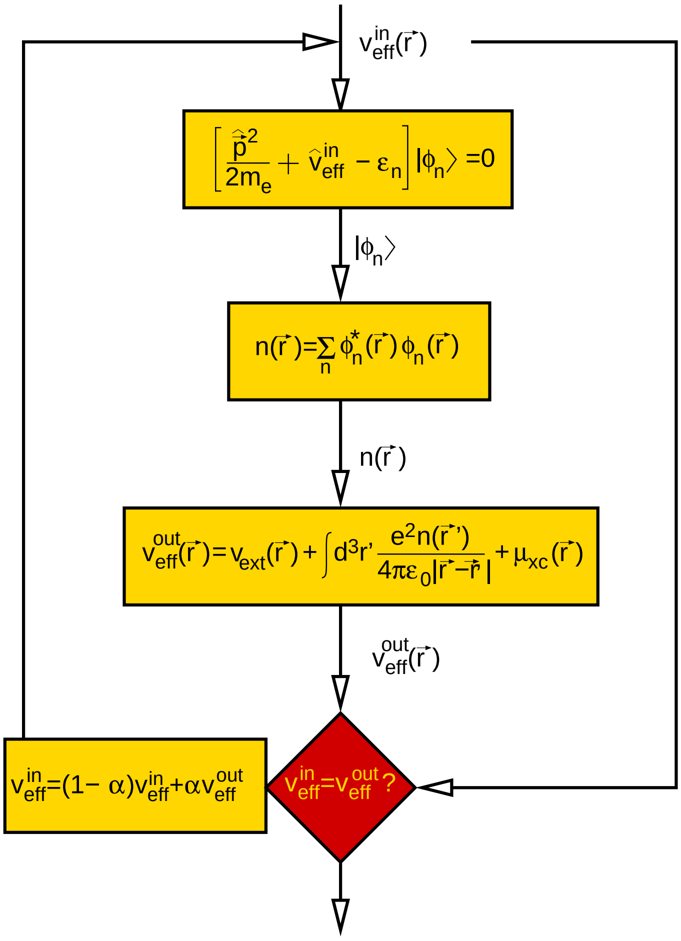


Fig. 7.6: Self-consistency cycle

7.4.5 Another Minimum principle

The total energy functional has the form

$$E_1[n] = F_0[n] + \frac{1}{2}\int d^3r \int d^3r' \frac{e^2 n(\vec{r})n(\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|} + \int d^3r\, n(\vec{r})v_{\rm ext}(\vec{r}) + E_{\rm xc}[n(\vec{r})]$$

Because the ground state of a non-interacting electron gas is a Slater determinant, we can define the functional $F_0[n]$ by minimizing over Slater determinants only. This allows us to describe F_0 as minimum over one-particle wave functions.

The functional F_0 can therefore be written as

$$F_0[\vec{n}] = \underset{\Lambda_{n,m}}{\mathrm{extr}} \underset{v_{eff}(\vec{r})}{\mathrm{extr}} \min_{|\phi_n\rangle} \Bigg\{ \sum_n \langle \phi_n | \frac{\hat{\vec{p}}^2}{2m_e} | \phi_n \rangle$$
$$+ \int d^3r \, v_{eff}(\vec{r}) \left[\sum_n \phi_n^*(\vec{r})\phi_n(\vec{r}) - n(\vec{r}) \right] - \sum_{n,m} \left(\langle \phi_n | \phi_m \rangle - \delta_{n,m} \right) \Lambda_{m,n} \Bigg\}$$

where v_{eff} plays the role of a Lagrange multiplier. This allows us to express the derivative of the universal density functional of the non-interacting electrons as

$$\frac{\delta F_0}{\delta n(\vec{r})} = v_{eff}(\vec{r})$$

In practical calculation we (1) chose an arbitrary potential v_{eff}. Then (2) we solve the Schrödinger equation, to obtain the one-particle wave functions, and the kinetic energy of the non-interacting electrons. Now (3) we evaluate the density. The kinetic energy and the density allows us (4) to evaluate the total energy. The density provides us (5) with a new expression for v_{eff}. Out of the last and the new effective potential one creates (6) a better choice for the effective potential, which is used as new start potential for (1). If the effective potentials, the one used for the Schrödinger equation and the one obtained from the density, are identical, we have found the desired ground-state energy and density.

Fertig: WS06/07 10 Doppelstunde 10.Dec.06

7.5 Adiabatic connection

We have seen so far that one can reasonably estimate the exchange-correlation hole, and evaluate the energy of the electron in the potential of the exchange-correlation hole. However, it is nearly hopeless to estimate the kinetic energy contribution to the exchange-correlation energy.

If there would be no cost to deform the exchange-correlation hole, the latter would simply adjust such that the electrostatic energy is optimized. One can easily estimate how the hole would look like: It would have the shape of a hard sphere centered at the reference electron. Its radius would be determined by the sum rule that says that there is exactly one positive charge in the exchange-correlation hole. The reason for this shape is that the electron would repel its neighbors as much as possible, in order to reduce its electrostatic interaction.

However as the exchange-correlation hole is deformed, there is a price to pay, namely that the kinetic energy goes up. In the extreme case of a hard-sphere like hole, the wave functions would have steps, that is infinite curvature. The kinetic energy is just a measure of the curvature of the wave function. Thus, the deformation of the hole costs kinetic energy. The final shape of the XC-hole is therefore a balance between electrostatic gain and kinetic energy cost.

The **adiabatic connection**[48, 49, 50] is a theorem that allows to relate the net energy gain, including the kinetic energy cost, to the electrostatic energy gain only, but for variable strengths of the interaction energy. While this theorem looks like mystery, it cannot be applied directly. Nevertheless it provides an important route that allows important approximations to be made. The theorem is demonstrated in the following:

We consider a given electron density for which the non-interacting functional $F_0[n(\vec{r})]$ is exactly known.

$$F_\lambda[n(\vec{r})] = \langle \Psi_\lambda | \hat{T} + \lambda \hat{W} | \Psi_\lambda \rangle$$
$$+ \int d^3r \, \mu_\lambda(\vec{r}) \Big(\langle \Psi_\lambda | \hat{n}(\vec{r}) | \Psi_\lambda \rangle - n(\vec{r}) \Big) - E_\lambda \Big(\langle \Psi_\lambda | \Psi_\lambda \rangle - 1 \Big)$$

where $|\Psi_\lambda\rangle$ is the many electron wave function that minimizes F_λ under the constraint that the density of the wave function is equal to $n(\vec{r})$. The function $\mu_\lambda(\vec{r})$ represents the Lagrange parameters. The operator

$$\hat{n}(\vec{r})\overset{\text{def}}{=}\int d^4x_1\cdots\int d^4x_N\,|\vec{x}_1,\ldots,\vec{x}_N\rangle\left(\sum_{i=1}^{N}\delta(\vec{r}-\vec{r}_i)\right)\langle\vec{x}_1,\ldots,\vec{x}_N|$$

is the operator that extracts the electron density from the wave function.

The wave functions that minimize the functional obey the Schrödinger equation

$$\left[\hat{T}+\lambda\hat{W}+\int d^3r\,\mu_\lambda(\vec{r})\hat{n}(\vec{r})-E_\lambda\right]|\Psi_\lambda\rangle=0 \tag{7.25}$$

This means that as the interaction strength is increased, a local potential, $\mu_\lambda(\vec{r})$, is switched on that ensures that the electron density remains unchanged.

If we know the derivatives of the functional with respect to the interaction strength λ, and if we know the functional for the non-interacting electron gas, we can determine the functional for the interacting electron gas as

$$\begin{aligned}F_1[n(\vec{r})]&=F_0[n(\vec{r})]+\int_0^1 d\lambda\,\frac{dF_\lambda[n(\vec{r})]}{d\lambda}\\&=\langle\Psi_0|\hat{T}|\Psi_0\rangle+\int_0^1 d\lambda\,\frac{dF_\lambda[n(\vec{r})]}{d\lambda}\end{aligned}$$

The first term is simply the kinetic energy of the non-interacting electron gas with the same density as the interacting system.

Now we can show that the derivative of the functional with respect to the interaction strength can be determined from the interaction energy alone. The proof rests on the **Hellmann-Feynman theorem**.

$$\begin{aligned}\frac{dF_\lambda[n(\vec{r})]}{d\lambda}&=\frac{d}{d\lambda}\Big[\langle\Psi_\lambda|\hat{T}+\lambda\hat{W}|\Psi_\lambda\rangle\\&\quad+\int d^3r\,\mu_\lambda(\vec{r})\Big(\langle\Psi_\lambda|\hat{n}(\vec{r})|\Psi_\lambda\rangle-n(\vec{r})\Big)-E_\lambda\Big(\langle\Psi_\lambda|\Psi_\lambda\rangle-1\Big)\Big]\\&=\langle\frac{d\Psi_\lambda}{d\lambda}|\hat{T}+\lambda\hat{W}|\Psi_\lambda\rangle+\langle\Psi_\lambda|\hat{W}|\Psi_\lambda\rangle+\langle\Psi_\lambda|\hat{T}+\lambda\hat{W}|\frac{d\Psi_\lambda}{d\lambda}\rangle\\&\quad+\int d^3r\,\frac{d\mu_\lambda(\vec{r})}{d\lambda}\underbrace{\left(\langle\Psi_\lambda|\hat{n}(\vec{r})|\Psi_\lambda\rangle-n(\vec{r})\right)}_{=0}\\&\quad+\int d^3r\,\mu_\lambda(\vec{r})\left(\langle\frac{d\Psi_\lambda}{d\lambda}|\hat{n}(\vec{r})|\Psi_\lambda\rangle+\langle\Psi_\lambda|\hat{n}(\vec{r})|\frac{d\Psi_\lambda}{d\lambda}\rangle\right)\\&\quad-\frac{dE_\lambda}{d\lambda}\underbrace{\Big(\langle\Psi_\lambda|\Psi_\lambda\rangle-1\Big)}_{=0}-E_\lambda\Big(\langle\frac{d\Psi_\lambda}{d\lambda}|\Psi_\lambda\rangle+\langle\Psi_\lambda|\frac{d\Psi_\lambda}{d\lambda}\rangle\Big)\\&=\langle\Psi_\lambda|\hat{W}|\Psi_\lambda\rangle\\&\quad+\langle\frac{d\Psi_\lambda}{d\lambda}|\underbrace{\Big[\hat{T}+\lambda\hat{W}+\int d^3r\,\mu_\lambda(\vec{r})\hat{n}(\vec{r})-E_\lambda\Big]|\Psi_\lambda\rangle}_{=0}\\&\quad+\underbrace{\langle\Psi_\lambda|\Big[\hat{T}+\lambda\hat{W}+\int d^3r\,\mu_\lambda(\vec{r})\hat{n}(\vec{r})-E_\lambda\Big]}_{=0}|\frac{d\Psi_\lambda}{d\lambda}\rangle\\&\overset{Eq.\ 7.25}{=}\langle\Psi_\lambda|\hat{W}|\Psi_\lambda\rangle\end{aligned}$$

Thus, we can express the functional for the finite interaction strength as

$$F_1[n(\vec{r})] = \langle\Psi_0|\hat{T}|\Psi_0\rangle + \underbrace{\int_0^1 d\lambda\ \langle\Psi_\lambda|\hat{W}|\Psi_\lambda\rangle}_{E_H[n(\vec{r})]+E_{xc}[n(\vec{r})]}$$

Thus, we have obtained an explicit expression for the exchange-correlation energy, that does not directly refer to the kinetic energy.

$$E_{xc}[n(\vec{r})] = \int_0^1 d\lambda\ \langle\Psi_\lambda|\hat{W}|\Psi_\lambda\rangle - \underbrace{\frac{1}{2}\int d^3r \int d^3r'\ \frac{e^2 n(\vec{r})n(\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}}_{E_H}$$

This expression has the problem that we still need the wave functions $|\Psi_\lambda\rangle$ for all possible values for the interaction strength.

The interaction energy can further be expressed by the exchange-correlation hole for a given value of the interaction, namely

$$\langle\Psi_\lambda|\hat{W}|\Psi_\lambda\rangle = \frac{1}{2}\int d^3r \int d^3r'\ \frac{e^2 n(\vec{r})n(\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|} + \int d^3r\ n(\vec{r})\frac{1}{2}\int d^3r'\ \frac{e^2 h_\lambda(\vec{r},\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}$$

so that

$$E_{xc} = \int d^3r\ n(\vec{r}) \underbrace{\int_0^1 d\lambda\ \frac{1}{2}\int d^3r'\ \frac{e^2 h_\lambda(\vec{r},\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}}_{=:\epsilon_{xc}} \tag{7.26}$$

The variable ϵ_{xc}, which itself is a functional of the density is the exchange-correlation energy per electron.

7.5.1 Screened interaction

It is possible[9] to reformulate the exchange-correlation energy by the exchange-only energy with a screened interaction. This is a form that reminds of the so-called **GW approximation**, where the self-energy is expressed by Green's-function times screened interaction.

$$\begin{aligned}E_{xc} &\overset{\text{Eq. 7.26}}{=} \frac{1}{2}\int d^3r \int d^3r'\ n(r)h_0(r,r') \underbrace{\int d\lambda \frac{h_\lambda(\vec{r},\vec{r'})}{h_0(\vec{r},\vec{r'})}\frac{e^2}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}}_{v_{scr}(\vec{r}-\vec{r'})} \\ &= \frac{1}{2}\sum_{i,j}\int d^3r \int d^3r'\ \phi_i^*(\vec{r})\phi_j(\vec{r})v_{scr}(\vec{r}-\vec{r'})\phi_j^*(\vec{r'})\phi_i(\vec{r'})\end{aligned}$$

where the screened interaction is defined as

$$v_{scr}(\vec{r}-\vec{r'}) \overset{\text{def}}{=} \int d\lambda\ \frac{h_\lambda(\vec{r},\vec{r'})}{h_0(\vec{r},\vec{r'})}\frac{e^2}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}$$

Note that, in contrast to the GW approximation, our expression is exact, but, on the other hand, it is limited to the ground state.

Note also that the screened interaction in this formulation only enters the exchange term, but not the Hartree term. It also includes the kinetic energy correction.

The qualitative behavior can be understood easily:

---

[9] This is an idea from the author, that has not been crosschecked properly. Therefore some caution is required.

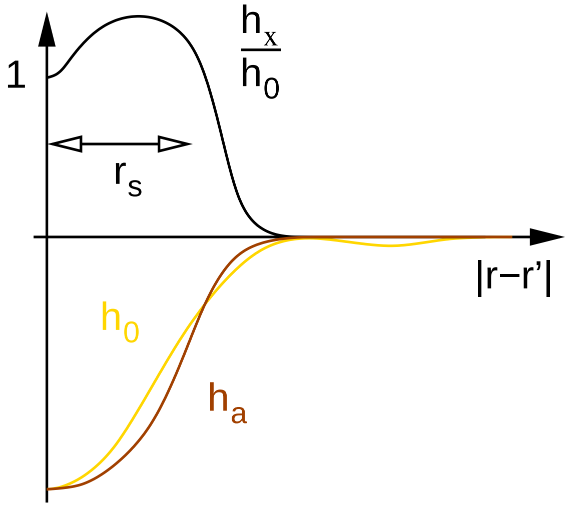


Fig. 7.7: Schematic diagram of the effect of the screening of the interaction. The golden line h_0 is the hole function of the non-interacting reference system, while the red line h_λ is the hole function in presence of a scaled interaction. The ratio scaled the Coulomb interaction in the exchange term. It suppresses the long-range tail of the interaction, and enhances the interaction for intermediate distances. r_s is the average electron radius $\frac{4\pi}{3}r_s^2 * n = 1$. ($h_a$ and h_x should be h_λ; Problem with the drawing program.)

- The hole function $h_\lambda(\vec{r}, \vec{r'})$ for a finite interaction is generally more compact than the on obtained without interaction. Thus, the screened interaction v_{scr} is generally more short ranged than the unscreened interaction.
- The screened interaction for $\vec{r} = \vec{r'}$ is identical to the unscreened interaction, because the hole function cancels exactly the spin density, independent of the strength of the interaction.
- Due to the compression of the hole upon increasing interaction and because of the sum-rule that the hole integrates to exactly one positive charge, the interaction must be enhanced for intermediate distances.
- The screening will be more effective, if the price for the distortion of the wave function is low, that is for a low density. This implies that the unscreened interaction is a reasonable approximation in the high-density region around the atomic center, while the tail region of a molecule is better described by a more short-ranged interaction.

7.5.2 Hybrid functionals

The adiabatic connection scheme is exploited in the construction of so-called **hybrid functionals**. These density functionals describe the exchange-correlation energy as a mixture of the correct exchange, obtained as in the Hartree-Fock method and another local or gradient corrected functional.

The performance of the hybrid functionals is probably among the best. They have the disadvantage that they require that the Hartree Fock exchange must be evaluated, which is computationally more challenging than the evaluation of the regular density functionals.

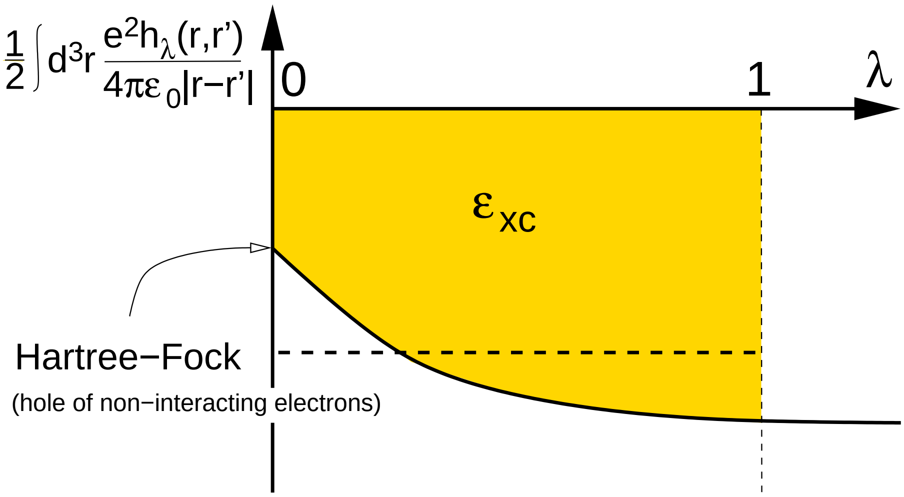
The idea behind the hybrid functionals is that the local functionals are appropriate if the interaction energy is very large. However if the interaction energy is small, the exchange hole is given by Hartree Fock exchange.

If for example the kinetic energy dominates, the shape of the exchange-correlation hole cannot adjust, and we can approximate $h_\lambda(\vec{r}, \vec{r'}) \approx h_0(\vec{r}, \vec{r'})$ by the hole of the non-interacting electron gas.

Thus, we obtain

$$\epsilon_{xc} \approx \frac{1}{2}\int d^3r' \; \frac{e^2 h_0(\vec{r},\vec{r'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}$$

If the electrostatic interaction dominates over the kinetic energy then $h_1(\vec{r},\vec{r'})$ would be a hard sphere, and therefore fairly local. Therefore, the hybrid functionals try to approximate the integral by a superposition of HF-exchange and a local density functional.



7.6 Functionals

According to the Bible, Jacob had a dream in which he saw a ladder descending from Heaven to Earth and angels climbing and descending the ladder[10]. In John Perdew's dream, the angels are users of DFT who climb the ladder to gain greater precision (at greater cost), but who also need to be able to descend the ladder depending upon their needs[**?**].

---

[10] Jacobs ladder has been described in the Book of Genesis (28:11âĂŞ19). It is a ladder from earth to heaven on which the angels descend and ascend.

JACOB'S LADDER OF DENSITY FUNCTIONALS

The different stairs of Jacobs Ladder of density functionals are

1. LDA (local density approximation): the exchange-correlation hole is taken from the free electron gas. These functionals exhibit strong overbinding. While the van-der Waals bond is not considered in these functionals, the overbinding mimicked van-der-Waals bonding, albeit for the wrong reason.
2. GGA (generalized gradient approximation): These functionals not only use the electron density but also its gradient to estimate the asymmetry of the exchange-correlation hole with respect to the reference electrons. As a result, surfaces are energetically favored compared to LDA and the overbinding is strongly reduced.
3. meta-GGA: in addition to the gradient also the kinetic energy density is used as a parameter. The kinetic energy is a measure for the flexibility of the electron gas.
4. Hybrid functionals: Hybrid functional include a fraction of exact exchange. This functionals improve the description of left-right correlations.
5. Exact: An exact density functional can be obtained using the constrained search formalism using many-particle techniques.

7.7 X_α method

Even before the invention of density-functional theory per se, the so-called X_α method has been introduced. Today, the X_α method has mostly historical value. The X_α method uses the expression for the exchange of a homogeneous electron gas instead of the exchange-correlation energy. However, the exchange energy has been scaled with a parameter, namely X_α, that has been adjusted to Hartree-Fock calculations. The results are shown in Fig. 7.8.

The rationale behind the X_α-method is a dimensional argument. Choose a given shape for the exchange-correlation hole, but scale it according to the density and the electron sum rule. Then the exchange-correlation energy per electron always scales like $n^{\frac{1}{3}}$. Each shape corresponds to a pre-factor.

Consider a given shape described by a function $f(\vec{r})$ with

$$\begin{aligned} f(\vec{0}) &= 1 \\ \int d^3r\; f(\vec{r}) &= 1 \end{aligned}$$

Now, we express the hole function by the function f by scaling its magnitude at the origin such that the amplitude of the hole cancels the electron density. Secondly, we stretch the function in space so that the sum rule, which says that the hole must integrate to -1, is fulfilled. These conditions yield the model for the exchange correlation hole.

$$h(\vec{r}_0, \vec{r}) = -n(\vec{r}_0) f\Big(\frac{\vec{r}-\vec{r}_0}{n(\vec{r}_0)^{\frac{1}{3}}}\Big)$$

The corresponding exchange-correlation energy per electron is

$$\epsilon_{xc}(\vec{r}) = -\frac{1}{2}\int d^3r'\; \frac{e^2}{4\pi\epsilon_0|\vec{r}-\vec{r'}|} f\Big(\frac{\vec{r}-\vec{r'}}{n(\vec{r})^{\frac{1}{3}}}\Big)$$

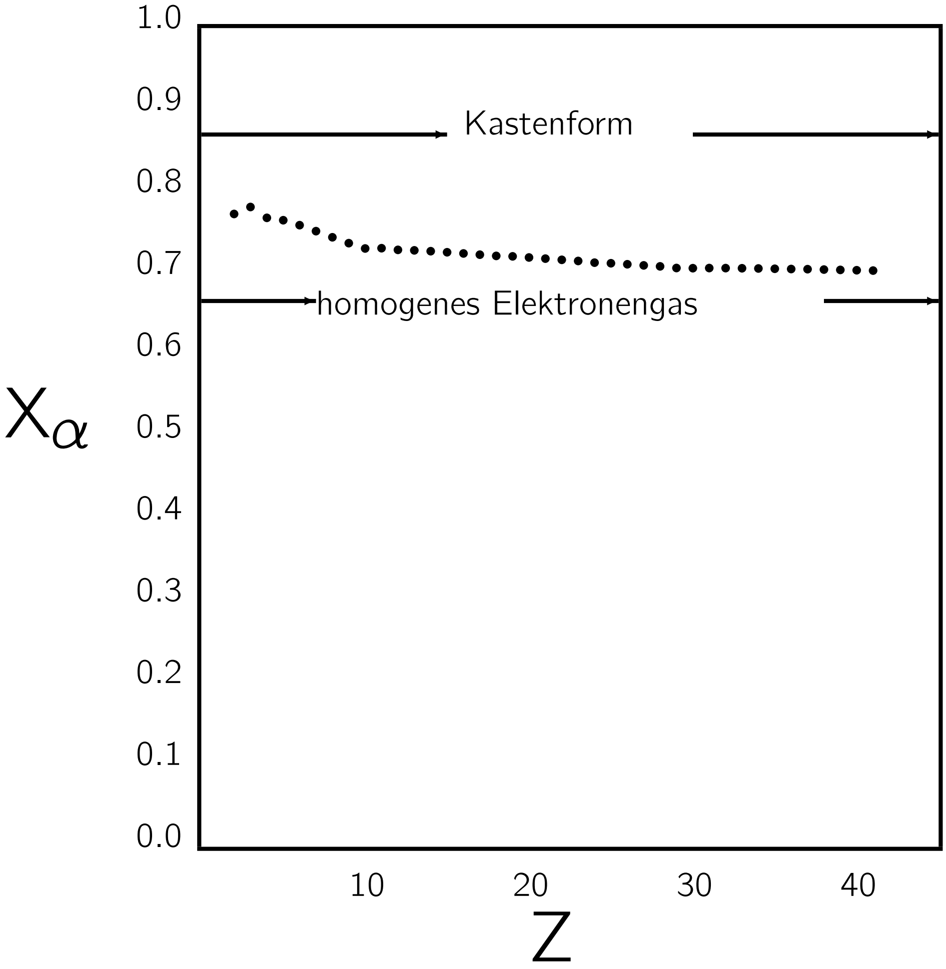


Fig. 7.8: X_α value obtained by comparison of the atomic energy with exact Hartree-Fock calculations as function of atomic number of the atom.[51, 52]

If we introduce a variable transform

$$\vec{y} = \frac{\vec{r} - \vec{r}'}{n(\vec{r})^{\frac{1}{3}}}$$

we obtain

$$\epsilon_{xc}(\vec{r}) = -\frac{n(\vec{r})^{\frac{1}{3}}}{2} \int d^3y' \, \frac{e^2}{4\pi\epsilon_0|\vec{y}|} f(\vec{y}) = -Cn^{\frac{1}{3}}$$

where C is a constant that is entirely defined by the shape function $f(\vec{r})$.

In general, the X_α method yields larger band gaps than density functional theory. The latter severely underestimates band gaps. This is in accord with the tendency of Hartree-Fock to overestimate band gaps. In contrast to Hartree-Fock, however, the X_α method is superior for the description of metals because it does not lead to a vanishing density of states at the Fermi level.

7.8 Local density functionals

The first true density functionals were constructed for the homogeneous electron gas, by extracting the exchange-correlation energy per electron ϵ_{xc} from a calculation of the interacting homogeneous electron gas.

A breakthrough came about when Ceperley and Alder[53] performed quantum Monte-Carlo calculations of the homogeneous electron gas as function of the density. Quantum Monte-Carlo calculations are computationally expensive, but provide the energy of an interacting electron gas in principle exactly, that is with the exceptions of numerical errors. These results have been parameterized by Perdew and Zunger[37] and combined with so-called RPA results for the high-density limit. **RPA** stands for **Random Phase Approximation**[54], which is accurate in the high density limit.

These density functionals exhibit very good results for solids. Electron densities are nearly perfect. Bond distances are typically underestimated by 1-3 % and bond angles agree with experiment within few degrees. Binding energies, on the other hand, are strongly overestimated. The errors are in the range of electron volts and thus comparable to bond energies. Thus, these functionals have been useless for studying chemical reactions. However, the results for solid state processes such as diffusion have been very good. This has lead to a long-standing misunderstanding between solid state physicists and chemists about the usefulness of density-functional theory. The density functionals worked find for solids, which was what the physicists are interested in, while they were a disaster for the binding energies of molecules, the major interest of chemists. This difficulty has been overcome by the gradient corrected density functionals discussed later.

The overestimate of the binding energies covers up the lack of Van der Waals interactions. For example, the binding energy and the bond distances of nobel gases are in very good agreement with experiment. However, the agreement is good for the wrong reason. The weak van der Waals interaction is not described properly in local density functionals, but the overbinding compensates for this fact.

For a long time it was surprising that the approximate density functionals work at all. They have been derived from a completely homogeneous electron gas and are applied to an electron density of real materials, which are far from homogeneous. Jones and Gunnarsson gave one explanation, namely that density functionals observe an important sum rule, namely that the exchange-correlation hole integrates to minus one electron charge. Fig. 7.9 compares the exchange hole of a free-electron gas with that of a nitrogen atom, demonstrating that despite their very different shape, their contributions to the exchange energy are similar.

7.9 Local spin-density approximation

In order to describe magnetic systems, or so-called open-shell molecules, one uses the local spin-density approximation, where the spin-dependent density $n(\vec{x}) = n(\vec{r}, \sigma)$ is used instead of the total density $n(\vec{r}) = \sum_\sigma n(\vec{x})$. As a result, we obtain one-particle wave functions with spin up and spin down character and one obtains two effective potentials, one for the spin-up electrons and another for the spin-down electrons.

The difference between the effective potentials acts like a magnetic field, even though its origin is purely electrostatic, namely exchange and correlation also called the **exchange interaction**.

7.10 Non-collinear local spin-density approximation

While the local spin-density formalism only allows one-particle wave function to have either purely spin-up or purely spin down character, the non-collinear formulation allows the wave functions to have a mixed spin-up and spin-down character. As a result, the magnetization can not only vary in magnitude but also in direction.

7.11 Generalized gradient functionals

For a long time one has hoped to obtain better functionals by including also a dependence on the gradient of the density. The first attempts started from a model with a weakly oscillating electron

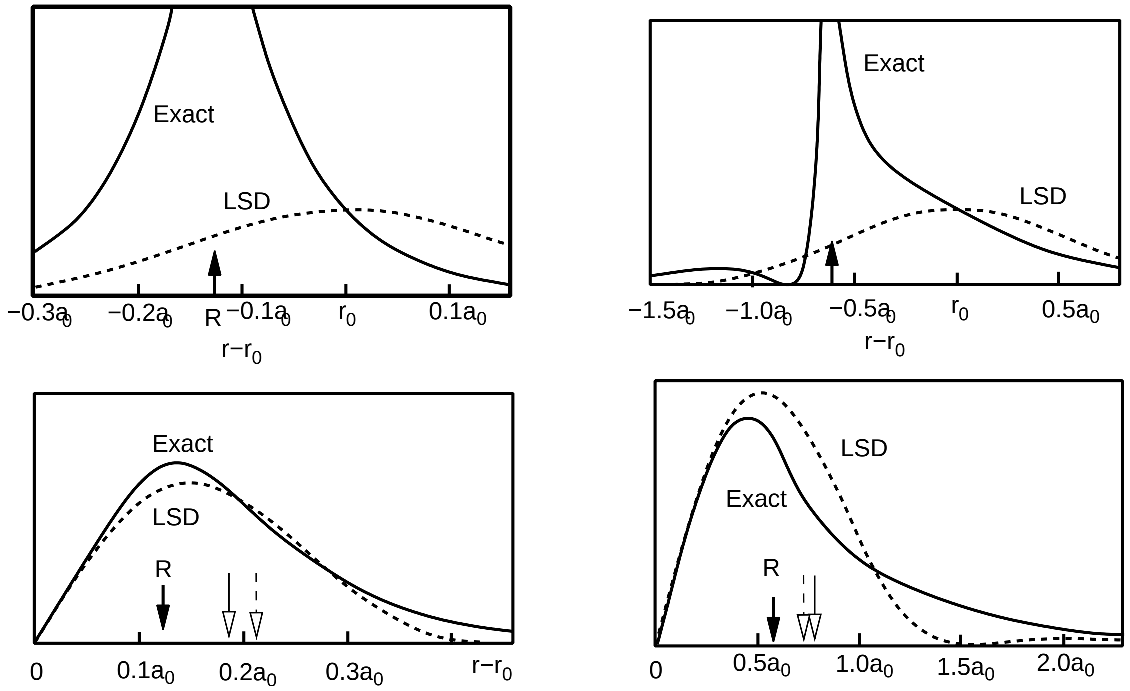


Fig. 7.9: exchange-correlation hole of an electron in the nitrogen atom (full line) and in the local density-functional approximation (dashed) [55]. In the left row, the reference electron is located 0.13 a_0 from the nucleus and on the right it is located further, at 0.63 a_0. Even though the shape of the DFT-exchange hole deviates grossly from the correct exchange hole, the spherical averages shown in the bottom figures are similar, which is attributed to the particle sum rule, obeyed in the local density approximation (LDA). The arrows pointing down indicate the exact and the DFT result of $\left[\int d^3r\, \frac{|h(\vec{r},\vec{r}_0)|}{|\vec{r}-\vec{r}_0|}\right]^{-1}$. Similar results are available for neon[37].

density. The resulting functionals however were worse than the local functionals.

The reason for this failure was that when the expansion for slowly varying densities are extrapolated to strongly oscillating densities, the density of the exchange-correlation hole was overcompensating the total density.

Later one realized that one should introduce a dimension-less scaled gradient defined as

$$s = \frac{r_s|\vec{\nabla} n|}{n}$$

The scaled gradient allows to switch off the gradient corrections in a physical sense, that is if the hole runs into the danger of producing a negative correlation function $g(\vec{x'},\vec{x})$.

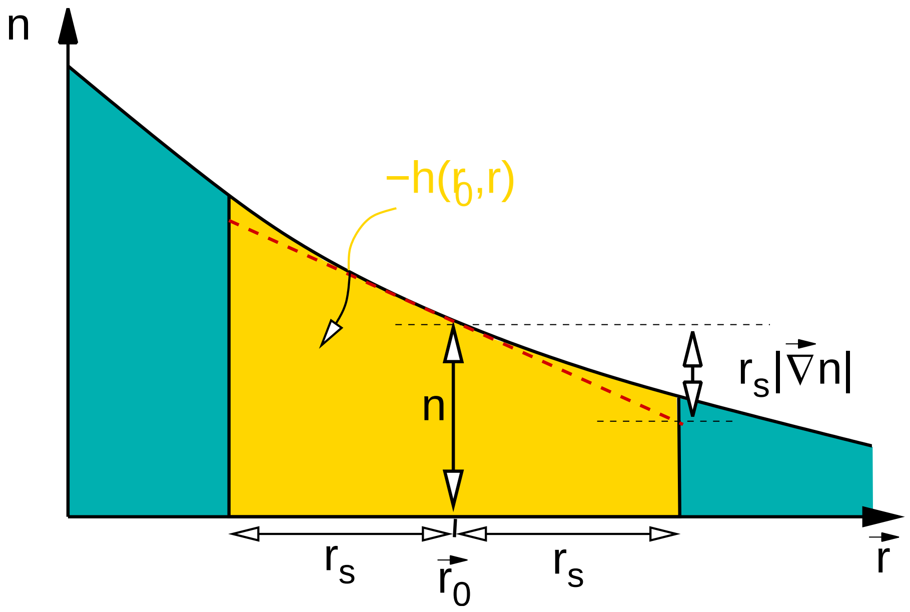
One way to construct the density is to use the results for slowly varying densities, which provides the parameter β.

$$\epsilon_{xc}(n,s) = \epsilon_{xc}^{hom}\Big[1+\beta(n)s^2\Big] \qquad \text{for} \qquad s << 1 \tag{7.27}$$

Surprisingly, the scaled gradient is largest in the tails of the wave functions, where the density falls off exponentially, that is

$$n(\vec{r}) \approx A\mathrm{e}^{-\lambda\vec{e}\vec{r}}$$

where $\vec{e}$ is the unit vector pointing into the direction in which the density falls off. Let us determine



the scaled gradient for this density

$$s=\sqrt[3]{\frac{3}{4\pi}}\,\frac{A\lambda\exp(-\lambda\vec{e}\vec{r})}{A^{\frac{4}{3}}\lambda\exp(-\frac{4}{3}\lambda\vec{e}\vec{r})}=\sqrt[3]{\frac{3}{4\pi}}\lambda n^{-\frac{1}{3}}\to\infty$$

If we use only the gradient correction with the small-s expansion, the exchange-correlation energy per particle clearly becomes infinite.

However, in the tail region we can make use of other information. Consider an electron that is far outside of a molecule. It is clear that it "sees" a molecule with $N-1$ electrons. Thus, the hole is entirely localized on the molecule and the center of the hole is far from the reference electron. If the hole is at the molecule and the reference electron is far from it, we can estimate its interaction by

$$\epsilon_{xc}=\frac{1}{2r} \tag{7.28}$$

where r is the distance of the electron from the molecule.

Let me demonstrate the construction of such a gradient corrected functional using the example of Becke's gradient correction for exchange[56]. We make an ansatz for the exchange energy per electron as

$$\epsilon_{xc}=Cn^{\frac{1}{3}}F(s) \tag{7.29}$$

where C is the pre-factor for the exchange energy. The function $F(s)$ is determined such that small gradient expansion, Eq. 7.27, is reproduced, and that the exchange energy per electron in the tail region is correct.

To consider the tail region we require

$$\epsilon_{xc}\overset{Eq.\ 7.29}{=}Cn^{\frac{1}{3}}F(s)\overset{Eq.\ 7.28}{=}\frac{1}{2r}\qquad\Rightarrow\qquad F(s)=\frac{1}{2Crn^{\frac{1}{3}}}\qquad\text{for}\quad s\to\infty$$

First we express the radius by the density and insert the result in Eq. 7.30 to obtain an expression for $F(n(s))$.

$$n(r)=Ae^{-\lambda r}\qquad\Rightarrow\qquad r(n)=-\frac{1}{\lambda}\ln\left[\frac{n}{A}\right]$$

$$\Rightarrow F(s)=\frac{1}{2Crn^{\frac{1}{3}}}=\frac{-\lambda}{2Cn^{\frac{1}{3}}\left(\ln[n]-\ln[A]\right)}$$

Next we express the density by the scaled gradient and insert the result in the above equation to obtain an expression for $F(s)$.

$$
\begin{aligned}
s(n) &= \sqrt[3]{\frac{3}{4\pi}}\lambda n^{-\frac{1}{3}} \qquad \Rightarrow \qquad n(s) = \frac{3\lambda^3}{4\pi}s^{-3} \\
\Rightarrow F(s) &= \frac{-\lambda}{2Cn^{\frac{1}{3}}\left(\ln[n]-\ln[A]\right)} \\
&= \frac{-\lambda}{2C\sqrt[3]{\frac{3\lambda^3}{4\pi}}\frac{1}{s}\left(\ln[\frac{3\lambda^3}{4\pi}s^{-3}]-\ln[A]\right)} \\
&= \frac{-s}{2C\sqrt[3]{\frac{3}{4\pi}}\left(\ln[\frac{3\lambda^3}{4\pi A^3}s^{-3}]-3\ln[s]\right)} \\
&\overset{s\to\infty}{\approx} \frac{1}{6C}\sqrt[3]{\frac{4\pi}{3}}\,\frac{s^2}{s\ln[s]}
\end{aligned}
$$

This equation gives us the large gradient limit of $F(s)$.

In order to connect the low gradient limit with the large gradient limit we choose

$$
F(s) = 1 + \frac{\beta s^2}{1 + 6C\beta s\sinh^{-1}(s)}
$$

The result is shown in Fig. 7.10.

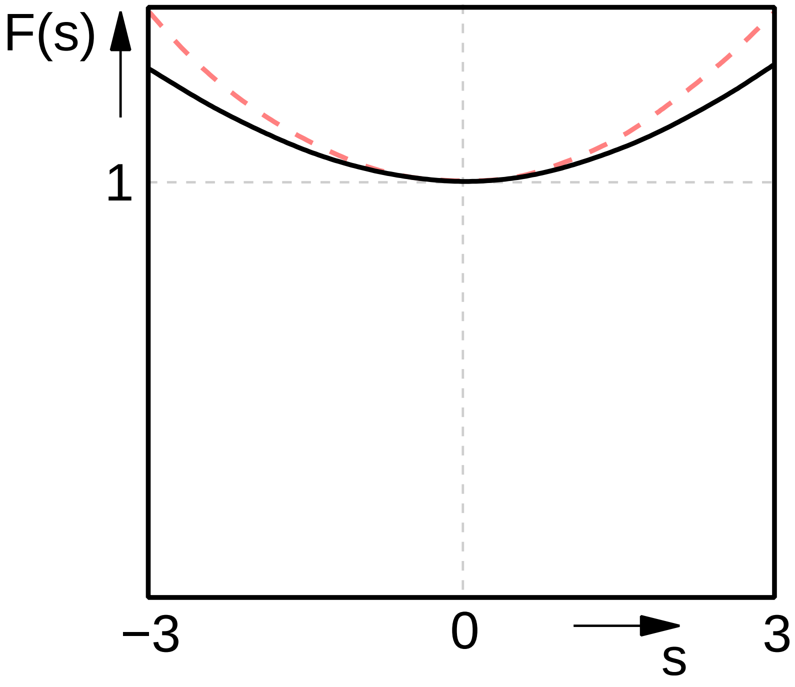


Fig. 7.10: The function $F(s)$ from Becke's gradient corrected functional [56] for the exchange energy and from the small gradient expansion (dashed). Note that the gradient correction of Becke is smaller than the expansion for large gradients.

The gradient correction plays the role of a surface energy, as it contributes mostly in the tail region. While the exchange-correlation energy per electron falls off exponentially in the local functionals, it falls off as $\frac{1}{2r}$ in the gradient corrected functionals. This effect lowers the energy in the tail region of a molecule compared to the local functionals. If we break a bond, the surface area of a molecule increases, because the bond is transformed into a tail region. Thus, the gradient correction favors the dissociation of the bonds. As a consequence, gradient corrected functionals avoid the artificial overbinding of the local functionals. This argument also explains that local functionals perform fairly well in solids: If an atom diffuses, there is no additional surface created so that the gradient correction is minor.

Up to know there is an entire suite of different gradient corrected functionals. They are called **Generalized Gradient Approximations (GGA)** to differentiate them from the gradient expansions. The most common functionals are the Perdew-Wang-91-GGA, which has been superseded by the simpler Perdew-Burke-Ernzerhof functional. Both yield nearly identical results.

As shown in Fig. 7.11, the performance of gradient corrected functionals is extremely good also for binding energies, which were unsatisfactory in the local density functionals. This has drawn also the chemists into the field of density-functional calculations.

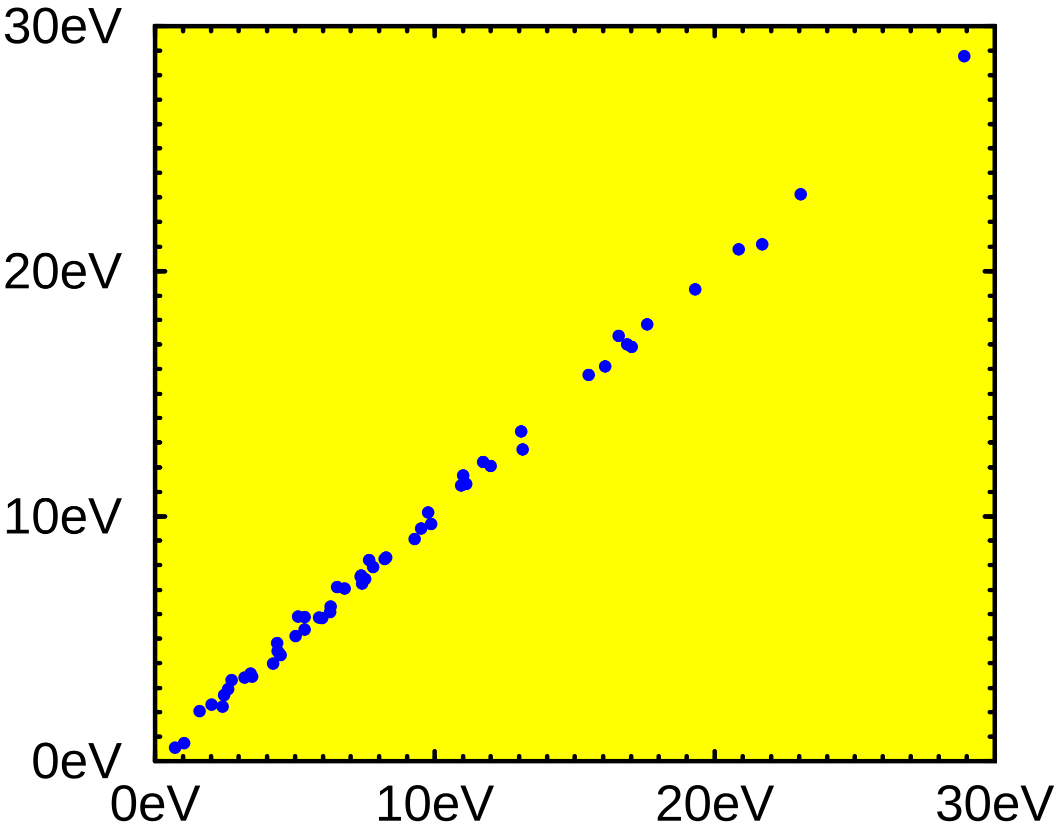


Fig. 7.11: Accuracy of the atomization energies calculated with the Becke88-Perdew86 density functional compared to experiment.

7.12 Additional material

7.12.1 Relevance of the highest occupied Kohn-Sham orbital

According to Perdew[57**?**], the energy of the highest occupied Kohn Sham orbital is the ionization potential of the material.

This statement has been disputed by Kleinman[**?**].

7.12.2 Correlation inequality and lower bound of the exact density functional

7.13 Functionals

- **[X-α]** Uses a local expression for exchange as derived from the free electron gas, and scaled with a factor α.
- **[Barth-Hedin]**
- **[PZ (Perdew-Zunger)]** parameterization of Ceperley and Alder's quantum Monte Carlo data of the free electron gas.
- **[BP86 (Becke-Perdew)]**
- **[PW91 (Perdew-Wang)]** First completely parameter free version of a gradient corrected density functional
- **[PBE (Perdew-Burke-Ernzerhof)]**
- **[B3LYP (Becke-Lee-Yang-Parr)]** Hybrid functional.
- **[GGA] (Generalized gradient approximation)** General expression for gradient corrected functionals.
- **[Hybrid functionals]** General expression functionals employing a mixture of Hartree-Fock and a density functional

Fertig: WS06/07 11 Doppelstunde 13.Dec.06

7.14 Reliability of DFT

If we knew the exact density functional, we would be able to determine exactly

- the total energy
- the electron density and
- the energy of the highest occupied orbital

The argument that DFT predicts the highest occupied orbital, while the orbital energies of all other one-particle orbitals are, in principle, without physical meaning goes as follows: Far from the surface of a crystal the density of a one-particle orbital falls off as $e^{-\lambda z}$ where λ is related to the orbital energy by $\lambda = \frac{1}{\hbar}\sqrt{-2m_e\epsilon}$, where the orbital energy is measured relative to the vacuum level, that is $\lim_{z\rightarrow\infty} v_{eff}(\vec{r})$. For large distances the highest occupied orbital, which has the slowest decay, will dominate the electron density. Thus, if we would know the electron density very accurately far from a surface, we would be able to determine the dominant exponential decay constant and thus the energy of the highest occupied orbital.[11] Thus, the energy of the highest occupied orbital can be determined from the density, which is an exact prediction of DFT.

Of course, even though exact DFT predicts these quantities accurately, it remains to be seen in each case if that also holds for the approximate functionals used in practice.

- Bond energies: Local functionals overestimate binding by often more than 1 eV. Energies are fairly good as long as the effective surface area of the system is not changed. This poor result is dramatically improved by the GGA's.
- Structures: bond lengths are underestimated in LDA by 0-2 %. They are overestimated by GGA's by 0-2 %.
- Dipole moments standard deviation 0.1 Debye[58]

7.15 Deficiencies of DFT

- **Missing Van der Waals interaction:** The van-der-Waals interaction is not included. Surprisingly the description of van der Waals bonds is nearly perfect with truly local density functionals. This however is an artifact of the tendency to overbinding in truly local density functionals on the one hand, and the lack of van der Waals interactions on the other. Since the overbinding of covalent bonds introduces errors in the range of 1 eV, it is generally not a good idea to use truly local density functionals instead of their gradient corrected counterparts.

  A density functional for the van der Waals interaction has been developed by Dion et al.[59]. It is named vdWDF (van-der-Waals density functional).
- **Band-gap Problem:** If we compare the spectrum of Kohn Sham states with optical absorption spectra we observe that the band gaps are too small. This is not a deficiency of density-functional theory, but an over-interpretation of the Kohn Sham spectrum. The excited states should not be calculated from the same potential as the ground state.
- **Reduced band width:** Band width of alkaline metals.
- **Mott insulators**
- **unstable negative ions**
- **Broken symmetry spin states:** singlet O_2 Molecule
- **Transition states:**
- **High-energy spectrum:** The high energy part if the Kohn-Sham spectrum is too low compared to the measured optical spectrum. The reason is that the exchange-correlation hole cannot follow the very fast electrons.????

---

[11] There is a caveat in the argument. Consider an insulator which may have surface states in the band gap. In that case the density would probe the highest occupied surface state, which may lie above the highest occupied bulk state.

Part II

Advanced solid state physics

Chapter 8

Second quantization

"Many workers, particularly those trained in field theory, advocate the language of second quantization. The author does not favor this method, and it will not be used in this article. He wishes to point out that nothing is accomplished by the use of the second quantization formalism except to express in a more compact language results which we meet in an equivalent way in the determinantal method/ Furthermore, he feels that the language of second quantization is likely to hide essential features such as the fundamentally nonlinear nature of the total energy, when expressed in terms of the occupation numbers, which are of primary importance, and which can be easily overlooked by one whose only training is in the field theory."

John C. Slater in "Statistical Exchange-correlation Correction in the self-consistent field".

8.1 Fock space

There are several ways towards second quantization. The most simple one ist to define the Fock space. The **Fock space** is the combination of the Hilbert spaces for 0-particle states, 1-particle states, 2-particle states, etc. If $\mathcal{H}_N$ is the Hilbert space of N-particle states the Fock space $\mathcal{F}$ is the unification of the Hilbert spaces for all integer N

$$\mathcal{F} = \mathcal{H}_0 \cup \mathcal{H}_1 \cup \mathcal{H}_2 \cup \ldots = \bigcup_{N=0}^{\infty} \mathcal{H}_N$$

What is missing is a definition of the scalar product between states with different particle numbers. We postulate that the scalar product of states with different particle numbers vanishes.[1]

As a curiosity, we introduced here also the 0-particle state, called **vacuum state**, that is not represented by a wave function at all.

8.2 Number representation

We are entering a new level of abstraction and it turned out to be useful to introduce a special notation for it. This will lead us to the number representation and to creation and annihilation operators.

In order to work within the Fock space we need to select a suitable basis set. We have already learned that Slater determinants resulting a complete orthonormal one-particle basis, form a complete set of many particle wave functions.

---

[1] This postulate is non-trivial and can be derived more naturally from the second quantization via introducing wave functions in the space of the fields describing the wave functions.

Each slater determinant is uniquely defined, up to a sign, by the one-particle states used in its construction. Therefore, the Slater determinants can be written economically in the **number representation**. For a given, ordered, one-particle basisset, we form a vector, where each position corresponds to a particular one-particle orbital. Each entry can be zero or one, and describes if that particular orbital is present in the Slater determinant or not. If the number is one, we say that the orbital is occupied, and if it is not, we say that it is unoccupied. Thus, a Slater determinant can be expressed in a form like

$$|0,0,0,\underbrace{1}_{i_1},0,\underbrace{1}_{i_2},0,\dots\rangle = |S_{i_1,i_2,\dots}\rangle \qquad \text{with} \quad i_1 < i_2 < \dots$$

where $S_{i_1,i_2,\dots}$ is a Slater determinant made of the one-particle orbitals $|\phi_{i_1}\rangle, |\phi_{i_2}\rangle, \dots$.

Because the Slater determinant changes its sign under permutation of two one-particle wave functions, we need to define the sign convention. The sign is fixed such that the one-particle orbitals must occur in the corresponding Slater determinant in the order of increasing index, that is $i_1 < i_2, \dots$. Thus

$$\begin{aligned} |1,1,0,0,\dots\rangle &= |S_{1,2}\rangle = -|S_{2,1}\rangle \\ \langle\vec{r},\vec{r}'|1,1,0,0,\dots\rangle &= \frac{1}{\sqrt{2!}}\left[\phi_1(\vec{r})\phi_2(\vec{r}') - \phi_2(\vec{r})\phi_1(\vec{r}')\right] \end{aligned}$$

Note also that there is an explicit **vacuum state**

$$|\mathcal{O}\rangle = |0,0,0,0,0,0,0,0,0,0,\dots\rangle$$

The vacuum state describes a system without particles. Note that the vacuum state, which we denote by a calligraphic "O", $|\mathcal{O}\rangle$, is different from the zero state $|0\rangle$. The latter is obtained by multiplying an arbitrary state with zero.

The most general antisymmetric wave function has the form

$$|\Phi\rangle = \sum_{n_1,n_2,\dots=0}^{1} |n_1,n_2,\dots\rangle c_{n_1,n_2,\dots}$$

Let us consider as one example a hydrogen molecule with only one s-orbital per site. The s-orbitals are orthogonalized so that we obtain a bonding and an antibonding orbital. Now we can form the limited Fock space for this model. The ordered one-particle basis is $|b,\uparrow\rangle, |b,\downarrow\rangle, |a,\uparrow\rangle, |a,\downarrow\rangle$, where $|b,\sigma\rangle$ describes the bonding and $|a,\sigma\rangle$ describes the antibonding orbital. There are $2^4 = 16$ Slater determinants as shown in Fig. 8.1.

- $|0000\rangle$ is the zero-particle state. It describes just the two protons without electrons.
- $|1000\rangle$, $|0100\rangle$, $|0010\rangle$ and $|0001\rangle$ are the one-particle states. Thus, they describe the H_2^+ ion. The states $|1000\rangle$ and $|0100\rangle$ correspond to the bound H_2^+ ion, while the remaining two states correspond to the electron in the antibonding state. The molecule in the latter configuration would be unstable.
- $|1100\rangle$ corresponds to the ground state of the H_2 molecule with two paired electrons in the bonding orbital. There are in total 6 two-electron slater determinants of which 5 are excited states.
- There are four three-electron states, namely $|1110\rangle$ $|1101\rangle$, $|1011\rangle$,$|01111\rangle$. The three particle states describe the negative H_2^- ion.
- finally there is one four-electron state, $|1111\rangle$ which describes the H_2^{2-} ion.

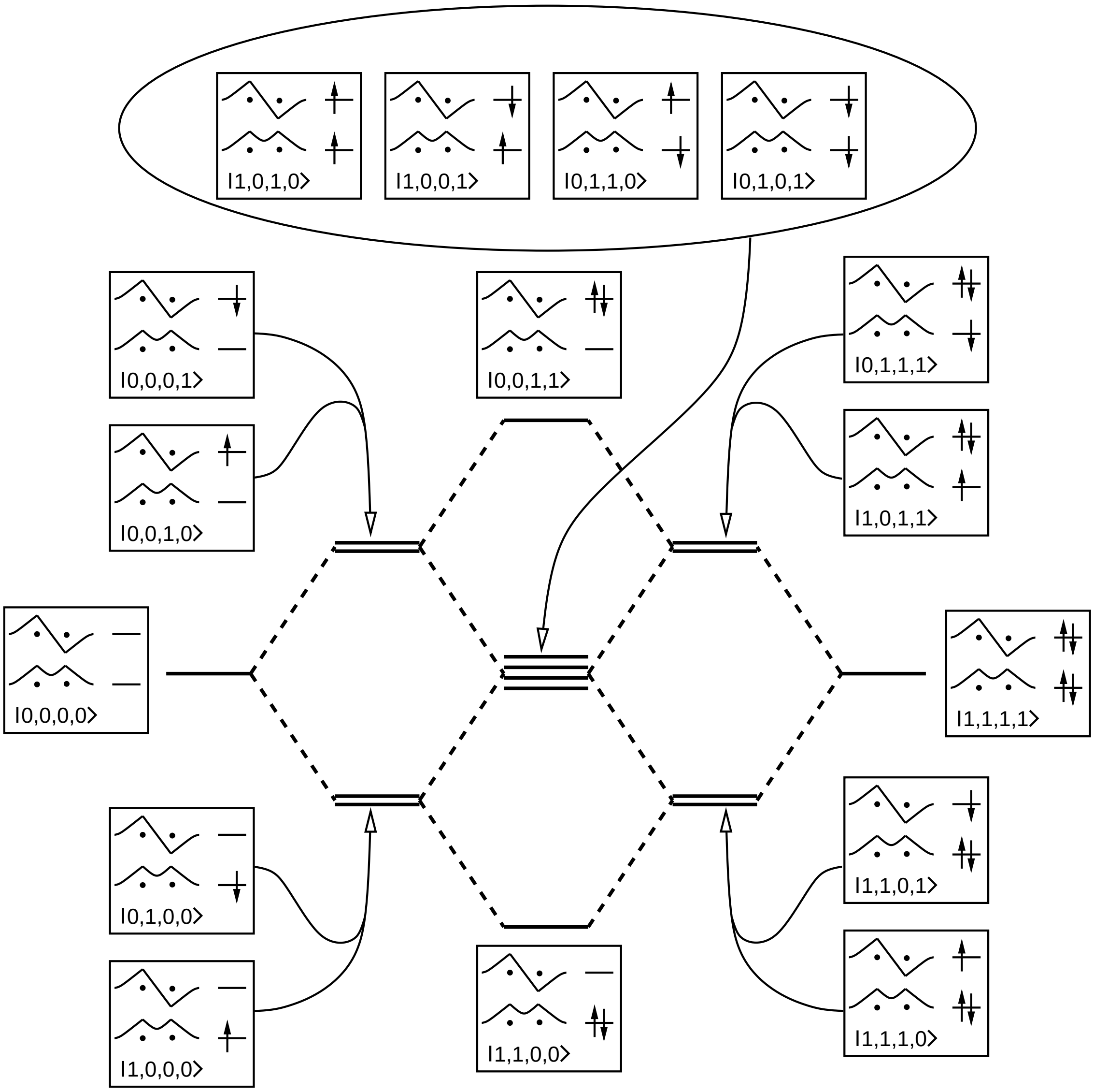


Fig. 8.1: Schematic representation of the Fock space build from the bonding and antibonding states of the hydrogen molecule. The middle diagram represents the many-particle energy levels assuming noninteracting electrons and with $\epsilon_a = -\epsilon_b$, where ϵ_a is the one-particle energy of the antibonding level and ϵ_b is the energy level of the bonding orbital. The insets correspond to the corresponding number representation and the corresponding one-electron energy level diagram. On the left it contains a schematic drawing of the one-particle wave functions.

8.2.1 Excursion: Integer representation representation of Slater determinants

Note that the sequence can be considered as a binary number with the digits. This allows to identify each Slater determinant uniquely with one integer number.

The conversion of an integer into the number representation is to

$$\begin{aligned}
0 &\rightarrow |0,0,0,0,0,0,0,0,0,0,\ldots\rangle\\
1 &\rightarrow |1,0,0,0,0,0,0,0,0,0,\ldots\rangle\\
2 &\rightarrow |0,1,0,0,0,0,0,0,0,0,\ldots\rangle\\
3 &\rightarrow |1,1,0,0,0,0,0,0,0,0,\ldots\rangle\\
4 &\rightarrow |0,0,1,0,0,0,0,0,0,0,\ldots\rangle\\
\sum_i n_i 2^{i-1} &\rightarrow |n_1,n_2,\ldots\rangle
\end{aligned}$$

In Fortran, bits can be handled in a simple manner. If I is the integer representation of a Slater determinant and i is the position of a particular one-particle orbital, we obtain its occupation number by

$$n_i = \mathtt{BTEST}(I-1, i-1)$$

The bits can be manipulated using the Fortran functions

$$\begin{aligned}
I' &= \mathtt{IBCLR}(I-1, i-1)\\
I'' &= \mathtt{IBSET}(I-1, i-1)
\end{aligned}$$

where I' is obtained from I by setting the occupation n_i to zero, and I'' is obtained by setting it to one.

In order to see the difficulty of many-particle physics let us count the number of states. If truncate the one-particle basisset to M orbitals, there are 2^M many-particle basisfunctions, because any one-particle state can be occupied or unoccupied. For $M = 10$ there are already $2^{10} = 1024$ states, which requires a kilobyte of memory to hold the coefficients. For $M = 20$ we already require a Megabyte and for 30 orbitals we need a Gigabyte.

8.3 Creation and annihilation operators

We define creation operators which should increase the particle number in a given one-particle state. This will allow to create all many-particle states from the vacuum state.

We define the creation operator for a particular orbital as follows: When we apply a creation operator to a Slater determinant, it adds a row with the new orbital at the top of the corresponding Matrix, and introduces a column with a new coordinate to the right. It is important to watch the sign of the state. Let us consider an example where a third particle in state $|\phi_2\rangle$ is created in a state that contains two particles, one in $|\phi_1\rangle$ and one in $|\phi_3\rangle$.

$$\begin{aligned}
\overbrace{\det\begin{vmatrix}\phi_1(\vec{x}_1) & \phi_1(\vec{x}_2)\\ \phi_3(\vec{x}_1) & \phi_3(\vec{x}_2)\end{vmatrix}}^{|1,0,1,0,\ldots\rangle}
&\xrightarrow{\hat{c}_2^\dagger}
\overbrace{\det\begin{vmatrix}\phi_2(\vec{x}_1) & \phi_2(\vec{x}_2) & \phi_2(\vec{x}_3)\\ \phi_1(\vec{x}_1) & \phi_1(\vec{x}_2) & \phi_1(\vec{x}_3)\\ \phi_3(\vec{x}_1) & \phi_3(\vec{x}_2) & \phi_3(\vec{x}_3)\end{vmatrix}}^{\hat{c}_2^\dagger|1,0,1,0,\ldots\rangle}\\
&= (-1)\underbrace{\det\begin{vmatrix}\phi_1(\vec{x}_1) & \phi_1(\vec{x}_2) & \phi_1(\vec{x}_3)\\ \phi_2(\vec{x}_1) & \phi_2(\vec{x}_2) & \phi_2(\vec{x}_3)\\ \phi_3(\vec{x}_1) & \phi_3(\vec{x}_2) & \phi_3(\vec{x}_3)\end{vmatrix}}_{|1,1,1,0,\ldots\rangle}
\end{aligned}$$

That is

$$\hat{c}_2^\dagger|1,0,1,0,0,\ldots\rangle = -|1,1,1,0,0,\ldots\rangle$$

We define the creation and annihilation operators by

$$\begin{aligned}
\hat{c}_i^\dagger|n_1, n_2, \ldots, \underbrace{0}_{n_i}, \ldots\rangle &= |n_1, n_2, \ldots, \underbrace{1}_{n_i}, \ldots\rangle\ \left[(-1)^{\sum_{j<i} n_j}\right] \\
\hat{c}_i^\dagger|n_1, n_2, \ldots, \underbrace{1}_{n_i} \ldots\rangle &= 0 \\
\hat{c}_i|n_1, n_2, \ldots, \underbrace{1}_{n_i}, \ldots\rangle &= |n_1, n_2, \ldots, \underbrace{0}_{i} \ldots\rangle\ (-1)^{\sum_{j<i} n_j} \\
\hat{c}_i|n_1, n_2, \ldots, \underbrace{0}_{n_i} \ldots\rangle &= 0
\end{aligned}$$

In short, we may write

FERMIONIC CREATION AND ANNIHILATION OPERATORS

$$\begin{aligned}
\hat{c}_i^\dagger|n_1, n_2, \ldots, n_i, \ldots\rangle &= |n_1, n_2, \ldots, n_i + 1, \ldots\rangle\ \left[(-1)^{\sum_{j<i} n_j}\right](1 - n_i) \\
\hat{c}_i|n_1, n_2, \ldots, n_i, \ldots\rangle &= |n_1, n_2, \ldots, n_i - 1 \ldots\rangle\ \left[(-1)^{\sum_{j<i} n_j}\right] n_i
\end{aligned} \tag{8.1}$$

Note that the state with $n_i + 1$ does not exist for $n_i = 1$. This does not matter because the state is multiplied with zero. A conceptionally elegant solution to this problem is to define a state with an occupation number other than zero or one as the zero state, that is.

$$|n_1, n_2, \ldots\rangle \overset{\text{def}}{=} |0\rangle \qquad \text{if there is an} \qquad n_i \notin \{0, 1\}$$

The matrix elements can be written as

$$\begin{aligned}
\begin{pmatrix} \langle \ldots 0 \ldots |\hat{c}_i^\dagger| \ldots 0 \ldots\rangle & \langle \ldots 0 \ldots |\hat{c}_i^\dagger| \ldots 1 \ldots\rangle \\ \langle \ldots 1 \ldots |\hat{c}_i^\dagger| \ldots 0 \ldots\rangle & \langle \ldots 1 \ldots |\hat{c}_i^\dagger| \ldots 1 \ldots\rangle \end{pmatrix} &= \left[(-1)^{\sum_{j<i} n_j}\right] \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \\
\begin{pmatrix} \langle \ldots 0 \ldots |\hat{c}_i| \ldots 0 \ldots\rangle & \langle \ldots 0 \ldots |\hat{c}_i| \ldots 1 \ldots\rangle \\ \langle \ldots 1 \ldots |\hat{c}_i| \ldots 0 \ldots\rangle & \langle \ldots 1 \ldots |\hat{c}_i| \ldots 1 \ldots\rangle \end{pmatrix} &= \left[(-1)^{\sum_{j<i} n_j}\right] \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}
\end{aligned}$$

In this equation, we used a notation where we dropped all occupation numbers except n_i.

From this equation it is immediately clear that $\hat{c}$ and $\hat{c}^\dagger$ are Hermitian conjugate to each other.

This immediately leads us to one other representation of the basis states, apart from the number representation and the Slater determinants. Each state can be identified with a product of creation operators. We start out from the **vacuum state** $|\mathcal{O}\rangle$. The vacuum state represents the vacuum, that is a system without particles.[2] Now we use the creation operators to fill in electrons.

The one-particle states are obtained from one creator

$$\hat{c}_{i_1}^\dagger|\mathcal{O}\rangle = |0, \ldots, 0, \underbrace{1}_{pos.i_1}, 0, \ldots\rangle$$

---

[2] Note, that the vacuum state differs from a zero state.

and the two-particle states are obtained form two creators

$$
\begin{aligned}
\hat{c}^\dagger_{i_2}\hat{c}^\dagger_{i_1}|\mathcal{O}\rangle &= \hat{c}^\dagger_{i_2}|\dots 0,\underbrace{1}_{pos.i_1},0\dots\rangle \\
&= \begin{cases}
+|\dots,0,\underbrace{1}_{pos.i_2},0,\dots,0,\underbrace{1}_{pos.i_1},0,\dots\rangle & \text{for } i_2 < i_1 \\
-|\dots,0,\underbrace{1}_{pos.i_1},0,\dots,0,\underbrace{1}_{pos.i_2},0,\dots\rangle & \text{for } i_2 > i_1 \\
0 & \text{for } i_i = i_2
\end{cases}
\end{aligned}
$$

For $i_2 < i_1$ the new one-particle orbital does not have to be commuted with the other orbital in order to bring the orbitals in the Slater determinant into the correct order. For $i_2 > i_1$, on the other hand there is exactly one permutation, which leads to the negative sign.

We observe that the states so created only differ by a sign from the states in the number representation. We can get rid of the sign by ordering the creators such that indices of the creation operators increase from left to right. In this case the creator for the highest orbital does not need to permuted with another one, because it is applied first. The second one does not undergo a sign change because there is no particle yet with a lower index. The argument just goes on. Thus

NUMBER REPRESENTATION AND CREATION OPERATORS

$$
\hat{c}^\dagger_{i_1}\hat{c}^\dagger_{i_2}\dots\hat{c}^\dagger_{i_N}|\mathcal{O}\rangle = |\dots,0,\underbrace{1}_{pos.i_1},0,\dots,0,\underbrace{1}_{pos.i_2},0,\dots,0,\underbrace{1}_{pos.i_N},0,\dots\rangle
$$

with $i_1 < i_2 < i_3,\dots,i_N$.

8.4 Commutator and anticommutator relations

In the following we determine the anticommutator relation between creation and annihilation operators, which will be essential for practical calculations.

The rules, we will arrive at are

ANTICOMMUTATION RULES FOR FERMIONIC CREATION AND ANNIHILATION OPERATORS

$$
\begin{aligned}
[\hat{c}^\dagger_i,\hat{c}^\dagger_j]_+ &= \hat{c}^\dagger_i\hat{c}^\dagger_j + \hat{c}^\dagger_j\hat{c}^\dagger_i = 0 \\
[\hat{c}_i,\hat{c}_j]_+ &= \hat{c}_i\hat{c}_j + \hat{c}_j\hat{c}_i = 0 \\
[\hat{c}^\dagger_i,\hat{c}_j]_+ &= \hat{c}^\dagger_i\hat{c}_j + \hat{c}_j\hat{c}^\dagger_i = \delta_{i,j}
\end{aligned}
\tag{8.2}
$$

Note that only the anticommutators are numbers, while the commutators are still operators, whose result depends on the state on which they are acting.

The reader will remember that the commutator relations

$$
\begin{aligned}
[\hat{x}_i,\hat{x}_j]_- &= 0 \\
[\hat{p}_i,\hat{p}_j]_- &= 0 \\
[\hat{p}_i,\hat{x}_j]_- &= \frac{\hbar}{i}\delta_{i,j}
\end{aligned}
$$

, between positions and momenta have been central to quantum mechanics. In the second quantization of Fermions the anticommutator relations above play a similarly important role.

If we consider the harmonic oscillator we found the commutator relations

$$
\begin{aligned}
[\hat{a}_i^\dagger, \hat{a}_j^\dagger]_- &= 0 \\
[\hat{a}_i, \hat{a}_j]_- &= 0 \\
[\hat{a}_i^\dagger, \hat{a}_j]_- &= \delta_{i,j}
\end{aligned}
$$

where the index refers to certain vibrational mode of a multidimensional harmonic oscillator. The resulting energy spectrum was equidistant and only bounded from below. In the second quantization of Bosons, we say that the creation operator $a_i^\dagger$ creates a particle in a certain vibrational mode. We can construct any number of excitations, or particles, in a given vibration. The quantization of Bosons follows naturally when one applies the correspondence principle to the wave function itself. Since the Lagrangian and the Hamiltonian of a one-particle system is quadratic in the wave function, it is natural that the quantization of a field leads to a harmonic oscillator. The difference between Fermions and Bosons shows up here that the creation and annihilation operators obey an anticommutator relation in one case and a commutator relation in the other.

Similarly to the normal, or Bosonic, harmonic oscillator one can also define a Fermionic harmonic oscillator by analogy. The Fermionic harmonic oscillator is described in ΦSX: Quantum Physics.

8.4.1 Anticommutator relations with the same one-particle orbital

We obtain with Eq. 8.1

$$
\begin{aligned}
\hat{c}_i^\dagger \hat{c}_i^\dagger |\ldots, n_i, \ldots\rangle &= 0 \\
\hat{c}_i \hat{c}_i |\ldots, n_i, \ldots\rangle &= 0 \\
\hat{c}_i^\dagger \hat{c}_i |\ldots, n_i, \ldots\rangle &= |\ldots, n_i, \ldots\rangle n_i \\
\hat{c}_i \hat{c}_i^\dagger |\ldots, n_i, \ldots\rangle &= |\ldots, n_i, \ldots\rangle (1 - n_i)
\end{aligned}
$$

Thus, we find for the pairs of creation and annihilation operators acting on the same one-particle state

$$
\begin{aligned}
[\hat{c}_i, \hat{c}_i]_+ &= 0 \\
[\hat{c}_i^\dagger, \hat{c}_i]_+ &= 1 \\
[\hat{c}_i, \hat{c}_i^\dagger]_+ &= 1 \\
[\hat{c}_i\dagger, \hat{c}_i\dagger]_+ &= 0
\end{aligned}
$$

which is consistent with Eq. 8.2 for $i = j$.

As a byproduct, we can also find

$$
\hat{c}_i^\dagger \hat{c}_i |\ldots, n_i, \ldots\rangle = n_i |\ldots, n_i, \ldots\rangle
$$

This shows that the Slater determinants are eigenstates of the operators $\hat{c}_i^\dagger \hat{c}_i$.

8.4.2 Anticommutator relations with different one-particle orbitals

For $i \neq j$ we obtain[3]

---

[3] The expressions are written such that the particle with the higher index is created or destroyed first, that is $j > i$. If the order is different, the number of states with index smaller than i has changed by one, which is taken care of by the sgn function.

with Eq. 8.1

$$
\begin{aligned}
\hat{c}_i^\dagger\hat{c}_j^\dagger|\ldots,n_i,\ldots,n_j,\ldots\rangle &= |\ldots,n_i+1,\ldots,n_j+1,\ldots\rangle\\
&\qquad\times\left[(-1)^{\sum_{k<i}n_k}\right](1-n_i)\left[(-1)^{\sum_{k<j}n_k}\right](1-n_j)\mathrm{sgn}(j-i)\\
\hat{c}_i^\dagger\hat{c}_j|\ldots,n_i,\ldots,n_j,\ldots\rangle &= |\ldots,n_i+1,\ldots,n_j-1,\ldots\rangle\\
&\qquad\times\left[(-1)^{\sum_{k<i}n_k}\right](1-n_i)\left[(-1)^{\sum_{k<j}n_k}\right]n_j\mathrm{sgn}(j-i)\\
\hat{c}_i\hat{c}_j^\dagger|\ldots,n_i,\ldots,n_j,\ldots\rangle &= |\ldots,n_i-1,\ldots,n_j+1,\ldots\rangle\\
&\qquad\times\left[(-1)^{\sum_{k<i}n_k}\right]n_i\left[(-1)^{\sum_{k<j}n_k}\right](1-n_j)\mathrm{sgn}(j-i)\\
\hat{c}_i\hat{c}_j|\ldots,n_i,\ldots,n_j,\ldots\rangle &= |\ldots,n_i-1,\ldots,n_j-1,\ldots\rangle\\
&\qquad\times\left[(-1)^{\sum_{k<i}n_k}\right]n_i\left[(-1)^{\sum_{k<j}n_k}\right]n_j\mathrm{sgn}(j-i)
\end{aligned}
$$

where $\mathrm{sgn}(x)$ gives back the sign of the argument.

$$
\mathrm{sgn}(x)=\begin{cases}-1 & \text{for} \quad x<0\\ +1 & \text{for} \quad x\geq 0\end{cases}
$$

The sign enters, because the number of occupation numbers has changed when the first operator has been applied. If the second orbital has a higher index than the first, this causes one additional sign change.

The expressions with two creators are identical in magnitude when i and j are interchanged except for the change of the sign. This is why the anticommutator of two creation operators vanishes. The same argument holds for the anticommutator of two annihilators. For the anticommutator of one creation and one annihilation operator, we have to interchange the indices i,j in one of the two expressions which introduces a sign change. Other than that the expressions obtained for $\hat{c}_i^\dagger\hat{c}_j$ and $\hat{c}_j\hat{c}_i^\dagger$ are identical.

We obtain

$$
.[\hat{c}_i^\dagger,\hat{c}_j^\dagger]_+=0;\qquad[\hat{c}_i^\dagger,\hat{c}_j]_+=0;\qquad[\hat{c}_i,\hat{c}_j^\dagger]_+=0;\qquad[\hat{c}_i,\hat{c}_j]_+=0\qquad\text{for}\qquad i\neq j
$$

which is consistent with Eq. 8.2 for $i\neq j$.

In the case of the product of two creation operators, the factor $\mathrm{sgn}(j-i)$ comes in for $i>j$, because there is already an additional particle to the left of the state i from the first creation operator, which contributes the additional sign.

8.5 Slater-Condon rules

Once we have defined our basis functions we need to be able to work out the matrix elements of operators in the Fock space. The evaluation of matrix elements is non trivial. In the following we assume, as usual, that the one-particle orbitals form an orthonormal set.

The rules to determine matrix elements of Slater determinants are the so-called **Slater rules** or **Slater-Condon rules**[60].

Before the Slater rules can be used, the Slater determinants must be brought into the form of **maximum coincidence**. That is, we ensure that all one-particle orbitals, which are shared by the two Slater determinants stand in the same positions in both determinants. For example the following determinants are in maximum coincidence with $|abcd\rangle$: $|efcd\rangle$, $|abcd\rangle$,$|aecd\rangle$. The following determinants are not: $|bacd\rangle$, $|eacd\rangle$. The determinants can be brought into maximum coincidence by permuting one-particle orbitals, where each permutation contributes a factor (-1).

Let us start with the definition of one- and two-particle matrix elements.

MATRIX ELEMENTS FOR ONE-PARTICLE OPERATORS AND A TWO-PARTICLE INTERACTION

$$\begin{aligned}A_{a,b}\overset{\text{def}}{=}\langle\phi_a|\hat{A}|\phi_b\rangle &\overset{\text{def}}{=} \int d^4x\ \phi^*_a(\vec{x})\hat{A}\phi_b(\vec{x}) = \int d^4x \int d^4x'\ \phi^*_a(\vec{x})A(\vec{x},\vec{x'})\phi_b(\vec{x'})\\ W_{a,b,c,d}\overset{\text{def}}{=}\langle\phi_a\phi_b|\hat{W}|\phi_c\phi_d\rangle &\overset{\text{def}}{=} \int d^4x \int d^4x'\ \phi^*_a(\vec{x})\phi^*_b(\vec{x'})W(\vec{x},\vec{x'})\phi_c(\vec{x})\phi_d(\vec{x'}) \qquad (8.3)\end{aligned}$$

There are different conventions in choosing the order of the indices in the tensor $\boldsymbol{W}$. Our choice is consistent with the "physicist notation" $< ab|cd >= W_{a,b,c,d}$ as used in the book of Szabo and Ostlund [61] (See their section 2.3.2 "Notations for One- and Two-Electron Integrals"). It is also consistent with the notation used in the Skript *Theoretische Festkörperphysik* (p.63, Eq. 15.47) by Franz Wegener. Szabo and Ostlund also refer to the "chemist notation" $[ij|kl]$ defined as $< ik|jl >= [ij|kl]$, which has lists first both x coordinates and then both x' coordinates. This chemists notation is also used by Mahan's Book[62] (see Eq. 1.157).

Fertig: WS06/07 17 Doppelstunde 17.Jan.07

Here, we only summarize the Slater-Condon rules. Their derivation can be found in appendix O on p. 299

SLATER-CONDON RULES

- Identical Slater determinants, i.e. $|\Psi\rangle = |\Phi\rangle$

$$\begin{aligned}\langle\Psi|\hat{A}|\Phi\rangle &= \sum_{n=1}^{N}\langle\phi_n|\hat{A}|\phi_n\rangle\\ \langle\Psi|\hat{W}|\Phi\rangle &= \frac{1}{2}\sum_{n\neq m}\Big[\langle\phi_n\phi_m|\hat{W}|\phi_n\phi_m\rangle - \langle\phi_n\phi_m|\hat{W}|\phi_m\phi_n\rangle\Big] \qquad (8.4)\end{aligned}$$

- Slater determinants that differ by a single one-particle orbital ϕ_a and ϕ_b

$$\begin{aligned}\langle\Psi|\hat{A}|\Phi\rangle &= \langle\phi_a|\hat{A}|\phi_b\rangle\\ \langle\Psi|\hat{W}|\Phi\rangle &= \sum_{n=1}^{N}\Big[\langle\phi_a\phi_n|\hat{W}|\phi_b\phi_n\rangle - \langle\phi_a\phi_n|\hat{W}|\phi_n\phi_b\rangle\Big] \qquad (8.5)\end{aligned}$$

- Slater determinants that differ by two one-particle orbitals ϕ_a, ϕ_b and ϕ_c, ϕ_d

$$\begin{aligned}\langle\Psi|\hat{A}|\Phi\rangle &= 0\\ \langle\Psi|\hat{W}|\Phi\rangle &= \Big[\langle\phi_a\phi_b|\hat{W}|\phi_c\phi_d\rangle - \langle\phi_a\phi_b|\hat{W}|\phi_d\phi_c\rangle\Big] \qquad (8.6)\end{aligned}$$

- Slater determinants that differ by more than two one-particle orbitals

$$\begin{aligned}\langle\Psi|\hat{A}|\Phi\rangle &= 0\\ \langle\Psi|\hat{W}|\Phi\rangle &= 0 \qquad (8.7)\end{aligned}$$

8.6 Operators expressed by field operators

We have seen that all states of the Fock space can be expressed by creation operators and the vacuum state. Hence, the creation and annihilation operators also provide a means to transform the states of the Fock space into each other. This shows that each operator acting on the Fock space can be expressed by creation and annihilation operators.

In the previous section we derived the Slater-Condon rules, which provide the matrix elements of one- and two-particle operators in terms of the Slater determinants that form the many particle basis-set. In appendix P on p. 303 we use the Slater-Condon rules to derive the explicit expressions for the one- and two-particle operators in this representation. Here, we just provide the results:

A one-particle operator $\hat{A}$ has the form

ONE-PARTICLE OPERATOR IN SECOND QUANTIZATION

$$\hat{A} = \sum_{i,j} A_{i,j} \hat{c}_i^\dagger \hat{c}_j \tag{8.8}$$

and a two-particle operator has the form

TWO-PARTICLE OPERATOR IN SECOND QUANTIZATION

$$\hat{W} = \frac{1}{2} \sum_{i,j,k,l} W_{i,j,k,l} \hat{c}_i^\dagger \hat{c}_j^\dagger \hat{c}_l \hat{c}_k \tag{8.9}$$

Note the reversed order of the annihilators relative to the indices.

The matrix elements $A_{i,j}$ and $W_{i,j,k,l}$ are defined analogously to the expressions for the Slater-Condon rules Eq. 8.3 on p. 161. Note, however, that often the matrix elements are defined differently. (See discussion following Eq. 8.3)

Fertig: WS06/07 18 Doppelstunde 22.Jan.07, Vormittags

Our approach will be to work out the matrix elements of the operators in the form given above in Eq. 8.8 and 8.9 between Slater determinants. Then it will be shown one-by-one that the result is identical to the Slater-Condon rules given above.

8.7 Real-space representation of field operators

Now we have the basic elements of the theory, namely states and operators. The formulation rests on the choice of a particular set of one-particle orbitals $|\phi_i\rangle$. Here, we change the representation from this set of one-particle orbitals to the real-space representation. Thus, we arrive at a formulation analogous to Schrödinger's wave mechanics. Here, we will see that the wave functions are replaced by operators, the field operators.

From the real-space representation, we can introduce any other basisset of one-particle orbitals.

Let us first define the creation and annihilation operators in the real space representation.

FIELD OPERATORS

$$\hat{\psi}^\dagger(\vec{x}) \overset{\text{def}}{=} \sum_i \phi_i^*(\vec{x})\hat{c}_i^\dagger \tag{8.10}$$

$$\hat{\psi}(\vec{x}) = \sum_i \phi_i(\vec{x})\hat{c}_i \tag{8.11}$$

The back transformations are obtained by multiplication of the first lines with $\phi_j(\vec{x})$, integration over space and summation over spin indices.

ORBITAL CREATION AND ANNIHILATION OPERATORS FROM FIELD OPERATORS

$$\hat{c}_i^\dagger = \int d^4x\; \phi_i(\vec{x})\hat{\psi}^\dagger(\vec{x}) \tag{8.12}$$

$$\hat{c}_i = \int d^4x\; \phi_i^*(\vec{x})\hat{\psi}(\vec{x}) \tag{8.13}$$

One-particle operators in real space

Let us now transform a general one-particle operator $\hat{A}$ into the real-space representation

$$\begin{aligned}
\hat{A} &= \sum_{i,j} \langle\phi_i|\hat{A}|\phi_j\rangle \hat{c}_i^\dagger \hat{c}_j \\
&= \sum_{i,j} \langle\phi_i| \underbrace{\int d^4x\; |\vec{x}\rangle\langle\vec{x}|}_{=1} \hat{A} \underbrace{\int d^4x'\; |\vec{x'}\rangle\langle\vec{x'}|}_{=1} |\phi_j\rangle \hat{c}_i^\dagger \hat{c}_j \\
&= \int d^4x \int d^4x' \sum_{i,j} \underbrace{\langle\phi_i|\vec{x}\rangle}_{\phi_i^*(\vec{x})} \underbrace{\langle\vec{x}|\hat{A}|\vec{x'}\rangle}_{A(\vec{x},\vec{x'})} \underbrace{\langle\vec{x'}|\phi_j\rangle}_{\phi_j(\vec{x'})} \hat{c}_i^\dagger \hat{c}_j \\
&= \int d^4x \int d^4x'\; A(\vec{x},\vec{x'}) \underbrace{\sum_i \phi_i^*(\vec{x})\hat{c}_i^\dagger}_{\hat{\psi}^\dagger(\vec{x})} \underbrace{\sum_j \phi_j(\vec{x'})\hat{c}_j}_{\hat{\psi}(\vec{x'})} \\
&= \int d^4x \int d^4x'\; A(\vec{x},\vec{x'})\hat{\psi}^\dagger(\vec{x})\hat{\psi}(\vec{x'})
\end{aligned}$$

Note that we deal here with two different types of operators: The creators and annihilators act on the Fock space. The operator $\hat{A}$ acts on a one-particle state if it is bracketed between two one-particle states.

Let us examine a few examples:

The one-particle part of the Hamiltonian has the form

$$\hat{h} = \int d^4x \int d^4x'\; |\vec{x}\rangle\delta(\vec{x}-\vec{x'}) \left(\frac{-\hbar^2}{2m_e}\vec{\nabla}'^2 + v_{ext}(\vec{x'})\right) \langle\vec{x'}|$$

Thus,

$$
\begin{aligned}
h(\vec{x},\vec{x'}) = \langle\vec{x}|\hat{h}|\vec{x'}\rangle &= \int d^4x'' \int d^4x''' \underbrace{\langle\vec{x}|\vec{x''}\rangle}_{\delta(\vec{x}-\vec{x''})} \delta(\vec{x''}-\vec{x'''}) \left(\frac{-\hbar^2}{2m_e}\vec{\nabla'''}^2 + v_{ext}(\vec{x'''})\right) \underbrace{\langle\vec{x'''}|\vec{x'}\rangle}_{\delta(\vec{x'''}-\vec{x'})} \\
&= \int d^4x'''\, \delta(\vec{x}-\vec{x'''}) \left(\frac{-\hbar^2}{2m_e}\vec{\nabla}'''^2 + v_{ext}(\vec{x'''})\right) \delta(\vec{x'''}-\vec{x'}) \\
&= \delta(\vec{x}-\vec{x'}) \left(\frac{-\hbar^2}{2m_e}\vec{\nabla}'^2 + v_{ext}(\vec{x'})\right)
\end{aligned}
$$

A word of caution is needed with in the last step when we removed the Delta function on the right side of the differential operator. Here, we used that there is always an integral, when a state is applied on the right hand side. In that case the Delta function drops out.

Finally we obtain the one-particle Hamiltonian in the form

$$
\hat{h} = \int d^4x\, \hat{\psi}^\dagger(\vec{x}) \left(\frac{-\hbar^2}{2m_e}\vec{\nabla}^2 + v_{ext}(\vec{x})\right) \hat{\psi}(\vec{x})
$$

The expression just looks like a normal expectation value of the one-particle Hamiltonian. However, since the wave functions are replaced by the creation and annihilation operators, the expression is an operator in Fock space.

Interaction operator in real space

Analogously, we can determine the interaction operator in real space as

$$
\begin{aligned}
\hat{W} &= \frac{1}{2}\sum_{i,j,k,l} W_{i,j,k,l}\hat{c}_i^\dagger \hat{c}_j^\dagger \hat{c}_l \hat{c}_k \\
&= \frac{1}{2}\sum_{i,j,k,l} W_{i,j,k,l} \underbrace{\int d^4x_1\, \phi_i(\vec{x_1})\hat{\psi}^\dagger(\vec{x_1})}_{\hat{c}_i^\dagger} \underbrace{\int d^4x_2\, \phi_j(\vec{x_2})\hat{\psi}^\dagger(\vec{x_2})}_{\hat{c}_j^\dagger} \\
&\quad \underbrace{\int d^4x_3\, \phi_l^*(\vec{x_3})\hat{\psi}(\vec{x_3})}_{\hat{c}_l} \underbrace{\int d^4x_4\, \phi_k^*(\vec{x_4})\hat{\psi}(\vec{x_4})}_{\hat{c}_k} \\
&= \frac{1}{2}\int d^4x_1 \int d^4x_2 \int d^4x_3 \int d^4x_4 \underbrace{\sum_{i,j,k,l} W_{i,j,k,l}\phi_i(\vec{x_1})\phi_j(\vec{x_2})\phi_l^*(\vec{x_3})\phi_k^*(\vec{x_4})}_{W(\vec{x}_1,\vec{x}_2,\vec{x}_3,\vec{x}_4)} \\
&\quad \cdot\hat{\psi}^\dagger(\vec{x_1})\hat{\psi}^\dagger(\vec{x_2})\hat{\psi}(\vec{x_3})\hat{\psi}(\vec{x_4})
\end{aligned}
\qquad (8.14)
$$

Now we need to use the expression for the matrix elements

$$
\begin{aligned}
W_{i,j,k,l} &= \int d^4x \int d^4x'\ \phi_i^*(\vec{x})\phi_j^*(\vec{x'})v_{int}(\vec{x},\vec{x'})\phi_k(\vec{x})\phi_l(\vec{x'}) \\
&\sum_{i,j,k,l} W_{i,j,k,l}\phi_i(\vec{x}_1)\phi_j(\vec{x}_2)\phi_l^*(\vec{x}_3)\phi_k^*(\vec{x}_4) \\
&= \sum_{i,j,k,l}\int d^4x \int d^4x'\ \phi_i^*(\vec{x})\phi_j^*(\vec{x'})v_{int}(\vec{x},\vec{x'})\phi_k(\vec{x})\phi_l(\vec{x'}) \\
&\quad \times\phi_i(\vec{x}_1)\phi_j(\vec{x}_2)\phi_l^*(\vec{x}_3)\phi_k^*(\vec{x}_4) \\
&= \int d^4x \int d^4x'\ \underbrace{\sum_i \phi_i^*(\vec{x})\phi_i(\vec{x}_1)}_{\delta(\vec{x}-\vec{x}_1)}\underbrace{\sum_j \phi_j^*(\vec{x'})\phi_j(\vec{x}_2)}_{\delta(\vec{x'}-\vec{x}_2)} \\
&\quad \times v_{int}(\vec{x},\vec{x'})\underbrace{\sum_k \phi_k(\vec{x})\phi_k^*(\vec{x}_4)}_{\delta(\vec{x}-\vec{x}_4)}\underbrace{\sum_l \phi_l(\vec{x'})\phi_l^*(\vec{x}_3)}_{\delta(\vec{x'}-\vec{x}_3)} \\
&= v_{int}(\vec{x}_1,\vec{x}_2)\delta(\vec{x}_1-\vec{x}_4)\delta(\vec{x}_2-\vec{x}_3)
\end{aligned}
$$

If we insert this result into the above expression Eq. 8.14, we obtain

$$
\begin{aligned}
\hat{W} &\stackrel{Eq.\ 8.14}{=} \frac{1}{2}\int d^4x_1 \int d^4x_2 \int d^4x_3 \int d^4x_4\ v_{int}(\vec{x}_1,\vec{x}_2)\delta(\vec{x}_1-\vec{x}_4)\delta(\vec{x}_2-\vec{x}_3) \\
&\quad \cdot\hat{\psi}^\dagger(\vec{x}_1)\hat{\psi}^\dagger(\vec{x}_2)\hat{\psi}(\vec{x}_3)\hat{\psi}(\vec{x}_4) \\
&= \frac{1}{2}\int d^4x_1 \int d^4x_2\ v_{int}(\vec{x}_1,\vec{x}_2)\hat{\psi}^\dagger(\vec{x}_1)\hat{\psi}^\dagger(\vec{x}_2)\hat{\psi}(\vec{x}_2)\hat{\psi}(\vec{x}_1) \\
&= \frac{1}{2}\int d^4x \int d^4x'\ \hat{\psi}^\dagger(\vec{x})\hat{\psi}^\dagger(\vec{x'})v_{int}(\vec{x},\vec{x'})\hat{\psi}(\vec{x'})\hat{\psi}(\vec{x}) \qquad (8.15)
\end{aligned}
$$

The many-electron Hamiltonian in real space

The Hamiltonian has the form

MANY-PARTICLE HAMILTONIAN

$$
\begin{aligned}
\hat{H} &= \int d^4x\ \hat{\psi}^\dagger(\vec{x})\left(\frac{-\hbar^2}{2m_e}\vec{\nabla}^2 + v_{ext}(\vec{x})\right)\hat{\psi}(\vec{x}) \\
&\quad + \frac{1}{2}\int d^4x \int d^4x'\ \hat{\psi}^\dagger(\vec{x})\hat{\psi}^\dagger(\vec{x'})\frac{e^2}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\hat{\psi}(\vec{x'})\hat{\psi}(\vec{x}) \qquad (8.16)
\end{aligned}
$$

Chapter 9

Ground state and excitations

This chapter is under construction. Proceed on p. 169

The ground state of a many particle state can be expressed as a superposition of all Slater determinants, i.e.

$$|\Psi\rangle=\prod_\alpha\left(\hat{c}_\alpha^\dagger\right)^{\sigma_\alpha}|\mathcal{O}\rangle c_{\vec{\sigma}}$$

The vector $\vec{\sigma}$ is nothing but the number representation of a Slater determinant. It is a vector composed out of zero's and one's for Fermions and of non-negative integers for Bosons.

To obtain the ground state, we minimize the total energy with respect to the wave functions.

$$E_0=\min_\Psi\left[\langle\Psi|\hat{H}|\Psi\rangle-\mathcal{E}\Big(\langle\Psi|\Psi\rangle-1\Big)\right]$$

The energy $\mathcal{E}$ is the Legendre multiplier for the norm constraint. The ground-state wave function $|\mathcal{G}\rangle$ is the minimizing wave function. (Called the **minimizer** of the above expression.) It obeys

$$\left[\hat{H}-\mathcal{E}\right]|\mathcal{G}\rangle=0$$

9.1 Hole excitations

When we discuss excitations, we need to consider particle excitations, hole excitations, particle-hole excitations etc. Different excitations are probed by different experiments.

Imagine a photo-emission experiment, which uses the energy of a photon to kick an electron out of the system into the vacuum. Clearly this process creates a hole. The absorption bands of the incoming photon reflect the energy of hole excitations. These are the energy differences of the $N-1$ states of the system and the ground state, which was present before the photon has been absorbed.

We construct the eigenstates of the $N-1$-electron system by constructing a basis from the N-particle ground state.

$$|\Psi_n\rangle=\sum_n\hat{c}_\alpha|\mathcal{G}\rangle c_{\alpha,n}$$

Now we can construct a Hamilton matrix and overlap matrix for the hole-states.

$$H^h_{\alpha,\beta}\stackrel{\text{def}}{=}\langle\mathcal{G}|\hat{c}_\alpha^\dagger\hat{H}\hat{c}_\beta|\mathcal{G}\rangle$$
$$S^h_{\alpha,\beta}\stackrel{\text{def}}{=}\langle\mathcal{G}|\hat{c}_\alpha^\dagger\hat{c}_\beta|\mathcal{G}\rangle$$

Note that these matrices have the same structure as a one-particle Hamilton- and overlap matrix. Solving the generalized eigenvalue equation yields the eigenstates of the $N-1$ particle system.

$$\sum_{\beta}\left[H^{h}_{\alpha,\beta}-\epsilon^{h}_{n}S^{h}_{\alpha,\beta}\right]c_{\beta,n}=0$$

The energy of the corresponding $N-1$ particle state is $\mathcal{E}_0+\epsilon^{h}_{n}$.

9.2 Two-particle excitations

Optical absorption probes a fundamentally different process as photoemission. Here, we create an electron-hole pair. For most electron hole pairs created the electrons and holes wander off in different directions, so that they do not feel each other. Therefore many excitation energies have the form $\epsilon^{e-h}_{n,m}=\epsilon^{h}_{n}+\epsilon^{e}_{m}$. However, it is also possible that electron and hole form a bound pair due to their charge attraction a so-called **exciton**.

In that case we have to set up a Hamiltonian from N electron states, as the particle number is not changed.

We can now form a basis of electron-hole states from

$$|\Psi_n\rangle=\sum_{n}\hat{c}^{\dagger}_{\alpha}\hat{c}_{\beta}|\mathcal{G}\rangle c_{\alpha,\beta,n}$$

and form a Hamiltonian and overlap matrix in this basis.

$$\begin{aligned}
H^{e-h}_{\alpha,\beta;\gamma,\delta}&\stackrel{\text{def}}{=}\langle\mathcal{G}|\underbrace{\hat{c}^{\dagger}_{\beta}\hat{c}_{\alpha}}_{(\hat{c}^{\dagger}_{\alpha}\hat{c}_{\beta})^{\dagger}}\hat{H}\hat{c}^{\dagger}_{\gamma}\hat{c}_{\delta}|\mathcal{G}\rangle\\
S^{e-h}_{\alpha,\beta;\gamma,\delta}&\stackrel{\text{def}}{=}\langle\mathcal{G}|\underbrace{\hat{c}^{\dagger}_{\beta}\hat{c}_{\alpha}}_{(\hat{c}^{\dagger}_{\alpha}\hat{c}_{\beta})^{\dagger}}\hat{c}^{\dagger}_{\gamma}\hat{c}_{\delta}|\mathcal{G}\rangle
\end{aligned}$$

Chapter 10

Green's functions in one-particle quantum mechanics

10.1 Green's function as inverse of a differential operator

Green's functions are a powerful tool of many-particle physics. However, it also plays an important role in one-particle quantum mechanics. Even more, we use Green's functions in many contexts, often without realizing that we deal with them.

Because explaining Green's functions in the context of many-particle physics is a double difficulty, let us first become familiar with Green's function in the context of one-particle quantum mechanics.

Definition

Consider a differential operator $D(\vec{x}, \vec{\nabla})$. The vector $\vec{x}$ shall be an n-dimensional vector, where n may also be one. It may describe time coordinates, spatial coordinates, both or any other set of variables. A Green's function $G(\vec{x}, \vec{x'})$ is defined by the equation

DEFINING EQUATION FOR A GREEN'S FUNCTION

$$D(\vec{x}, \vec{\nabla})G(\vec{x}, \vec{x}_0) = \delta(\vec{x} - \vec{x}_0)$$

Green's function and inhomogeneous differential equations

A Green's function can be used to solve a general inhomogeneous differential equation

$$\begin{aligned} D(\vec{x}, \vec{\nabla})f(\vec{x}) &= I(\vec{x}) \\ f(\vec{x}) &= \int d^n x_0 \; G(\vec{x}, \vec{x}_0) I(\vec{x}_0) \end{aligned}$$

Examples:

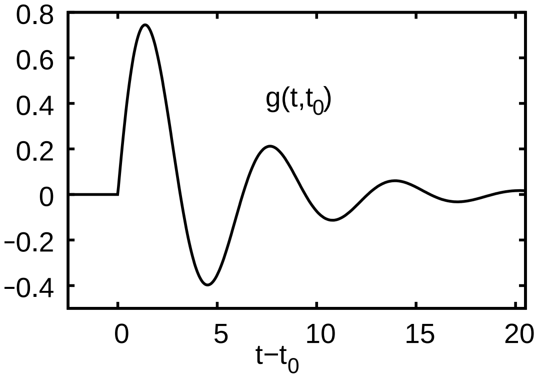
- The Green's function can be used to solve the driven harmonic oscillator

$$m\ddot{x} = -cx - \alpha\dot{x} + f(t)$$
$$\underbrace{\left(m\partial_t^2 + c + \alpha\partial_t\right)}_{D(t,\partial_t)} G(t,t_0) = \delta(t-t_0)$$

  In ΦSX:Klassische Mechanik we determined the corresponding Green's function, which had the form

$$G(t,t') = \frac{-i}{m(\omega_1-\omega_2)}\left(\mathrm{e}^{i\omega_1(t-t')} - \mathrm{e}^{i\omega_2(t-t')}\right)\theta(t-t')$$
$$= \frac{1}{m\sqrt{\frac{c}{m}-\left(\frac{\alpha}{2m}\right)^2}}\sin\left(\sqrt{\frac{c}{m}-\left(\frac{\alpha}{2m}\right)^2}\,(t-t')\right)\mathrm{e}^{-\frac{\alpha}{2m}(t-t')}$$

  where ω_1 and ω_2 are the complex eigenfrequencies of the harmonic oscillator.[1]



  Once the Green's function is known, we can immediately write down the solution for the driven harmonic oscillator, which fulfills the corresponding inhomogeneous equations of motion. The solution is

$$x(t) = \int dt'\, G(t,t')f(t')$$

  In this case the Green's function is the response to a Kraftstoß[2]

- The Green's function has been used to solve the Poisson equation

$$\vec\nabla^2\Phi(\vec r) = -\frac{1}{\epsilon_0}\rho(\vec r)$$

  The Green's function for this problem has the form

$$\nabla^2 G(\vec r,\vec r') = \delta(\vec r-\vec r') \qquad \Rightarrow \qquad G(\vec r,\vec r') = \frac{-1}{4\pi|\vec r-\vec r'|}$$

  so that we obtain the solution as

$$\Phi(\vec r) = \int d^3r'\, G(\vec r,\vec r')\left(\frac{-\rho(\vec r')}{\epsilon_0}\right) = \int d^3r'\,\frac{\rho(\vec r')}{4\pi\epsilon_0|\vec r-\vec r'|}$$

---

[1] For the undamped harmonic oscillator the frequencies are $\omega_1 = \sqrt{\frac{c}{m}}$ and $\omega_2 = -\sqrt{\frac{c}{m}}$

[2] There does not seem to exist an english word for the german word "Kraftstoß". It describes a force peak in the form of a δ function, which transfers a finite momentum in an infinitesimal time interval.

- The Green's function of the Schrödinger equation with a time-independent Hamilton operator

$$\left[i\hbar\partial_t - \left(\frac{-\hbar^2}{2m}\vec{\nabla}^2 + v(\vec{r})\right)\right]\psi(\vec{r},t) = 0$$

  is defined by

$$\left[i\hbar\partial_t - \left(\frac{-\hbar^2}{2m}\vec{\nabla}^2 + v(\vec{r})\right)\right]G(\vec{r},t,\vec{r'},t') = \delta(t-t')\delta(\vec{r}-\vec{r'})$$

  Using the eigenfunctions $\phi_n(\vec{r})$ of the Hamilton operator and its eigenvalues ϵ_n we can represent the Green's function as

$$G(\vec{r},t,\vec{r'},t') = \frac{1}{i\hbar}\sum_n \phi_n(\vec{r})\phi_n^*(\vec{r'})\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}\theta(t-t')$$

Green's function as inverse of an operator, a differential operator or a matrix

In the following we will show the role of the Green's function in different contexts, that is once for a differential equation, once for an abstract operator and once for a matrix equation. The analogy between these version will be used frequently in the future.

$$D(x,\vec{\nabla})G(\vec{x},\vec{x}_0) = \delta(\vec{x}-\vec{x}_0) \quad : \quad D(x,\vec{\nabla})f(\vec{x}) = I(x) \quad \Rightarrow \quad f(\vec{x}) = \int d^nx'\; G(\vec{x},\vec{x'})I(\vec{x'}) \tag{10.1}$$

$$\hat{D}\hat{G} = \hat{1} \quad : \quad \hat{D}|f\rangle = |I\rangle \quad \Rightarrow \quad |f\rangle = \hat{G}|I\rangle \tag{10.2}$$

$$\boldsymbol{DG} = \boldsymbol{1} \quad : \quad \boldsymbol{D}\vec{f} = \vec{I} \quad \Rightarrow \quad \vec{f} = \boldsymbol{G}\vec{I} \tag{10.3}$$

The analogy with the matrix form shows that the Green's function can be considered the inverse of a differential operator, a general operator or a matrix. In particular, we see that a differential equation is closely related to a system of regular equations.

The various forms can be transformed into each other. We start with the most general form, the abstract notation

$$\hat{D}\hat{G} = \hat{1}$$

and derive the corresponding other equations.

$$\langle\phi_i|\hat{D}\underbrace{\sum_k |\phi_k\rangle\langle\phi_k|}_{\hat{1}}\hat{G}|\phi_j\rangle = \langle\phi_i|\hat{1}|\phi_j\rangle \qquad \langle\vec{x}|\hat{D}\underbrace{\int d^4x'\;|\vec{x'}\rangle\langle\vec{x'}|}_{\hat{1}}\hat{G}|\vec{x}_0\rangle = \langle\vec{x}|\hat{1}|\vec{x}_0\rangle$$

$$\sum_k D_{i,k}G_{k,j} = \delta_{i,j} \qquad \int d^4x'\;\langle\vec{x}|\hat{D}|\vec{x'}\rangle G(\vec{x},\vec{x}_0) = \delta(\vec{x}-\vec{x}_0) \tag{10.4}$$

Here, we used the matrix elements of the Green's function

$$G_{i,j} \overset{\text{def}}{=} \langle\phi_i|\hat{G}|\phi_j\rangle \qquad G(\vec{x},\vec{x'}) \overset{\text{def}}{=} \langle\vec{x}|\hat{G}|\vec{x'}\rangle$$

A differential operator can be written as matrix elements of a special operator

$$\hat{D} = \int d\vec{x}\;|\vec{x}\rangle D(\vec{x},\vec{\nabla})\langle\vec{x}|$$

which has the matrix elements

$$\langle\vec{x}|\hat{D}|\vec{x'}\rangle = \int d\vec{x''}\;\langle\vec{x}|\vec{x''}\rangle D(\vec{x''},\vec{\nabla}'')\langle\vec{x''}|\vec{x'}\rangle = \delta(\vec{x}-\vec{x'})D(\vec{x'},\vec{\nabla}')$$

Insertion into Eq. 10.4 yields the first equations in Eq. 10.1 and Eq. 10.3 respectively.

10.2 Green's function of a time-independent Hamiltonian

Consider a time-independent Hamiltonian H, which has eigenstates $|\psi_n\rangle$ with eigenvalues ϵ_n. The differential operator describing the Schrödinger equation has the form

$$\hat{D} = \left[i\hbar\partial_t - \hat{H}\right] = i\hbar\partial_t - \underbrace{\sum_n |\psi_n\rangle\epsilon_n\langle\psi_n|}_{\hat{H}} = \sum_n |\psi_n\rangle \left(i\hbar\partial_t - \epsilon_n\right)\langle\psi_n| \tag{10.5}$$

The Green's function, in its operator form with respect to the spatial variables, is defined by

$$\hat{D}\hat{G}(t, t_0) = \hat{1}\delta(t - t_0) \tag{10.6}$$

We can represent the Green's function in eigenstates $|\psi_n\rangle$ of the Hamiltonian

$$\hat{G} = \underbrace{\sum_n |\psi_n\rangle\langle\psi_n|}_{\hat{1}}\,\hat{G}\underbrace{\sum_m |\psi_m\rangle\langle\psi_m|}_{\hat{1}} = \sum_{n,m}|\psi_n\rangle\underbrace{\langle\psi_n|\hat{G}|\psi_m\rangle}_{G_{n,m}}\langle\psi_m| = \sum_{n,m}|\psi_n\rangle G_{n,m}\langle\psi_m|$$

where $G_{n,m}(t, t')$ are the corresponding matrix elements of the Green's function.

Now we can work out the Green's function

$$\begin{aligned}
\hat{D}\hat{G}(t, t') &= \hat{1}\delta(t - t')\\
\underbrace{\sum_j |\psi_j\rangle\left(i\hbar\partial_t - \epsilon_j\right)\langle\psi_j|}_{=\hat{D}\quad\text{(See Eq. 10.5)}}\underbrace{\sum_{n,m}|\psi_n\rangle G_{n,m}(t, t')\langle\psi_m|}_{\hat{G}} &= \hat{1}\delta(t - t')\\
\sum_{j,n,m}|\psi_j\rangle\left(i\hbar\partial_t - \epsilon_j\right)\underbrace{\langle\psi_j|\psi_n\rangle}_{\delta_{j,n}}G_{n,m}(t, t')\langle\psi_m| &= \hat{1}\delta(t - t')\\
\sum_{n,m}|\psi_n\rangle\left(i\hbar\partial_t - \epsilon_n\right)G_{n,m}(t, t')\langle\psi_m| &= \underbrace{\sum_{n,m}|\psi_n\rangle\delta_{n,m}\langle\psi_m|}_{\hat{1}}\\
\Rightarrow \left(i\hbar\partial_t - \epsilon_n\right)G_{n,m}(t, t') &= \delta_{n,m}\delta(t - t')
\end{aligned} \tag{10.7}$$

In the last step we exploited that the eigenstates are linear independent, which implies that each operator cannot be represented by two different sets of matrix elements.

Thus, we arrived at an ordinary differential equation for the matrix elements, namely Eq. 10.7, which has the solution[3]

$$G_{n,m}(t, t') = \frac{1}{i\hbar}\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}\theta(t - t')\delta_{n,m}$$

with the matrix elements we can also express the Green's function in operator form

---

[3] Let us make the test (proof may continue on the next page):

$$\begin{aligned}
&\left[i\hbar\partial_t - \epsilon_n\right]\frac{1}{i\hbar}\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}\theta(t - t')\delta_{n,m}\\
&= \left(\partial_t\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}\right)\theta(t - t')\delta_{n,m} + \mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}\left(\partial_t\theta(t - t')\right)\delta_{n,m} - \frac{1}{i\hbar}\epsilon_n\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}\theta(t - t')\delta_{n,m}\\
&= \underbrace{\left(\frac{1}{i\hbar}\epsilon_n\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}\right)\theta(t - t')\delta_{n,m}}_{X} + \underbrace{\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}}_{=1\text{ for }t=t'}\left(\delta(t - t')\right)\delta_{n,m} - \underbrace{\frac{1}{i\hbar}\epsilon_n\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}\theta(t - t')\delta_{n,m}}_{X}\\
&= \delta(t - t')\delta_{n,m}
\end{aligned}$$

$$\boxed{\hat{G}(t,t')=\frac{1}{i\hbar}\theta(t-t')\sum_n|\psi_n\rangle \mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}\langle\psi_n| \qquad (10.8)}$$

and we can determine the Green's function in real space representation

$$\begin{aligned}
G(\vec{x},t,\vec{x'},t') &\overset{\text{def}}{=} \langle\vec{x}|\hat{G}(t,t')|\vec{x'}\rangle = \sum_n \langle\vec{x}|\psi_n\rangle\frac{1}{i\hbar}\theta(t-t_0)\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}\langle\psi_n|\vec{x'}\rangle \\
&= \frac{1}{i\hbar}\sum_n \psi_n(\vec{x})\psi_n^*(\vec{x'})\theta(t-t_0)\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}
\end{aligned}$$

Fertig: WS06/07 19 Doppelstunde 22.Jan.07, Nachmittags

The Green's function introduced here is the **retarded Green's function**. The retarded Green's function $\hat{G}(t,t')$ vanishes for $t<t'$. There are also other Green's function such as the advanced Green's function, which will become important in the many-particle formalism.

Note the Green's function given above is not the only solution of the differential equation. We are free to add a solution of the homogeneous equation to obtain another solution. For example we can subtract

$$\frac{1}{i\hbar}\sum_n|\psi_n\rangle\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}\langle\psi_n|$$

to obtain the advanced Green's function

$$\hat{G}^A(t,t')=\frac{-1}{i\hbar}\theta(t'-t)\sum_n|\psi_n\rangle\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}\langle\psi_n|$$

10.2.1 Green's function in the energy representation

The Green's function only depends on the difference of its time arguments, which is a consequence of our requirement that the Hamiltonian is time-independent.

$$G(\vec{x},t,\vec{x'},t_0)=G(\vec{x},\vec{x'},t-t_0)$$

This allows to Fourier transform the Green's function in the time-difference argument.

$$\begin{aligned}
G(\vec{x},\vec{x'},\epsilon) &\overset{\text{def}}{=} \int_{-\infty}^{\infty}dt\; G(\vec{x},\vec{x'},t)\;\mathrm{e}^{\frac{i}{\hbar}\epsilon t} \\
&\overset{\text{Eq. 10.8}}{=} \frac{1}{i\hbar}\sum_n\psi_n(\vec{x})\psi_n^*(\vec{x'})\int_{-\infty}^{\infty}dt\;\theta(t)\mathrm{e}^{-\frac{i}{\hbar}(\epsilon_n-\epsilon)t} \\
&= \frac{1}{i\hbar}\sum_n\psi_n(\vec{x})\psi_n^*(\vec{x'})\int_0^{\infty}dt'\;\mathrm{e}^{-\frac{i}{\hbar}(\epsilon_n-\epsilon)t'} \qquad (10.9)
\end{aligned}$$

The integral does not converge if we let the upper limit of the integral go to infinity, while ϵ is real. If we consider the energy argument of the Green's function to be a complex number, we see that the integral converges only, if the imaginary part of ϵ is positive. If it is real or if it has a negative imaginary part, the result is undefined.

$$\begin{aligned}
G(\vec{x},\vec{x}',\epsilon+i\eta) &\overset{\text{Eq. 10.9}}{=} \frac{1}{i\hbar}\sum_n \psi_n(\vec{x})\psi_n^*(\vec{x}')\int_0^\infty dt'\ \mathrm{e}^{-\frac{i}{\hbar}(\epsilon_n-\epsilon-i\eta)t'} \\
&= \frac{1}{i\hbar}\sum_n \psi_n(\vec{x})\psi_n^*(\vec{x}')\left[\frac{i\hbar}{\epsilon_n-\epsilon-i\eta}\mathrm{e}^{-\frac{i}{\hbar}(\epsilon_n-\epsilon-i\eta)t'}\right]_0^\infty \\
&\overset{\eta>0}{=} \frac{1}{i\hbar}\sum_n \psi_n(\vec{x})\psi_n^*(\vec{x}')\frac{-i\hbar}{\epsilon_n-\epsilon-i\eta} \\
&= -\sum_n \frac{\psi_n(\vec{x})\psi_n^*(\vec{x}')}{\epsilon_n-\epsilon-i\eta} \\
&= \sum_n \frac{\psi_n(\vec{x})\psi_n^*(\vec{x}')}{\epsilon-\epsilon_n+i\eta}
\end{aligned}$$

Notice the restriction that the imaginary part of $\epsilon+i\eta$ must be positive. The Green's function is only defined in the upper half-plane of the complex plane. While we can evaluate the final expression also in the lower half plane, the Green's function is simply not defined there.

Thus, we can express the Green's function as

GREEN'S FUNCTION IN THE ENERGY REPRESENTATION

$$\hat{G}(\epsilon+i\eta) = \sum_n |\psi_n\rangle\frac{1}{\epsilon-\epsilon_n+i\eta}\langle\psi_n| = (\epsilon+i\eta-\hat{H})^{-1} \qquad \text{for} \quad \eta>0 \tag{10.10}$$

Real and imaginary part of a schematic Green's function are shown in Fig. 10.1.

The Green's function as function of energy obey's the equation

$$\left(\epsilon-\hat{H}\right)\hat{G}(\epsilon) = \hat{1} \tag{10.11}$$

which can be verified by insertion

$$\begin{aligned}
\left(\epsilon-\hat{H}\right)\hat{G}(\epsilon) &= \left(\epsilon-\hat{H}\right)\sum_n |\psi_n\rangle\frac{1}{\epsilon-\epsilon_n}\langle\psi_n| \\
&= \sum_n |\psi_n\rangle\frac{\epsilon-\epsilon_n}{\epsilon-\epsilon_n}\langle\psi_n| = \sum_n |\psi_n\rangle\langle\psi_n| = \hat{1}
\end{aligned}$$

10.3 Projected Density of States

So far the Green's function is still a fairly abstract object. Let us therefore show a few physical quantities that can be obtained from the Green's function.

The Green's function provides us with the density of states projected onto a state $|\chi\rangle$ as

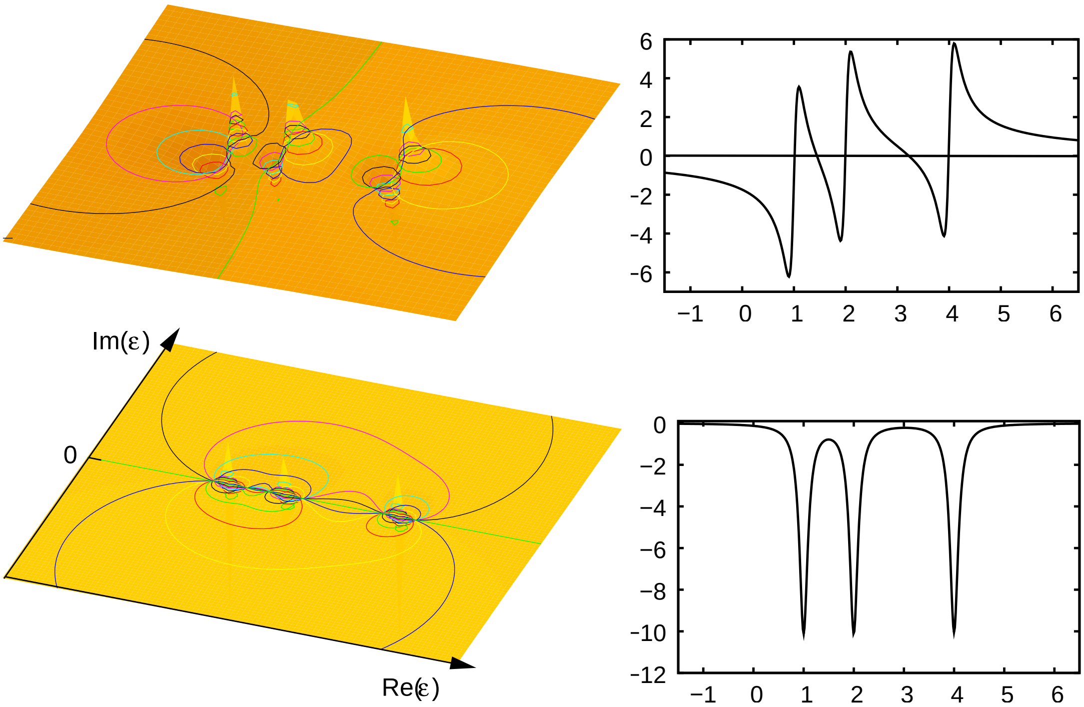


Fig. 10.1: Green's function in the complex plane (left) and on the real axis of the energy for a system with three eigenvalues. Top: real part of $G(\epsilon)$. Bottom: imaginary part.

PROJECTED DENSITY OF STATES

$$D_\chi(\epsilon)=\sum_n\langle\chi|\psi_n\rangle\delta(\epsilon-\epsilon_n)\langle\psi_n|\chi\rangle=-\frac{1}{\pi}\lim_{\eta\to0^+}\mathrm{Im}[\langle\chi|\hat{G}(\epsilon+i\eta)|\chi\rangle] \tag{10.12}$$

The proof will be given below. Note that we cannot determine off-site elements of the density of states from the Green's function.

Total density of states

Thus, we gain access to the excitation spectrum of the Hamiltonian. The total density of states can be obtained

$$\begin{aligned}
D_{tot}(\epsilon) &\overset{Eq.\ 4.57}{=} \sum_n \delta(\epsilon-\epsilon_n)\\
&= \sum_n\int d^4x\ \langle\vec{x}|\psi_n\rangle\delta(\epsilon-\epsilon_n)\langle\psi_n|\vec{x}\rangle\\
&\overset{Eq.\ 10.12}{=} -\frac{1}{\pi}\mathrm{Im}[\int d^4x\ \langle\vec{x}|\hat{G}(\epsilon+i\eta)|\vec{x}\rangle]\\
&= -\frac{1}{\pi}\lim_{\eta\to0^+}\mathrm{Im}[Tr[\hat{G}(\epsilon+i\eta)]]
\end{aligned} \tag{10.13}$$

Electron density

Integration over the occupied states provides us with the probability that an electron is in a given orbital. The orbital $|\chi\rangle$ could also be $|\vec{x}\rangle$, in which case we obtain the electron density.

$$\begin{aligned} n(\vec{r}) &= \sum_{\sigma}\sum_{n}\int^{\epsilon_F} d\epsilon\ \langle\vec{r},\sigma|\psi_n\rangle\delta(\epsilon-\epsilon_n)\langle\psi_n|\vec{r},\sigma\rangle \\ &= -\frac{1}{\pi}\sum_{\sigma}\int^{\epsilon_F} d\epsilon\ \mathrm{Im}[\langle\vec{r},\sigma|\hat{G}(\epsilon)|\vec{r},\sigma\rangle \end{aligned}$$

Proof of Eq. 10.12

Let us derive this Eq. 10.12: First we determine the imaginary part of the expectation value of the Green's function with a one-particle orbital $|\chi\rangle$

$$\begin{aligned} \langle\chi|\hat{G}|\chi\rangle &= \sum_n \langle\chi|\psi_n\rangle\frac{1}{\epsilon-\epsilon_n+i\eta}\langle\psi_n|\chi\rangle \\ &= \sum_n \langle\chi|\psi_n\rangle\frac{\epsilon-\epsilon_n-i\eta}{(\epsilon-\epsilon_n)^2+\eta^2}\langle\psi_n|\chi\rangle \\ \Rightarrow \mathrm{Im}[\langle\chi|\hat{G}(\epsilon)|\chi\rangle] &= \sum_n \langle\chi|\psi_n\rangle\frac{-\eta}{(\epsilon-\epsilon_n)^2+\eta^2}\langle\psi_n|\chi\rangle \end{aligned}$$

The last step works only for expectation values, that is the bra applied to the left must be correspond to the same orbital as the ket applied to the left. Only then we obtain a real contribution from the orbitals, namely $|\langle\chi|\psi_n\rangle|^2$.

Below we will show that

$$\lim_{\eta\to 0}\frac{\eta}{x^2+\eta^2} = \pi\delta(x) \tag{10.14}$$

With the identity Eq.10.14 we obtain directly Eq. 10.12.

In the following we will prove Eq. 10.14. We can first show that the expression vanishes except at the origin.

$$\lim_{\eta\to 0}\frac{\eta}{x^2+\eta^2} \overset{x\neq 0}{=} \lim_{\eta\to 0}\frac{\eta}{x^2} = 0 \qquad \text{for} \qquad x\neq 0$$

For $x=0$ the function diverges

$$\lim_{\eta\to 0}\frac{\eta}{x^2+\eta^2} \overset{x=0}{=} \lim_{\eta\to 0}\frac{1}{\eta} = \infty \qquad \text{for} \qquad x=0$$

Now we need to show that the integral converges to a constant

$$\begin{aligned} \int dx\ \frac{\eta}{x^2+\eta^2} &= \int \frac{dx}{\eta}\ \frac{1}{\left(1+\frac{x}{\eta}\right)^2} \\ &\overset{z\overset{\text{def}}{=}\frac{x}{\eta}}{=} \int dz\ \frac{1}{1+z^2} \end{aligned}$$

Now we introduce another variable transform from z to q

$$\begin{aligned} q &\overset{\text{def}}{=} \mathrm{atan}(z) \\ z &= \tan(q) \\ \frac{dz}{dq} &= \frac{\sin^2(q)+\cos^2(q)}{\cos^2(q)} = \frac{\sin^2(q)}{\cos^2(q)}+1 = 1+\tan^2(q) = 1+z^2 \end{aligned}$$

which yields

$$\int dz \, \frac{1}{1+z^2} = \int_{-\infty}^{\infty} dz \, \frac{dq}{dz} = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} dq = \pi$$

Thus, we have shown that $\frac{1}{\pi} \lim_{\eta \to 0} \frac{\eta}{x^2+\eta^2}$ vanishes everywhere except at the origin, and that it integrates to one, which shows that it is the δ-function [4]. This concludes the proof of Eq. 10.12.

---

[4] A proper proof would need to show $f(0) = \frac{1}{\pi} \lim_{\eta \to 0} \int dx \; f(x) \frac{\eta}{x^2+\eta^2}$, which we have not done here. One can construct distributions that fulfill the equations we have shown, that however are not the δ function. Thus, we cannot say that we have proven the statement but shown it beyond reasonable doubt, whatever that means.

Chapter 11

Perturbation theory in the one-particle theory

One of the main advantages of the Green's function is that a perturbative treatment is relatively simple.

Consider a Hamiltonian, that consists of an unperturbed part H_0 and a perturbation $\hat{W}$.

$$\Big(\epsilon-\hat{H}_0-\hat{W}\Big)\hat{G}(\epsilon)\overset{Eq.\ 10.11}{=}\hat{1}$$

We assume that we know the Green's function $G_0(\epsilon)$ of the unperturbed system, which fulfills

$$\Big(\epsilon-\hat{H}_0\Big)\hat{G}_0(\epsilon)\overset{Eq.\ 10.11}{=}\hat{1}$$

Thus, we can determine the perturbed Green's function as

$$\begin{aligned}
\Big(\epsilon-\hat{H}_0-\hat{W}\Big)\hat{G}(\epsilon)&=\hat{1}\\
\overset{\hat{G}_0\cdot}{\Rightarrow}\hat{G}_0(\epsilon)\Big(\epsilon-\hat{H}_0-\hat{W}\Big)\hat{G}(\epsilon)&=\hat{G}_0(\epsilon)\\
\Rightarrow\underbrace{\hat{G}_0\Big(\epsilon-\hat{H}_0\Big)}_{=\hat{1}}\hat{G}(\epsilon)-\hat{G}_0\hat{W}\hat{G}(\epsilon)&=\hat{G}_0\\
\Rightarrow\Big(\hat{1}-\hat{G}_0\hat{W}\Big)\hat{G}(\epsilon)&=\hat{G}_0
\end{aligned}$$

From this equation we obtain the Green's function for the perturbed system as

DYSON'S EQUATION

$$\Rightarrow\hat{G}(\epsilon)=\Big(\hat{1}-\hat{G}_0\hat{W}\Big)^{-1}\hat{G}_0=\Big(\hat{G}_0^{-1}-\hat{W}\Big)^{-1}=\hat{G}_0+\hat{G}_0\hat{W}\hat{G}\qquad(11.1)$$

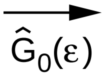
I am showing here three different expressions for the full Green's function, that are all equivalent. The first expression suggests immediately a perturbation series, using the Taylor expansion of $(1+x)^{-1}=\sum_{n=0}^{\infty}(-x)^n$.

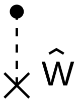
Thus, we obtain

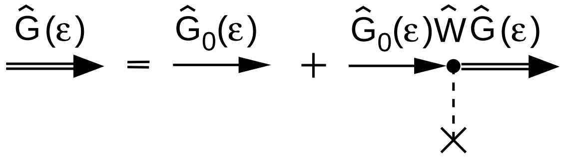
$$
\begin{aligned}
\hat{G}(\epsilon) &= \sum_{n=0}^{\infty}\Big(\hat{G}_0(\epsilon)\hat{W}\Big)^n\hat{G}_0(\epsilon)\\
&= \hat{G}_0(\epsilon)+\hat{G}_0(\epsilon)\hat{W}\hat{G}_0(\epsilon)+\hat{G}_0(\epsilon)\hat{W}\hat{G}_0(\epsilon)\hat{W}\hat{G}_0(\epsilon)+\dots
\end{aligned}
\tag{11.2}
$$

11.1 Feynman diagrams

Here, we can already come into contact with **Feynman diagrams**. As shown below, the unperturbed Green's function is represented by a simple arrow, and the perturbed Green's function by a double arrow. The perturbation is replaced by a point connected to a cross. With this notation, Eq. 11.1 has the form.







Feynman diagrams play an important role in the evaluation of the Green's function in many-particle theory. While each symbol is only a representation of a mathematical object, the Feynman diagrams offer an extraordinarily intuitive notation, so that one begins, after a while, to develop a physical intuition about them.

Fertig: WS06/07 20 Doppelstunde 24.Jan.07

11.2 Time-dependent Perturbation

Definition of the propagator

Let us define the time propagation operator[1]

$$|\phi(t)\rangle = \hat{U}(t,t')|\phi(t')\rangle$$

The time propagation operator fulfills the equation

$$\Big(i\hbar\partial_t-\hat{H}\Big)\hat{U}(t,t')=\hat{0}$$

The time propagation operator can formally be written[2] as

$$\hat{U}(t,t')=\sum_n|\psi_n(t)\rangle\langle\psi_n(t')| \tag{11.3}$$

---

[1] germ: Zeitentwicklungsoperator

[2] Let us start with a complete orthonormal system of states $|\chi_n\rangle$ and let us solve the Schrödinger equation

$$\left(i\hbar\partial_t-\hat{H}\right)|\psi_n(t)\rangle=0$$

with the initial condition $|\psi(t_0)\rangle=|\chi_n\rangle$.

Then we can show that, for an arbitrary state $|\phi_0\rangle$, the wave function

$$|\phi(t)\rangle=\sum_n|\psi_n(t)\rangle\langle\psi_n(t_0)|\phi_0\rangle$$

where the states $|\psi_n(t)\rangle$ are a complete orthonormal basis set which fulfills the time-dependent Schrödinger equation.

Properties of the Propagator

We can immediately write down some properties of $\hat{U}_I(t)$

$$\hat{U}(t,t) = \hat{1} \tag{11.4}$$

$$\hat{U}(t,t'')\hat{U}(t'',t') = \hat{U}(t,t') \tag{11.5}$$

$$\hat{U}(t',t)\hat{U}(t,t') \overset{Eq.\ 11.5,}{=} \hat{1} \tag{11.6}$$

$$\hat{U}(t',t) \overset{Eq.\ 11.6}{=} \hat{U}^{-1}(t,t') \tag{11.7}$$

$$\hat{U}(t',t) \overset{Eq.\ 11.3}{=} \hat{U}^{\dagger}(t,t') \tag{11.8}$$

$$\hat{U}^{\dagger}(t,t') \overset{Eq.\ 11.7,11.8}{=} \hat{U}^{-1}(t,t') \qquad \text{Unitarity} \tag{11.9}$$

11.2.1 Perturbation series

Let us now consider a perturbation of the Hamiltonian, so that

$$\Big(i\hbar\partial_t - \hat{H}_0 - \lambda\hat{W}(t)\Big)\hat{U}(t,t') = \hat{0}$$

and let us assume that the propagator for the unperturbed system is known

$$\Big(i\hbar\partial_t - \hat{H}_0\Big)\hat{U}^{(0)}(t,t') = \hat{0}$$

11.2.2 Propagator of a time-independent Hamiltonian

While the formalism is more general, one usually uses a time-independent system as unperturbed system. Let us therefore examine the propagator for a time-independent Hamiltonian $\hat{H}^{(0)}$.

$$\hat{U}^{(0)}(t,t') = \sum_n |\psi_n^{(0)}\rangle e^{-\frac{i}{\hbar}\epsilon_n^{(0)}(t-t')}\langle\psi_n^{(0)}| \tag{11.10}$$

where $|\psi_n^{(0)}\rangle$ are the eigenstates of the unperturbed Hamiltonian

$$\hat{H}_0|\psi_n^{(0)}\rangle = |\psi_n^{(0)}\rangle\epsilon_n^{(0)}$$

and $\epsilon_n^{(0)}$ are the corresponding energy eigenvalues. The time dependent wave functions have the form.

$$|\psi_n^{(0)}(t)\rangle = |\psi_n^{(0)}\rangle e^{-\frac{i}{\hbar}\epsilon_n^{(0)}t}$$

---

solves the Schrödinger equation

$$\left(i\hbar\partial_t - \hat{H}\right)|\phi(t)\rangle = \sum_n \underbrace{\left(i\hbar\partial_t - \hat{H}\right)|\psi_n(t)\rangle}_{|0\rangle}\langle\psi_n(t_0)|\phi_0\rangle = 0$$

with the initial condition $|\phi(t_0)\rangle = |\phi_0\rangle$

$$|\phi(t_0)\rangle = \sum_n |\psi_n(t_0)\rangle\langle\psi_n(t_0)|\phi_0\rangle \underbrace{\sum_n |\chi_n\rangle\langle\chi_n|}_{=\hat{1}}\phi_0\rangle = |\phi_0\rangle$$

Thus, the operator

$$\hat{U}(t,t_0) = \sum_n |\psi_n(t)\rangle\langle\psi_n(t_0)|$$

propagates any state from t_0 to the time t.

which can be checked by insertion into the time-dependent Schrödinger equation.

This propagator can be written[3] as

$$\hat{U}^{(0)}(t,t') = \mathrm{e}^{-\frac{i}{\hbar}\hat{H}^{(0)}(t-t')}$$

11.2.3 Interaction picture

We would like to get rid of the unperturbed Hamiltonian, by using the unperturbed propagator.

$$\Big(i\hbar\partial_t \overbrace{-\hat{H}_0 - \hat{W}(t)}^{-\hat{H}(t)}\Big)\underbrace{\hat{U}^{(0)}(t,0)\hat{U}^{(0)}(0,t)}_{=\hat{1}}\hat{U}(t,t') = \hat{0}$$

$$\hat{U}^{(0)}(t,0)i\hbar\partial_t\Big(\hat{U}^{(0)}(0,t)\hat{U}(t,t')\Big) = \hat{W}(t)\hat{U}^{(0)}(t,0)\hat{U}^{(0)}(0,t)\hat{U}(t,t')$$

We can give this expression a more symmetric form by multiplying the unperturbed propagator on the left and the right.

$$\begin{aligned}&\underbrace{\hat{U}^{(0)}(0,t)\hat{U}^{(0)}(t,0)}_{\hat{1}}i\hbar\partial_t\underbrace{\hat{U}^{(0)}(0,t)\hat{U}(t,t')\hat{U}^{(0)}(t',0)}_{\hat{U}_I}\\ =&\underbrace{\hat{U}^{(0)}(0,t)\hat{W}(t)\hat{U}^{(0)}(t,0)}_{\hat{W}_I(t)}\underbrace{\hat{U}^{(0)}(0,t)\hat{U}(t,t')\hat{U}^{(0)}(t',0)}_{\hat{U}_I(t,t')}\end{aligned}$$

INTERACTION PICTURE

We introduce here the **interaction picture** where an operator has the form

$$\hat{A}_I(t) = \hat{U}^{(0)}(0,t)\hat{A}(t)\hat{U}^{(0)}(t,0) \tag{11.11}$$

When we change the representation of the operators, we also need to change the representation of the wave functions

$$|\psi_I(t)\rangle = \hat{U}^{(0)}(0,t)|\psi(t)\rangle \tag{11.12}$$

so that the expectation values are unchanged by changing the representation, that is

$$\langle\Psi_I(t)|\hat{A}_I(t)|\Psi_I(t)\rangle = \langle\Psi(t)|\hat{A}(t)|\Psi(t)\rangle$$

11.2.4 Series expansion

The propagator in the interaction picture obeys the simple equation

---

[3] The proof that this result is identical to Eq. 11.10 goes as follows:

$$\begin{aligned}\mathrm{e}^{-\frac{i}{\hbar}\hat{H}^{(0)}(t-t')} &= \overbrace{\sum_n|\psi_n^{(0)}\rangle\langle\psi_n^{(0)}|}^{=\hat{1}}\mathrm{e}^{-\frac{i}{\hbar}\hat{H}^{(0)}(t-t')}\overbrace{\sum_m|\psi_m^{(0)}\rangle\langle\psi_m^{(0)}|}^{=\hat{1}} = \sum_{n,m}|\psi_n^{(0)}\rangle\delta_{n,m}\mathrm{e}^{-\frac{i}{\hbar}\epsilon_m^{(0)}(t-t')}\langle\psi_m^{(0)}|\\ &= \sum_n|\psi_n^{(0)}\rangle\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n^{(0)}(t-t')}\langle\psi_n^{(0)}|\end{aligned}$$

DYNAMICS OF THE PROPAGATOR IN THE INTERACTION PICTURE

$$i\hbar\partial_t \hat{U}_I(t,t') = \hat{W}_I(t)\hat{U}_I(t,t')$$

which can be solved as

$$\begin{aligned}\hat{U}_I(t_f,t_i) &= \underbrace{\hat{U}_I(t_i,t_i)}_{=\hat{1}} + \frac{1}{i\hbar}\int_{t_i}^{t_f} dt'\, \hat{W}_I(t')\hat{U}_I(t',t_i)\\ &= \hat{1} + \frac{1}{i\hbar}\int_{t_i}^{t_f} dt'\, \hat{W}_I(t')\hat{U}_I(t',t_i)\end{aligned}$$

We can solve this equation iteratively. We introduce the propagator up to n-th order in $\hat{W}$ as $\hat{U}_I^{(n)}(t,t')$.

In zeroth order we obtain the unperturbed propagator

$$\hat{U}_I^{(0)}(t,t') = \hat{1}$$

which shows that the unperturbed wave functions in the interaction picture do not change with time.

The higher orders fulfill

$$\hat{U}_I^{(n)}(t_f,t_i) = \hat{1} + \frac{1}{i\hbar}\int_{t_i}^{t_f} dt'\, \hat{W}_I(t')\hat{U}_I^{(n-1)}(t',t_i)$$

which is obtained by removing all terms that are of higher order than $\hat{W}^n$. This equation allows us to determine the propagator for the perturbed system to all orders by an iterative procedure.

Let us write down the first few terms, in order to understand how the terms look like.

1. first order

$$\begin{aligned}\hat{U}_I^{(1)}(t_f,t_i) &= \hat{1} + \frac{1}{i\hbar}\int_{t_i}^{t_f} dt'\, \hat{W}_I(t')\underbrace{\hat{U}_I^{(0)}(t',t_i)}_{=\hat{1}}\\ &= \hat{1} + \frac{1}{i\hbar}\int_{t_i}^{t_f} dt'\, \hat{W}_I(t')\end{aligned}$$

2. second order

$$\begin{aligned}\hat{U}_I^{(2)}(t_f,t_i) &= \hat{1} + \frac{1}{i\hbar}\int_{t_i}^{t_f} dt'\, \hat{W}_I(t')\hat{U}_I^{(1)}(t',t_i)\\ &= \hat{1} + \frac{1}{i\hbar}\int_{t_i}^{t_f} dt'\, \hat{W}_I(t') + \frac{1}{i\hbar}\int_{t_i}^{t_f} dt'\, \hat{W}_I(t')\frac{1}{i\hbar}\int_{t_i}^{t'} dt''\, \hat{W}_I(t'')\\ &= \hat{1} + \frac{1}{i\hbar}\int_{t_i}^{t_f} dt'\, \hat{W}_I(t') + \left(\frac{1}{i\hbar}\right)^2\int_{t_i}^{t_f} dt'\int_{t_i}^{t'} dt''\, \hat{W}_I(t')\hat{W}_I(t'')\end{aligned}$$

We can already generalize the building principle and construct the higher orders

$$\hat{U}_I(t_f,t_i) = \sum_{n=0}^{\infty}\left(\frac{1}{i\hbar}\right)^n \int_{t_i}^{t_f} dt_1 \int_{t_i}^{t_1} dt_2 \ldots \int_{t_i}^{t_{n-1}} dt_n\, \hat{W}_I(t_1)\cdots\hat{W}_I(t_n) \tag{11.13}$$

Note that the interactions are occur such that the earlier interaction stands right from the later interaction. Thus, the time increases from right to left.

Let us now introduce the **time-ordered product** of operators.

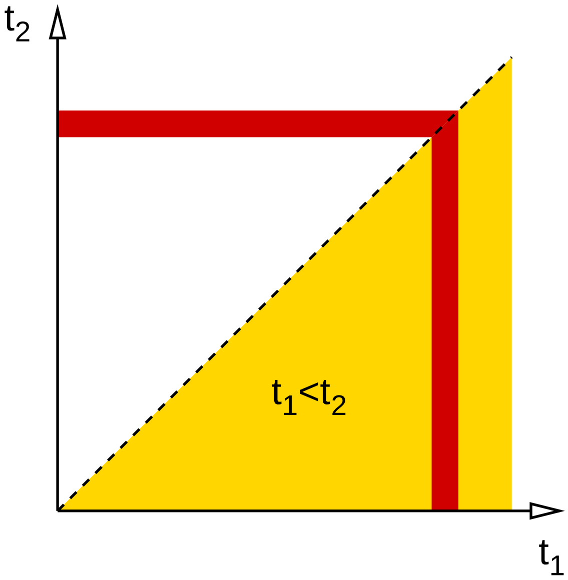
TIME-ORDERED PRODUCT

$$T[\hat{A}(t),\hat{B}(t')]\stackrel{\text{def}}{=}\left\{\begin{array}{lll}\hat{A}(t)\hat{B}(t') & \text{for} & t\geq t' \\ \hat{B}(t')\hat{A}(t) & \text{for} & t'>t\end{array}\right. \qquad (11.14)$$

The time-ordered product of many operators is obtained by simply placing the operators such that the time increases from right to left.

The double integral can be written with the help of the time ordered product as

$$\begin{aligned}
&\int_{t_i}^{t_f} dt_1 \int_{t_0}^{t_1} dt_2\, \hat{W}_I(t_1)\hat{W}_I(t_2)\\
=\;&\int_{t_i}^{t_f} dt_1 \int_{t_0}^{t_f} dt_2\, \hat{W}_I(t_1)\hat{W}_I(t_2)\theta(t_2-t_1)\\
\stackrel{t_1\leftrightarrow t_2}{=}\;&\int_{t_0}^{t_f} dt_1 \int_{t_0}^{t_f} dt_2\, \frac{1}{2}\left(\hat{W}_I(t_1)\hat{W}_I(t_2)\theta(t_2-t_1)+\hat{W}_I(t_2)\hat{W}_I(t_1)\theta(t_1-t_2)\right)\\
=\;&\frac{1}{2}\int_{t_i}^{t_f} dt_1 \int_{t_i}^{t_f} dt_2\, T[\hat{W}_I(t_1)\hat{W}_I(t_2)]
\end{aligned}$$



If we extend integration bounds in all an n-th order order terms, we obtain for each time ordered product in total $n!$ permutations of the operators. Thus, we need to divide the result by $n!$. Thus, we obtain

$$\hat{U}_I(t_f,t_i) = \sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{1}{i\hbar}\right)^n\int_{t_i}^{t_f}dt_1\ldots\int_{t_i}^{t_f}dt_n\,T[\hat{W}_I(t_1)\cdots\hat{W}_I(t_n)] \tag{11.15}$$

$$= T\left[\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{1}{i\hbar}\right)^n\int_{t_i}^{t_f}dt_1\ldots\int_{t_i}^{t_f}dt_n\,\hat{W}_I(t_1)\cdots\hat{W}_I(t_n)\right]$$

$$= T\left[\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{1}{i\hbar}\int_{t_i}^{t_f}dt'\,\hat{W}_I(t')\right)^n\right]$$

$$= T\left[\mathrm{e}^{\frac{1}{i\hbar}\int_{t_i}^{t_f}dt'\,\hat{W}_I(t')}\right] \tag{11.16}$$

Fertig: WS06/07 21 Doppelstunde 8.Feb.07

11.3 Green's function as propagator

The Green's function is closely related to the propagator. Let us remember the equation Eq. 10.6 that defined the Green's function.

$$\left(i\hbar\partial_t - \hat{H}\right)\hat{G}(t,t') = \hat{1}\delta(t-t')$$

The propagator fulfills

$$\left(i\hbar\partial_t - \hat{H}\right)\hat{U}(t,t') = \hat{0}$$

This suggests that we can express the Green's function by the propagator as

GREEN'S FUNCTION AND PROPAGATOR

$$\hat{G}(t,t') = \frac{1}{i\hbar}\theta(t-t')\hat{U}(t,t') \tag{11.17}$$

Thus, we have a physical interpretation of the Green's function:

INTERPRETATION OF THE GREEN'S FUNCTION

The Green's function is related to the probability density $P(x,t,x',t')$ of finding a particle at $\vec{x}$ at time t, if it has been at time t' at position $\vec{x'}$.

$$P(x,t,x',t') = |i\hbar G(x,t,x',t')|^2 \tag{11.18}$$

Note however that this simplistic sentence does not capture that the Green's function contains a phase information in addition to the probability.
More accurately, the Green's function propagates a wave function $|\psi_0\rangle$ from time t' to its new state at time t, that is

$$|\psi(t)\rangle = i\hbar\hat{G}(t,t')|\Psi_0\rangle \qquad \text{if} \qquad |\psi(t')\rangle = |\psi_0\rangle \qquad \text{and} \qquad t > t'$$

Let us now return to Eq. 11.13 and determine the propagator in the regular Schrödinger picture. This will show that the Green's function plays a central role for the perturbative treatment of any quantity.

$$\hat{U}_I(t_f,t_i)=\hat{U}^{(0)}(0,t_f)\hat{U}(t_f,t_i)\hat{U}^{(0)}(t_i,0) \qquad \Rightarrow \qquad \hat{U}(t_f,t_i)=\hat{U}^{(0)}(t_f,0)\hat{U}_I(t_f,t_i)\hat{U}^{(0)}(0,t_i)$$

$$\begin{aligned}
\hat{G}(t_f,t_i) &= \frac{1}{i\hbar}\theta(t_f-t_i)\hat{U}^{(0)}(t_f,0)\hat{U}_I(t_f,t_i)\hat{U}^{(0)}(0,t_i) \\
&\overset{\text{Eq. 11.13}}{=} \frac{1}{i\hbar}\theta(t_f-t_i)\hat{U}^{(0)}(t_f,0) \\
&\qquad \cdot \underbrace{\sum_{n=0}^{\infty}\left(\frac{1}{i\hbar}\right)^n \int_{t_i}^{t_f} dt_1 \int_{t_i}^{t_1} dt_2 \ldots \int_{t_i}^{t_{n-1}} dt_n\, \hat{W}_I(t_1)\cdots\hat{W}_I(t_n)}_{\hat{U}_I(t_f,t_i)}\,\hat{U}^{(0)}(0,t_i)
\end{aligned}$$

Now, we replace the integration bounds

$$\begin{aligned}
&\theta(t_f-t_i)\int_{t_i}^{t_f} dt_1 \int_{t_i}^{t_1} dt_2 \cdots \int_{t_i}^{t_{n-1}} dt_n \\
&= \int_{-\infty}^{\infty} dt_1 \cdots \int_{-\infty}^{\infty} dt_n\; \theta(t_f-t_1)\theta(t_1-t_2)\cdots\theta(t_{n-1}-t_n)\theta(t_n-t_i)
\end{aligned} \tag{11.19}$$

Later we will use the form of the Green's function from Eq. 11.17 in slightly modified form

$$i\hbar\hat{G}(t,t') \overset{\text{Eq. 11.17}}{=} \theta(t-t)\hat{U}(t,t') = \hat{U}(t,0)\theta(t-t)\hat{U}(0,t') \tag{11.20}$$

With these considerations in mind, we proceed

$$\begin{aligned}
\hat{G}(t_f,t_i) &\overset{\text{Eq. 11.19}}{=} \sum_{n=0}^{\infty}\left(\frac{1}{i\hbar}\right)^{n+1} \int_{-\infty}^{\infty} dt_1 \cdots \int_{-\infty}^{\infty} dt_n\; \theta(t_f-t_1)\cdots\theta(t_n-t_i) \\
&\quad \cdot \hat{U}^{(0)}(t_f,0)\underbrace{\hat{U}^{(0)}(0,t_1)\hat{W}_S(t_1)\hat{U}^{(0)}(t_1,0)}_{\hat{W}_I(t_1)}\cdots\underbrace{\hat{U}^{(0)}(0,t_n)\hat{W}_S(t_1)\hat{U}^{(0)}(t_n,0)}_{\hat{W}_I(t_n)}\hat{U}^{(0)}(0,t_i) \\
&= \sum_{n=0}^{\infty}\left(\frac{1}{i\hbar}\right)^{n+1} \int_{-\infty}^{\infty} dt_1 \cdots \int_{-\infty}^{\infty} dt_n \\
&\qquad \cdot \underbrace{\hat{U}^{(0)}(t_f,0)\theta(t_f-t_1)\hat{U}^{(0)}(0,t_1)}_{i\hbar\hat{G}^{(0)}(t_f,t_1)}\hat{W}_S(t_1)\hat{U}^{(0)}(t_1,0)\cdots \\
&\qquad \cdots\hat{U}^{(0)}(0,t_n)\hat{W}_S(t_1)\underbrace{\hat{U}^{(0)}(t_n,0)\theta(t_n-t_i)\hat{U}^{(0)}(0,t_i)}_{i\hbar\hat{G}^{(0)}(t_n,t_i)} \\
&\overset{\text{Eq. 11.20}}{=} \sum_{n=0}^{\infty}\int_{-\infty}^{\infty} dt_1 \cdots \int_{-\infty}^{\infty} dt_n\; \hat{G}^{(0)}(t_f,t_1)\hat{W}_S(t_1)\hat{G}^{(0)}(t_1,t_2)\hat{W}_S(t_2)\hat{G}^{(0)}(t_2,t_3) \\
&\qquad \cdots\hat{G}^{(0)}(t_{n-1},t_n)\hat{W}_S(t_n)\hat{G}^{(0)}(t_n,t_i)
\end{aligned}$$

The Green's function came into play via the integration bounds for the time integration and the conversion from the interaction picture to the Schrödinger picture.

We obtain realizing that our expression describes the Green's function for the perturbed system

$$\begin{aligned}
\hat{G}(t_f,t_i) = &\sum_{n=0}^{\infty}\int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2 \ldots \int_{-\infty}^{\infty} dt_n \\
&\cdot \hat{G}^{(0)}(t_f,t_1)\hat{W}(t_1)\hat{G}^{(0)}(t_1,t_2)\hat{W}(t_2)\cdots\hat{W}(t_n)\hat{G}^{(0)}(t_n,t_i)
\end{aligned}$$

Note that this is identical to Eq. 11.2 after we Fourier transform from a time dependence to an energy dependence. Remember that a convolution is converted by a Fourier transform into a simple product.

11.4 Retarded Potentials

This is just a sketch of an idea and not intended for the reader....

Consider a Schrödinger equation that has a nonlocal potential not only in space but also in time. That is it would have a retarded potential.

$$\left[i\hbar\partial_t - \hat{H}_0\right]|\psi(t)\rangle - \int dt'\hat{v}(t-t')|\psi(t')\rangle = 0$$

Let us make an Ansatz

$$|\psi(t)\rangle = \int d\epsilon|\psi(\epsilon)\rangle \mathrm{e}^{-\frac{i}{\hbar}\epsilon t}$$

and insert it into the above equation.

$$\begin{aligned}
\int d\epsilon\ \left[i\hbar\partial_t - \hat{H}_0\right]|\psi(\epsilon)\rangle \mathrm{e}^{-\frac{i}{\hbar}\epsilon t} - \int d\epsilon \int dt'\ \hat{v}(t-t')|\psi(\epsilon)\rangle \mathrm{e}^{-\frac{i}{\hbar}\epsilon t'} &= 0\\
\int d\epsilon\ \left[\epsilon - \hat{H}_0\right]|\psi(\epsilon)\rangle \mathrm{e}^{-\frac{i}{\hbar}\epsilon t} - \int d\epsilon\ \underbrace{\left(\int dt'\ \hat{v}(t-t')\mathrm{e}^{\frac{i}{\hbar}\epsilon(t-t')}\right)}_{=:\hat{v}(\epsilon)}|\psi(\epsilon)\rangle \mathrm{e}^{-\frac{i}{\hbar}\epsilon t} &= 0\\
\int d\epsilon\ \left[\epsilon - \hat{H}_0 - \hat{v}(\epsilon)\right]|\psi(\epsilon)\rangle \mathrm{e}^{-\frac{i}{\hbar}\epsilon t} &= 0
\end{aligned}$$

Thus, we obtain en energy dependent potential. If the potential were local in time, its Fourier transform would be a constant in ϵ and therefore the potential would be energy independent.

In order to solve the above equation we must fulfill

$$\left[\epsilon - \hat{H}_0 - v(\epsilon)\right]|\psi(\epsilon)\rangle = 0$$

Note however, that the energies are, in general, complex and that the wave functions are not necessarily orthonormal. (I do not have a proof either way.)

Questions: Compare this to the Lehmann representation. Can one identify the wave functions with those of the Lehmann representation

$$\phi_n(x) \overset{?}{=} \langle\mathcal{G}|\hat{\psi}(x)|\Psi_n\rangle$$

Chapter 12

Green's functions in many-particle physics

12.1 Heisenberg picture

There are three pictures to represent the dynamics of expectation values. The Schrödinger picture, which is the one we are used to, the interaction picture that we have introduced previously (see p. 182) and the Heisenberg picture.

In the **Schrödinger picture**, the wave function is time dependent, but the operators are normally time independent. The dynamics of the system is described by the Schrödinger equation

$$i\hbar\partial_t|\Psi_S(t)\rangle = \hat{H}|\Psi_S(t)\rangle$$

The expectation value of an operator $\hat{A}_S(t)$ has the form

$$\langle A\rangle = \langle\Psi(t)|\hat{A}_S(t)|\Psi(t)\rangle$$

There is no dynamical equation for the operators. The time dependence of a Schrödinger operator may, for example, describe the time dependence of the Hamilton operator due to external interactions.

Now we use again the propagator $\hat{U}(t,0)$

$$|\Psi_S(t)\rangle = \hat{U}(t,0)|\Psi_S(0)\rangle$$

Using the Schrödinger equation we obtain an equation for $\hat{U}(t,0)$

$$\begin{aligned}
i\hbar\partial_t\hat{U}(t,0)|\psi_0\rangle &= \hat{H}_S(t)\hat{U}(t,0)|\psi_0\rangle \\
\Rightarrow\qquad i\hbar\partial_t\hat{U}(t,0) &= \hat{H}_S(t)\hat{U}(t,0) &\quad (12.1)\\
i\hbar\partial_t\hat{U}(0,t) &= i\hbar\partial_t\hat{U}^\dagger(t,0) = \left(-i\hbar\partial_t\hat{U}^\dagger(t,0)\right) = -\left(\hat{H}_S(t)\hat{U}(t,0)\right)^\dagger \\
&= -\hat{U}(t,0)^\dagger\hat{H}_S(t)^\dagger = -\hat{U}(0,t)\hat{H}_S(t) &\quad (12.2)
\end{aligned}$$

A state $|\Psi_H(t)\rangle$ in the **Heisenberg picture** is related to one in the Schrödinger picture by the propagator $\hat{U}(t,0)$

$$|\Psi_S(t)\rangle = \hat{U}(t,0)|\Psi_H(t)\rangle \qquad (12.3)$$

We can easily see that the state in the Heisenberg picture is the Schrödinger state at $t=0$. Thus, it is time independent.

HEISENBERG STATE

$$|\Psi_H\rangle = |\Psi_S(t=0)\rangle$$

The dynamical equation for the Heisenberg states, which replaces the Schrödinger equation is

$$i\hbar\partial_t|\Psi_H\rangle = 0$$

If we want to express an expectation value as

$$\langle A\rangle = \langle\Psi(t)|\hat{A}_S(t)|\Psi(t)\rangle = \langle\Psi_H|\hat{A}_H(t)|\Psi_H\rangle$$

we need to introduce

HEISENBERG OPERATORS

$$\hat{A}_H(t) = \hat{U}(0,t)\hat{A}_S(t)\hat{U}(t,0) \tag{12.4}$$

There is a non-trivial equation of motion for the Heisenberg operator, which replaces the Schrödinger equation.

$$\begin{aligned}
i\hbar\partial_t\hat{A}_H(t) &= \left[i\hbar\partial_t\hat{U}(0,t)\right]\hat{A}_S(t)\hat{U}(t,0) + \hat{U}(0,t)\left[i\hbar\partial_t\hat{A}_S(t)\right]\hat{U}(t,0) + \hat{U}(0,t)\hat{A}_S(t)\left[i\hbar\partial_t\hat{U}(t,0)\right] \\
&= -\hat{U}(0,t)\hat{H}_S(t)\hat{A}_S(t)\hat{U}(t,0) + i\hbar\hat{U}(0,t)\left(\frac{\partial\hat{A}_S}{\partial t}\right)\hat{U}(t,0) + \hat{U}(0,t)\hat{A}_S(t)\hat{H}_S(t)\hat{U}(t,0) \\
&= \left([\hat{A}_S(t),\hat{H}_S(t)]_- + i\hbar\frac{\partial\hat{A}_S}{\partial t}\right)_H \\
&= [\hat{A}_H(t),\hat{H}_H(t)]_- + i\hbar\left(\frac{\partial\hat{A}_S}{\partial t}\right)_H
\end{aligned}$$

12.1.1 Pictures in comparison

$$\begin{aligned}
\left(i\hbar\partial_t - \hat{H}\right)|\psi_S\rangle &= 0 \\
\left(i\hbar\partial_t - \hat{W}\right)|\psi_I\rangle &= 0 \\
\left(i\hbar\partial_t\right)|\psi_H\rangle &= 0
\end{aligned}$$

$$\begin{aligned}
|\psi_S(t)\rangle &= \hat{1}|\psi_S(t)\rangle \\
|\psi_I(t)\rangle &= \hat{U}^{(0)}(0,t)|\psi_S(t)\rangle \\
|\psi_H(t)\rangle &= \hat{U}(0,t)|\psi_S(t)\rangle = |\psi_S(0)\rangle
\end{aligned}$$

$$\langle\hat{A}\rangle = \langle\psi_S(t)|\hat{A}_S(t)|\psi_S(t)\rangle = \langle\psi_I(t)|\hat{A}_I(t)|\psi_I(t)\rangle = \langle\psi_H(t)|\hat{A}_H(t)|\psi_I(t)\rangle$$

$$\begin{aligned}
\partial_t \hat{A}_S &= \frac{d\hat{A}_S}{dt} \\
\partial_t \hat{A}_I &= \left(\frac{d\hat{A}_S}{dt}\right)_I + \frac{i}{\hbar}[\hat{H}_I, \hat{A}_I]_- \\
\partial_t \hat{A}_H &= \left(\frac{d\hat{A}_S}{dt}\right)_H + \frac{i}{\hbar}[\hat{H}_H, \hat{A}_H]_-
\end{aligned}$$

$$\begin{aligned}
\hat{A}_I(t) &= \hat{U}^{(0)}(0,t)\hat{A}_S(t)\hat{U}^{(0)}(t,0) \\
\hat{A}_H(t) &= \hat{U}(0,t)\hat{A}_S(t)\hat{U}(t,0)
\end{aligned}$$

12.2 One-particle Green's function expressed by the many-particle Hamiltonian

Here, we want to generalize the result Eq. 10.8 from p. 173 for the Green's function in the Hilbert space (one-particle quantum mechanics)

$$G(\vec{x},t,\vec{x'},t') = \frac{1}{i\hbar}\theta(t-t')\sum_n \langle\vec{x}|\psi_n\rangle \mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}\langle\psi_n|\vec{x'}\rangle$$

to the Fock-space (many-particle quantum mechanics).

We start out with the definition of the one-particle Green's function Eq. 10.8 on p. 173. Here, we show the principle for a time-independent Hamiltonian. The generalization to a time-dependent Hamiltonian should not be too difficult.

We take Eq. 10.8 and rewrite it to make use of creation and annihilation operators. The Hamiltonian of a one-particle system can be written as

$$\hat{H} = \sum_n \epsilon_n \hat{c}_n^\dagger \hat{c}_n$$

where $\hat{c}_n^\dagger$ and $\hat{c}_n$ are the creation and annihilation operators in the representation of energy eigenstates. If we apply this Hamiltonian to a one-particle state $\hat{c}_n^\dagger|\mathcal{O}\rangle$, that is an eigenstate of the Hamiltonian with eigenvalue ϵ_n we obtain

$$\hat{H}\hat{c}_n^\dagger|\mathcal{O}\rangle = \hat{c}_n^\dagger|\mathcal{O}\rangle\epsilon_n \tag{12.5}$$

The Hamiltonian applied to a many-particle state gives

$$\hat{H}\hat{c}_{i_1}^\dagger\cdots\hat{c}_{i_N}^\dagger|\mathcal{O}\rangle = \hat{c}_{i_1}^\dagger\cdots\hat{c}_{i_N}^\dagger|\mathcal{O}\rangle\sum_{j=1}^N \epsilon_{i_j}$$

This Hamiltonian is identical to the one-particle Hamiltonian $\hat{H} = \sum_n |\psi_n\rangle\epsilon_n\langle\psi_n| = \sum_n \hat{c}_n^\dagger|\mathcal{O}\rangle\epsilon_n\langle\mathcal{O}|\hat{c}_n$ in the one-particle Hilbert space. For two or more particles, however, it differs from the one-particle Hamiltonian, because the form $\sum_n |\psi_n\rangle\epsilon_n\langle\psi_n|$ can only act on one-particle states.

Let us now turn to the one-particle Green's function Eq. 10.8

$$
\begin{aligned}
G(\vec{x},t,\vec{x},t') &\overset{Eq.\ 10.8}{=} \langle\vec{x}|\left[\frac{1}{i\hbar}\theta(t-t')\sum_n|\psi_n\rangle e^{-\frac{i}{\hbar}\epsilon_n(t-t')}\langle\psi_n|\right]|\vec{x'}\rangle \\
&\overset{1}{=} \frac{1}{i\hbar}\theta(t-t')\sum_{n,m}\psi_n(\vec{x})\,\overbrace{\langle\mathcal{O}|\hat{c}_n e^{-\frac{i}{\hbar}\hat{H}(t-t')}\hat{c}_m^\dagger|\mathcal{O}\rangle}^{\delta_{n,m}e^{-\frac{i}{\hbar}\epsilon_n(t-t')}}\,\psi_m^*(\vec{x'}) \\
&\overset{2}{=} \frac{1}{i\hbar}\theta(t-t')\sum_{n,m}\psi_n(\vec{x})\langle\mathcal{O}|\underbrace{e^{\frac{i}{\hbar}\hat{H}t}\hat{c}_n e^{-\frac{i}{\hbar}\hat{H}t}}_{\hat{c}_{H,n}(t)}\underbrace{e^{\frac{i}{\hbar}\hat{H}t'}\hat{c}_m^\dagger e^{-\frac{i}{\hbar}\hat{H}t'}}_{\hat{c}_{H,m}^\dagger(t')}|\mathcal{O}\rangle\psi_m^*(\vec{x'}) \\
&= \frac{1}{i\hbar}\theta(t-t')\sum_{n,m}\psi_n(\vec{x})\langle\mathcal{O}|\hat{c}_{H,n}(t)\hat{c}_{H,m}^\dagger(t')|\mathcal{O}\rangle\psi_m^*(\vec{x'}) \\
&\overset{Eq.\ 8.11}{=} \frac{1}{i\hbar}\theta(t-t')\langle\mathcal{O}|\hat{\psi}_H(\vec{x},t)\hat{\psi}_H^\dagger(\vec{x},t')|\mathcal{O}\rangle
\end{aligned}
\tag{12.6}
$$

- In step 1 (I refer to the number on top of the equal sign) we made use of the fact that $\hat{c}_m^\dagger|\mathcal{O}\rangle$ is a one-particle state with quantum number m. This state is an eigenstate of the Hamilton operator $\hat{H}=\sum_n\epsilon_n\hat{c}_n^\dagger\hat{c}_n$. We also exploited the orthonormality of the one-particle states, namely $\langle\mathcal{O}|\hat{c}_n^\dagger\hat{c}_m|\mathcal{O}\rangle=\langle\psi_n|\psi_m\rangle=\delta_{n,m}$.
- In step 2 we made use that the Hamiltonian of Eq. 12.5 applied to the vacuum state $|\mathcal{O}\rangle$ gives the zero state $|\varnothing\rangle$, that is $\hat{H}|\mathcal{O}\rangle=|\varnothing\rangle$. As a consequence, we obtain

  $$e^{\frac{i}{\hbar}\hat{H}t}|\mathcal{O}\rangle=\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{it}{\hbar}t\right)^n\hat{H}^n|\mathcal{O}\rangle=|\mathcal{O}\rangle+\sum_{n=1}^{\infty}|\varnothing\rangle=|\mathcal{O}\rangle$$

  which is then exploited in step 2.

The Green's function can also be expressed in a representation of energy eigenstates as

$$
\begin{aligned}
G(\vec{x},t,\vec{x'},t') &= \sum_{n,m}\langle\vec{x}|\psi_n\rangle G_{n,m}(t,t')\langle\psi_m|\vec{x'}\rangle \\
G_{n,m}(t,t') &= \frac{1}{i\hbar}\theta(t-t')\langle\mathcal{O}|\hat{c}_{H,n}(t)\hat{c}_{H,m}^\dagger(t')|\mathcal{O}\rangle
\end{aligned}
\tag{12.7}
$$

The physical meaning of the Green's function is completely analogous to that of the one-particle Green's function: It is the probability amplitude for a particle, which has been created at time t' at position x', to turn up at time t at the position x.

12.3 Many-particle Green's function

We define the Green's function analogously to Eq. 12.6. However, now the vacuum state is replaced by the vacuum state for electrons and holes, which is the ground state $|\mathcal{G}\rangle$ of the electron gas. The Green's function is then generalized to electron and hole excitations.

$$
\begin{aligned}
G_{n,m}(t,t') &= \frac{1}{i\hbar}\Big[\theta(t-t')\langle\mathcal{G}|\hat{e}_{H,n}(t)\hat{e}_{H,m}^\dagger(t')|\mathcal{G}\rangle-\theta(t'-t)\langle\mathcal{G}|\hat{h}_{H,m}(t')\hat{h}_{H,n}^\dagger(t)|\mathcal{G}\rangle\Big] \\
&\overset{Eq.\ ??}{=} \frac{1}{i\hbar}\Big[\theta(t-t')\langle\mathcal{G}|\hat{c}_{H,n}(t)\hat{c}_{H,m}^\dagger(t')|\mathcal{G}\rangle-\theta(t'-t)\langle\mathcal{G}|\hat{c}_{H,m}^\dagger(t')\hat{c}_{H,n}(t)|\mathcal{G}\rangle\Big]
\end{aligned}
$$

We simply mapped the electron and the hole Green's function together so that we obtain the electron propagator for $t>t'$ and the hole propagator for $t<t'$. The physical picture is that a hole propagating forward in time is identical to an electron propagating backward in time[63, 64].

At first sight the negative sign in front of the hole contribution to the Green's function seems arbitrary. As we will see later, the negative sign is required because the non-interacting Green's function shall still obey the inhomogeneous differential equation Eq. 12.23.

This expression of the Green's function can be written in a more compact form using the **time-ordered product** T. The time ordered product orders the field operators from the right to the left in ascending time. For each interchange of two Fermionic operators the result obtains a factor -1.

WICK'S TIME-ORDERED PRODUCT OF FERMION OPERATORS

For two Fermion operators $\hat{A}$, $\hat{B}$ the time-ordered product is

$$T[\hat{A}(t)\hat{B}(t')] = \begin{cases} \hat{A}(t)\hat{B}(t') & \text{for} \quad t \geq t' \\ -\hat{B}(t')\hat{A}(t) & \text{for} \quad t < t' \end{cases}$$

A Fermion operator is either a Fermionic[a] creation or annihilation operator.[b]
T is **Wick's time-ordering operator** , which acts on all operators on the right of it. In general Wicks time-ordering operator reorders a product so that the time arguments ascent from right to left and it applies a sign change or every permutation of two fermion operators.

---

[a] For bosonic creation and annihilation operators the minus sign would be replaced by a plus sign.

[b] In the previous definition, Eq. 11.14, the minus sign was missing. The reason is that it was specific to operators, that contain products of an even number of Fermion operators. With this limitation the two definitions are identical.

We have introduced the time-ordered product earlier Eq. 11.14. In that case we did not mention the sign change. The reason is that in Eq. 11.14 the operators corresponded to even products of fermionic operators, so that the sign changes canceled.

Thus, we arrive at the final definition of the **Green's function** as

$$G(\vec{x}, t, \vec{x'}, t') = \frac{1}{i\hbar} \frac{\langle \mathcal{G}|T[\hat{\psi}_H(\vec{x}, t)\hat{\psi}^\dagger_H(\vec{x'}, t')]|\mathcal{G}\rangle}{\langle \mathcal{G}|\mathcal{G}\rangle} \tag{12.8}$$

where the field operators are represented in the Heisenberg picture and $\mathcal{G}$ is the ground state of the interacting electron gas.

We assumed here that the ground state is normalized. i.e. $\langle \mathcal{G}|\mathcal{G}\rangle = 1$.

12.4 Differential equation for the many-particle Green's function

The Green's function of a many-particle system is not a Green's function in the mathematical sense. In mathematics, a Green's function is the inverse of a differential operator. This applies only to the Green function of one-particle systems. In solid state physics, only the physical meaning of the Green's function has been kept, namely as probability amplitude between two events.

When we try to set up a differential equation analogous to the one-particle case, we are lead to higher Green's functions involving more and more particles. Nevertheless, we can define a retarded one-particle potential that replaces the action of the many-particle system. This potential is the **self energy**. If the self-energy is included, the Green's function obeys a differential equation as in the one-particle case.

In the following we derive a differential equation for the interacting many-body Green's function.

The first step is to find a differential equation for a Heisenberg operator $\hat{c}_\alpha(t)$, because the time derivative will act on this operator within the expression for the Green's function

$$i\hbar\partial_t \hat{c}_\zeta(t) = \mathrm{e}^{\frac{i}{\hbar}\hat{H}t}\Big[\hat{c}_\zeta, \hat{H}\Big]_- \mathrm{e}^{-\frac{i}{\hbar}\hat{H}t} \tag{12.9}$$

In order to continue, we need the commutator of the annihilation operator in the Schrödinger picture with the Hamiltonian.

$$\hat{H}=\sum_{\alpha,\beta}h_{\alpha,\beta}\hat{c}^\dagger_\alpha\hat{c}_\beta+\frac{1}{2}\sum_{\alpha,\beta,\gamma,\delta}W_{\alpha,\beta,\delta,\gamma}\hat{c}^\dagger_\alpha\hat{c}^\dagger_\beta\hat{c}_\gamma\hat{c}_\delta \tag{12.10}$$

We begin with the one-particle operator

$$\begin{aligned}\Big[\hat{c}_\zeta,\sum_{\alpha,\beta}h_{\alpha,\beta}\hat{c}^\dagger_\alpha\hat{c}_\beta\Big]_-&=\sum_{\alpha,\beta}h_{\alpha,\beta}\Big(\underbrace{\hat{c}_\zeta\hat{c}^\dagger_\alpha}_{\delta_{\zeta,\alpha}-\hat{c}^\dagger_\alpha\hat{c}_\zeta}\hat{c}_\beta-\hat{c}^\dagger_\alpha\underbrace{\hat{c}_\beta\hat{c}_\zeta}_{-\hat{c}_\zeta\hat{c}_\beta}\Big)=\sum_{\alpha,\beta}h_{\alpha,\beta}\delta_{\zeta,\alpha}\hat{c}_\beta\\&=\sum_\beta h_{\zeta,\beta}\hat{c}_\beta\end{aligned} \tag{12.11}$$

and continue with the interaction

$$\begin{aligned}&\Big[\hat{c}_\zeta,\frac{1}{2}\sum_{\alpha,\beta,\gamma,\delta}W_{\alpha,\beta,\delta,\gamma}\hat{c}^\dagger_\alpha\hat{c}^\dagger_\beta\hat{c}_\gamma\hat{c}_\delta\Big]_-\\&\quad=\frac{1}{2}\sum_{\alpha,\beta,\gamma,\delta}W_{\alpha,\beta,\delta,\gamma}\Big(\hat{c}_\zeta\hat{c}^\dagger_\alpha\hat{c}^\dagger_\beta\hat{c}_\gamma\hat{c}_\delta\underbrace{-\hat{c}^\dagger_\alpha\hat{c}^\dagger_\beta\hat{c}_\zeta\hat{c}_\gamma\hat{c}_\delta+\hat{c}^\dagger_\alpha\hat{c}^\dagger_\beta\hat{c}_\zeta\hat{c}_\gamma\hat{c}_\delta}_{=0}-\hat{c}^\dagger_\alpha\hat{c}^\dagger_\beta\hat{c}_\gamma\hat{c}_\delta\hat{c}_\zeta\Big)\\&\quad=\frac{1}{2}\sum_{\alpha,\beta,\gamma,\delta}W_{\alpha,\beta,\delta,\gamma}\Big(\Big(\underbrace{\hat{c}_\zeta\hat{c}^\dagger_\alpha}_{\delta_{\zeta,\alpha}-\hat{c}^\dagger_\alpha\hat{c}_\zeta}\hat{c}^\dagger_\beta-\hat{c}^\dagger_\alpha\underbrace{\hat{c}^\dagger_\beta\hat{c}_\zeta}_{\delta_{\beta,\zeta}-\hat{c}_\zeta\hat{c}^\dagger_\beta}\Big)\hat{c}_\gamma\hat{c}_\delta+\hat{c}^\dagger_\alpha\hat{c}^\dagger_\beta\Big(\underbrace{\hat{c}_\zeta\hat{c}_\gamma}_{-\hat{c}_\gamma\hat{c}_\zeta}\hat{c}_\delta-\hat{c}_\gamma\underbrace{\hat{c}_\delta\hat{c}_\zeta}_{-\hat{c}_\zeta\hat{c}_\delta}\Big)\Big)\\&\quad=\frac{1}{2}\sum_{\alpha,\beta,\gamma,\delta}W_{\alpha,\beta,\delta,\gamma}\Big(\delta_{\zeta,\alpha}\hat{c}^\dagger_\beta-\hat{c}^\dagger_\alpha\delta_{\beta,\zeta}\Big)\hat{c}_\gamma\hat{c}_\delta\\&\quad=\frac{1}{2}\sum_{\beta,\gamma,\delta}W_{\zeta,\beta,\delta,\gamma}\hat{c}^\dagger_\beta\hat{c}_\gamma\hat{c}_\delta-\sum_{\alpha,\gamma,\delta}W_{\alpha,\zeta,\delta,\gamma}\hat{c}^\dagger_\alpha\hat{c}_\gamma\hat{c}_\delta\\&\quad=\frac{1}{2}\sum_{\beta,\gamma,\delta}\Big(W_{\zeta,\beta,\delta,\gamma}-W_{\beta,\zeta,\delta,\gamma}\Big)\hat{c}^\dagger_\beta\hat{c}_\gamma\hat{c}_\delta\\&\quad=\sum_{\beta,\gamma,\delta}W_{\zeta,\beta,\delta,\gamma}\hat{c}^\dagger_\beta\hat{c}_\gamma\hat{c}_\delta\end{aligned} \tag{12.12}$$

With the results Eqs. 12.11,12.12 we can return to Eq. 12.9 and evaluate the time derivative of the annihilator in the Heisenberg picture as

$$\begin{aligned}i\hbar\partial_t\hat{c}_\zeta(t)&=\mathrm{e}^{\frac{i}{\hbar}\hat{H}t}\Big[\hat{c}_\zeta,\hat{H}\Big]_-\mathrm{e}^{-\frac{i}{\hbar}\hat{H}t}\\&\overset{Eqs.\ 12.11,12.12}{=}\mathrm{e}^{\frac{i}{\hbar}\hat{H}t}\Big(\sum_\beta h_{\zeta,\beta}\hat{c}_\beta+\sum_{\gamma,\delta,\rho}W_{\zeta,\beta,\delta,\gamma}\hat{c}^\dagger_\beta\hat{c}_\gamma\hat{c}_\delta\Big)\mathrm{e}^{-\frac{i}{\hbar}\hat{H}t}\\&=\sum_\beta h_{\zeta,\beta}\hat{c}_\beta(t)+\sum_{\beta,\gamma,\delta}W_{\zeta,\beta,\delta,\gamma}\hat{c}^\dagger_\beta(t)\hat{c}_\gamma(t)\hat{c}_\delta(t)\end{aligned} \tag{12.13}$$

Now we can evaluate the time derivative of the Green's function

$$
\begin{aligned}
i\hbar\partial_t G_{\alpha,\beta}(t,t')= &\quad i\hbar\partial_t\left[\frac{\theta(\vec{t}-\vec{t'})}{i\hbar}\langle\hat{c}_\alpha(t)\hat{c}^\dagger_\beta(t')\rangle-\frac{\theta(t'-t)}{i\hbar}\langle\hat{c}^\dagger_\beta(t')\hat{c}_\alpha(t)\rangle\right]\\
= &\quad \delta(\vec{t}-\vec{t'})\langle\hat{c}_\alpha(t)\hat{c}^\dagger_\beta(t')\rangle+\delta(t'-t)\langle\hat{c}^\dagger_\beta(t')\hat{c}_\alpha(t)\rangle\\
+ &\quad \frac{\theta(t-t')}{i\hbar}\langle\Big(i\hbar\partial_t\hat{c}_\alpha(t)\Big)\hat{c}^\dagger_\beta(t')\rangle-\frac{\theta(t'-t)}{i\hbar}\langle\hat{c}^\dagger_\beta(t')\Big(i\hbar\partial_t\hat{c}_\alpha(t)\Big)\rangle\\
\overset{\text{Eq. 12.13}}{=} &\quad \delta(\vec{t}-\vec{t'})\langle\Big[\hat{c}_\alpha(t),\hat{c}^\dagger_\beta(t')\Big]_+\rangle\\
+ &\quad \frac{\theta(t-t')}{i\hbar}\langle\Big(\sum_b h_{\alpha,b}\hat{c}_b(t)+\sum_{b,c,d}W_{\alpha,b,d,c}\hat{c}^\dagger_b(t)\hat{c}_c(t)\hat{c}_d(t)\Big)\hat{c}^\dagger_\beta(t')\rangle\\
- &\quad \frac{\theta(t'-t)}{i\hbar}\langle\hat{c}^\dagger_\beta(t')\Big(\sum_{\beta'} h_{\alpha,b}\hat{c}_b(t)+\sum_{b,c,d}W_{\alpha,b,d,c}\hat{c}^\dagger_b(t)\hat{c}_c(t)\hat{c}_d(t)\Big)\rangle\\
= &\quad \delta(t-t')\langle\Big[\hat{c}_\alpha(t),\hat{c}^\dagger_\beta(t')\Big]_+\rangle\\
+ &\quad \sum_b h_{\alpha,b}\Big(\frac{\theta(t-t')}{i\hbar}\langle\hat{c}_b(t)\hat{c}^\dagger_\beta(t')\rangle-\frac{\theta(t'-t)}{i\hbar}\langle\hat{c}^\dagger_\beta(t')\hat{c}_b(t)\rangle\Big)\\
+ &\quad \sum_{b,c,d}W_{\alpha,b,d,c}\Big(\frac{\theta(t-t')}{i\hbar}\langle\hat{c}^\dagger_b(t)\hat{c}_c(t)\hat{c}_d(t)\hat{c}^\dagger_\beta(t')\rangle-\frac{\theta(t'-t)}{i\hbar}\langle\hat{c}^\dagger_\beta(t')\hat{c}^\dagger_b(t)\hat{c}_c(t)\hat{c}_d(t)\rangle\Big)\\
\overset{\text{Eq. 12.15}}{=} &\quad \delta(\vec{t}-\vec{t'})\delta_{\alpha,\beta}+\sum_b h_{\alpha,b}G_{b,\beta}(t,t')\\
+ &\quad \sum_{b,c,d}W_{\alpha,b,d,c}\underbrace{\Big(\frac{\theta(t-t')}{i\hbar}\langle\overbrace{\hat{c}^\dagger_b(t)\hat{c}_c(t)\hat{c}_d(t)}\,\hat{c}^\dagger_\beta(t')\rangle-\frac{\theta(t'-t)}{i\hbar}\langle\hat{c}^\dagger_\beta(t')\,\overbrace{\hat{c}^\dagger_b(t)\hat{c}_c(t)\hat{c}_d(t)}\rangle\Big)}_{\frac{1}{i\hbar}\langle\mathcal{T}\hat{c}^\dagger_b(t)\hat{c}_c(t)\hat{c}_d(t)\hat{c}^\dagger_\beta(t')\rangle}
\end{aligned}
\tag{12.14}
$$

The anticommutator relations for the Schrödinger operators do not hold for Heisenberg operators. Nevertheless, we have made use use of the fact that the anticommutator is only required for equal times

$$
\Big[\hat{c}^\dagger_\beta(t),\hat{c}_\alpha(t)\Big]_+=\mathrm{e}^{\frac{i}{\hbar}\hat{H}t}\Big[\hat{c}^\dagger_\beta,\hat{c}_\alpha\Big]_+\mathrm{e}^{-\frac{i}{\hbar}\hat{H}t'}=\mathrm{e}^{\frac{i}{\hbar}\hat{H}t}\delta_{\alpha,\beta}\mathrm{e}^{-\frac{i}{\hbar}\hat{H}t'}=\delta_{\alpha,\beta}
\tag{12.15}
$$

Eq. 12.14 can be brought into a form similar to the defining equation of the Green's function for non-interacting systems is

QUASI EQUATION OF MOTION FOR THE INTERACTING GREEN'S FUNCTION

$$
\sum_\gamma\Big(\delta_{\alpha,\gamma}i\hbar\partial_t-h_{\alpha,\gamma}\Big)G_{\gamma,\beta}(t,t')=\delta(t-t')\delta_{\alpha,\beta}+\sum_{b,c,d}W_{\alpha,b,d,c}\frac{1}{i\hbar}\langle\mathcal{T}\hat{c}^\dagger_b(t)\hat{c}_c(t)\hat{c}_d(t)\hat{c}^\dagger_\beta(t')\rangle
\tag{12.16}
$$

[a]

---

[a] Eq. 12.16 is analogous to Eq. 10.11 in the book of Abrikosov, Gorkov Dzyaloshinski[65].

12.5 Self energy

Eq. 12.16 shows that unlike the non-interacting case the generalization of the Green's function to interacting systems does not obey an analogous differential equation. In order to obtain the equation for the Green's function *matG*, one needs the two-particle Green's function, which contains four creation and annihilation operators. In order to determine the latter we obtain a differential equation, where even higher-order Green's functions enter. We end up with a infinite sequence of increasingly more complicated differential equations.

In order to recover a form more closely related to the defining equation of a Green's function we define a **self energy** via

DEFINITION OF THE SELF ENERGY

$$\sum_{\gamma}\int dt''\,\Sigma_{\alpha,\gamma}(t,t'')G_{\gamma,\beta}(t'',t')=\sum_{b,c,d}W_{\alpha,b,d,c}\frac{1}{i\hbar}\langle\mathcal{T}\hat{c}^{\dagger}_{b}(t)\hat{c}_{c}(t)\hat{c}_{d}(t)\hat{c}^{\dagger}_{\beta}(t')\rangle \tag{12.17}$$

so that

$$\sum_{\gamma}\int dt''\Big(\delta(t-t'')\Big(\delta_{\alpha,\gamma}i\hbar\partial_{t''}-h_{\alpha,\gamma}\Big)-\Sigma_{\alpha,\gamma}(t,t'')\Big)G_{\gamma,\beta}(t'',t')=\delta(t-t')\delta_{\alpha,\beta} \tag{12.18}$$

The self-energy acts like a retarded one-particle potential that describes the forces of remaining particles system.

The defining equation Eq. 12.17 provides us with an expression for the interaction energy, namely

$$\frac{-i\hbar}{2}\sum_{\alpha,\gamma}\int dt''\,\Sigma_{\alpha,\gamma}(t,t'')G_{\gamma,\alpha}(t'',t^{+})=\Big\langle\frac{1}{2}\sum_{\alpha,b,c,d}W_{\alpha,b,d,c}\hat{c}^{\dagger}_{\alpha}\hat{c}^{\dagger}_{b}\hat{c}_{c}\hat{c}_{d}\Big\rangle \tag{12.19}$$

This is a very strange way for calculating an energy: We use the information about the dynamics of particles immersed in the electron gas for all time, to extract a total energy at some initial time.

Eq. 12.18 is a **Dyson equation**

$$\begin{aligned}&\int dt''\,\Big(\hat{G}_0^{-1}(t,t'')-\hat{\Sigma}(t,t'')\Big)\hat{G}(t'',t')=\hat{1}\delta(t-t')\\ \Rightarrow\qquad &\hat{G}(t,t')=\hat{G}_0(t,t')+\int dt''\int d\bar{t}\,\hat{G}_0(t,t'')\hat{\Sigma}(t'',\bar{t})\hat{G}(\bar{t},t')\end{aligned} \tag{12.20}$$

12.6 Physical meaning of the Green's function

Note that the many-particle Green's function can no more be written as the inverse of the Schrödinger equation.

12.6.1 Expectation value of single particle operators

Consider a one-particle operator, which can always be expressed as

$$\hat{A}(t)=\sum_{n,m}A_{n,m}\hat{c}^{\dagger}_{n}\hat{c}_{m}=\int d^4x\int d^4x'\,A(\vec{x},\vec{x'},t)\hat{\psi}^{\dagger}(x)\hat{\psi}(\vec{x'})$$

The ground-state expectation value is

$$\begin{aligned}
\langle\mathcal{G}|\hat{A}|\mathcal{G}\rangle &= \int d^4x \int d^4x'\; A(\vec{x},\vec{x'},t)\langle\mathcal{G}|\hat{\psi}^\dagger(x)\hat{\psi}(\vec{x'})|\mathcal{G}\rangle \\
&= \int d^4x \int d^4x'\; A(\vec{x},\vec{x'})\langle\mathcal{G}|\underbrace{\hat{U}(0,t)}_{\mathrm{e}^{\frac{i}{\hbar}E_{\mathcal{G}}t}}\hat{\psi}^\dagger(x)\underbrace{\hat{U}(t,0)\hat{U}(0,t)}_{\hat{1}}\hat{\psi}(\vec{x'})\underbrace{\hat{U}(t,0)}_{\mathrm{e}^{-\frac{i}{\hbar}E_{\mathcal{G}}t}}|\mathcal{G}\rangle \\
&= \int d^4x \int d^4x'\; A(\vec{x},\vec{x'})\langle\mathcal{G}|\hat{\psi}^\dagger_H(\vec{x},t)\hat{\psi}_H(\vec{x'},t)|\mathcal{G}\rangle \\
&= \int d^4x \int d^4x'\; A(\vec{x},\vec{x'})\lim_{t'\to t^+}(-1)\langle\mathcal{G}|T[\hat{\psi}_H(\vec{x},t)\hat{\psi}^\dagger_H(\vec{x'},t')]|\mathcal{G}\rangle \\
&= -i\hbar\lim_{t'\to t^+}\int d^4x \int d^4x'\; A(\vec{x},\vec{x'})G(\vec{x},t,\vec{x'},t')
\end{aligned}$$

Thus, we indirectly obtain the one-particle density matrix of the many-particle system as

$$\boxed{\rho^{(1)}(\vec{x},\vec{x'}) = -i\hbar\lim_{t'\to t^+}G(\vec{x},t,\vec{x'},t')}$$

12.6.2 Excitation spectrum and Lehmann Representation

The Green's function contains the information on the complete one-particle spectrum of the many-particle system.

Let us rewrite Eq. 12.8

$$\begin{aligned}
G(\vec{x},t,\vec{x'},t') &= \frac{1}{i\hbar}\frac{\langle\mathcal{G}|T[\hat{\psi}_H(\vec{x},t)\hat{\psi}^\dagger_H(\vec{x'},t')]|\mathcal{G}\rangle}{\langle\mathcal{G}|\mathcal{G}\rangle} \\
&= \frac{1}{i\hbar}\theta(t-t')\frac{\langle\mathcal{G}|\hat{\psi}_H(\vec{x},t)\hat{\psi}^\dagger_H(\vec{x'},t')|\mathcal{G}\rangle}{\langle\mathcal{G}|\mathcal{G}\rangle} - \frac{1}{i\hbar}\theta(t'-t)\frac{\langle\mathcal{G}|\hat{\psi}^\dagger_H(\vec{x'},t')\hat{\psi}_H(\vec{x},t)|\mathcal{G}\rangle}{\langle\mathcal{G}|\mathcal{G}\rangle}
\end{aligned}$$

Now we insert a complete many-particle basis $|\Psi_n\rangle$ of eigenstates of the many-particle Hamiltonian, that is

$$\hat{H}|\Psi_n\rangle = |\Psi_n\rangle E_n$$

Thus, we can write

$$\hat{U}(t,0)|\Psi_n\rangle = |\Psi_n\rangle\mathrm{e}^{-\frac{i}{\hbar}E_nt}$$

This basis set is complete in the Fock space, that is it contains the N-particle states for all N. The unit operator in Fock space can then be written as

$$\hat{1} = \sum_n |\Psi_n\rangle\langle\Psi_n|$$

$$\begin{aligned}
G(\vec{x},t,\vec{x'},t') &= \frac{1}{i\hbar}\sum_n \theta(t-t')\langle\mathcal{G}|\hat{\psi}_H(\vec{x},t)|\Psi_n\rangle\langle\Psi_n|\hat{\psi}^\dagger_H(\vec{x'},t')|\mathcal{G}\rangle \\
&\quad - \frac{1}{i\hbar}\sum_n \theta(t'-t)\langle\mathcal{G}|\hat{\psi}^\dagger_H(\vec{x'},t')|\Psi_n\rangle\langle\Psi_n|\hat{\psi}_H(\vec{x},t)|\mathcal{G}\rangle \\
&= \frac{1}{i\hbar}\sum_n \theta(t-t')\mathrm{e}^{-\frac{i}{\hbar}(E_n-E_{\mathcal{G}})(t-t')}\langle\mathcal{G}|\hat{\psi}(\vec{x})|\Psi_n\rangle\langle\Psi_n|\hat{\psi}^\dagger(\vec{x'})|\mathcal{G}\rangle \\
&\quad - \frac{1}{i\hbar}\sum_n \theta(t-t')\mathrm{e}^{\frac{i}{\hbar}(E_n-E_{\mathcal{G}})(t-t')}\langle\mathcal{G}|\hat{\psi}^\dagger(\vec{x'})|\Psi_n\rangle\langle\Psi_n|\hat{\psi}(\vec{x})|\mathcal{G}\rangle
\end{aligned}$$

If the ground state has N particles, the creation operator makes a $N+1$ particle state out of it and the annihilator makes an $N-1$ particle state out of it. Since states with different particle numbers are orthogonal, only states $|\Psi_n\rangle$ with either $N+1$ or $N-1$ particles contribute to the Green's function.

In our case the Hamiltonian is time independent. Therefore the Green's function only depends on the relative time argument $t-t'$. In order to extract the spectral properties we Fourier transform the above expression for the Green's function in the time coordinate.

We use the equation Eq. B.2 derived in Appendix B.1 (p. 243)

$$\int dt\; \theta(t) e^{\frac{i}{\hbar}(\epsilon+i\eta)t} = \frac{i\hbar}{\epsilon+i\eta}$$
$$\int dt\; \theta(-t) e^{\frac{i}{\hbar}(\epsilon-i\eta)t} = \frac{-i\hbar}{\epsilon-i\eta}$$

Thus, we obtain the

LEHMANN REPRESENTATION OF THE GREEN'S FUNCTION

$$G(\vec{x},\vec{x'},\omega) \overset{\text{def}}{=} \int dt\; e^{\frac{i}{\hbar}(\hbar\omega t+i\eta|t|)} G(\vec{x},t,\vec{x'},0)$$
$$= \sum_n \Big[\underbrace{\frac{\langle\mathcal{G}|\hat\psi(\vec{x})|\Psi_n\rangle\langle\Psi_n|\hat\psi^\dagger(\vec{x'})|\mathcal{G}\rangle}{\hbar\omega-(E_n-E_\mathcal{G})+i\eta}}_{\text{electron like}} + \underbrace{\frac{\langle\mathcal{G}|\hat\psi^\dagger(\vec{x'})|\Psi_n\rangle\langle\Psi_n|\hat\psi(\vec{x})|\mathcal{G}\rangle}{\hbar\omega+(E_n-E_\mathcal{G})-i\eta}}_{\text{hole like}} \Big]$$

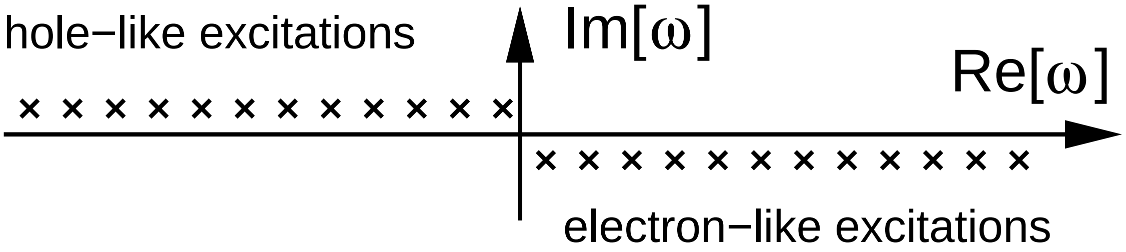


Fig. 12.1: Position of the poles of the interacting Green's function for non-interacting system. The poles are located in the complex ω plane at the positions given by the excitation energy (real part) and $\pm\eta$ (imaginary part). For an interacting system the excitations may move a finite distance away form the real axis, which corresponds to a finite life time of the corresponding quasi particles.

12.7 Non-interacting Green's function

Later we will express the interacting Green's functions by non-interacting Green's functions. Hence, we need to know the non-interacting Green's function.

The Green's function is expressed by Heisenberg operators. The Heisenberg operators with the non-interacting Hamiltonian are the operators in the interaction picture. $|\mathcal{G}^{(0)}\rangle$ is the ground state of the non-interacting system.

The field operators in the interaction picture have a particularly simple form in the representation of eigenstates of the unperturbed Hamiltonian, namely

$$
\begin{aligned}
\hat{c}_{n,I}(t) &= \mathrm{e}^{-\frac{i}{\hbar}\epsilon_n t}\hat{c}_n \\
\hat{c}^\dagger_{n,I}(t) &= \mathrm{e}^{+\frac{i}{\hbar}\epsilon_n t}\hat{c}^\dagger_n
\end{aligned}
\tag{12.21}
$$

In contrast to the similar expression in the one-particle picture, the derivation of the Eq. 12.21 is non-trivial. It is shown in appendix S on p. 319. The derivation is based on the Baker-Campbell-Hausdorff theorem[66, 67, 68, 69, 70, 71]. This theorem, which is used to derive a number of useful relations involving exponentials of operators is also shown in appendix. T.2.

Let us first bring the Green's function in a representation of eigenstates of the non-interacting Hamiltonian.

$$
\begin{aligned}
G^{(0)}(\vec{x},t,\vec{x}',t') &\overset{Eq.\ 12.8}{=} \frac{1}{i\hbar}\frac{\langle\mathcal{G}^{(0)}|T[\hat{\psi}_I(\vec{x},t)\ \hat{\psi}^\dagger_I(\vec{x}',t')]|\mathcal{G}^{(0)}\rangle}{\langle\mathcal{G}^{(0)}|\mathcal{G}^{(0)}\rangle} \\
&= \frac{1}{i\hbar}\frac{\langle\mathcal{G}^{(0)}|T[\sum_n\phi_n(\vec{x})\hat{c}_{n,I}(t)\sum_m\phi^*_m(\vec{x}')\hat{c}^\dagger_{m,I}(t')]|\mathcal{G}^{(0)}\rangle}{\langle\mathcal{G}^{(0)}|\mathcal{G}^{(0)}\rangle} \\
&= \sum_{n,m}\phi_n(\vec{x})\phi^*_m(\vec{x}')\underbrace{\left(\frac{1}{i\hbar}\frac{\langle\mathcal{G}^{(0)}|T[\hat{c}_{n,I}(t)\hat{c}^\dagger_{m,I}(t')]|\mathcal{G}^{(0)}\rangle}{\langle\mathcal{G}^{(0)}|\mathcal{G}^{(0)}\rangle}\right)}_{G^{(0)}_{n,m}(t,t')}
\end{aligned}
$$

The Green's function $G^{(0)}_{n,m}(t,t')$ can now be worked out using Eq. 12.21

$$
\begin{aligned}
G^{(0)}_{n,m}(t,t') &= \frac{1}{i\hbar}\frac{\langle\mathcal{G}^{(0)}|T[\hat{c}_{n,I}(t)\hat{c}^\dagger_{m,I}(t')]|\mathcal{G}^{(0)}\rangle}{\langle\mathcal{G}^{(0)}|\mathcal{G}^{(0)}\rangle} \\
&= \frac{1}{i\hbar}\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}\delta_{n,m}\left(\theta(\mu-\epsilon_n)\theta(t-t')-\theta(\epsilon_n-\mu)\theta(t'-t)\right)
\end{aligned}
$$

Thus, we obtain the

NON-INTERACTING GREEN'S FUNCTION

$$
\begin{aligned}
G^{(0)}(\vec{x},t,\vec{x}',t') &= \underbrace{\frac{1}{i\hbar}\sum_n\theta(\mu-\epsilon_n)\phi_n(\vec{x})\theta(t-t')\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}\phi^*_n(\vec{x}')}_{\text{electron part}} \\
&\quad\underbrace{-\frac{1}{i\hbar}\sum_n\theta(\epsilon_n-\mu)\phi_n(\vec{x})\theta(t'-t)\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}\phi^*_n(\vec{x}')}_{\text{hole part}}
\end{aligned}
\tag{12.22}
$$

This expression is very similar to the unperturbed Green's function Eq. 10.8 of one-particle quantum mechanics. The main difference is that the contribution for the occupied states is similar to the advanced Green's function.

It is left to the reader to verify that the non-interacting Green's function obeys the same differential equation

DIFFERENTIAL EQUATION FOR THE NON-INTERACTING GREEN'S FUNCTION

$$\left(i\hbar\partial_t - \hat{H}_0\right)\hat{G}^{(0)}(t,t') = \hat{1}\delta(t-t') \tag{12.23}$$

as the one for the one-particle system given in Eq. 10.6. Remember, however, that they are not identical! (A differential equation has several solutions, which depend on the boundary conditions. The latter differ.) The derivation will make clear why the hole contribution of the Green's function enters with a negative sign.

12.8 Migdal-Galitskii-Koltun (MGK) sum rule and total energy

Eq. 12.14 leads to the **Migdal-Galitskii-Koltun sum rule**[1], which gives the total energy in terms of the Green function.

$$\lim_{t'\to t^+} \delta_{\alpha,\beta} i\hbar\partial_t G_{\beta,\alpha}(t,t') = \frac{1}{i\hbar}\langle \hat{H}_0 + 2\hat{W}\rangle$$

The contribution with the delta function does not contribute because we choose $t' \to t^+ \neq t$.

MIGDAL-GALITSKII-KOLTUN SUM RULE

$$\lim_{t'\to t^+} \frac{1}{2}\Big(\delta_{\alpha,\beta} i\hbar\partial_t + h_{\alpha,\beta}\Big)\Big(i\hbar G_{\beta,\alpha}(t,t')\Big) = \langle \hat{H}_0 + \hat{W}\rangle = E_{tot} \tag{12.24}$$

The positive sign is correct!

The Migdal-Galitsky sum rule can be used to obtain the total energy,if the self energy is known. For this purpose we use Eq. 12.18 to replace the time derivative in Eq. 12.24. This yields

$$E_{tot} = \sum_{\alpha,\beta} h_{\alpha,\beta}\Big(i\hbar G_{\beta,\alpha}(t,t^+)\Big) + \frac{1}{2}\sum_{\alpha,\beta}\int dt'\, \Sigma_{\alpha,\beta}(t,t')\Big(i\hbar G_{\beta,\alpha}(t',t^+)\Big)$$

12.9 Spectral function

Do not read: sketch only

$$A_{\alpha,\beta}(t,t') = \Big\langle \mathcal{G}\Big|\Big[\hat{\psi}(\vec{x},t), \hat{\psi}^\dagger(\vec{x'},t')\Big]_+\Big|\mathcal{G}\Big\rangle$$

12.10 Variations of the Green's function

Do not read: sketch only

The Green's function we introduced so-far is the so-called causal or time-ordered Green's function. There are other types of Green functions. All of these Green's function contain the same information. That is they can be converted into each other

---

[1] V.M. Galitskii, A.B. Migdal, Application of quantum field theory methods to the many body problem, Sov. Phys., JETP 7, 96-104 (1958)

The Green's functions are the **retarded Green's function** G^R, the **advanced Green's function** G^A, the **greater Green's function** $G^>$, and the **lesser Green's function** $G^<$.

$$G^T(\vec{x}, t, \vec{x'}, t') = \frac{1}{i\hbar}\langle\mathcal{G}|\mathcal{T}_\zeta\hat{\psi}(\vec{x}, t)\hat{\psi}^\dagger(\vec{x'}, t'\Big|\mathcal{G}\rangle \tag{12.25}$$

$$G^R(\vec{x}, t, \vec{x'}, t') = \frac{1}{i\hbar}\theta(t-t')\langle\mathcal{G}|\Big[\hat{\psi}(\vec{x}, t), \hat{\psi}^\dagger(\vec{x'}, t'\Big]_\zeta|\mathcal{G}\rangle \tag{12.26}$$

$$G^A(\vec{x}, t, \vec{x'}, t') = \frac{-1}{i\hbar}\theta(t'-t)\langle\mathcal{G}|\Big[\hat{\psi}(\vec{x}, t), \hat{\psi}^\dagger(\vec{x'}, t'\Big]_\zeta|\mathcal{G}\rangle \tag{12.27}$$

$$G^>(\vec{x}, t, \vec{x'}, t') = \frac{1}{i\hbar}\langle\mathcal{G}|\hat{\psi}(\vec{x}, t), \hat{\psi}^\dagger(\vec{x'}, t'|\mathcal{G}\rangle \tag{12.28}$$

$$G^<(\vec{x}, t, \vec{x'}, t') = \frac{\zeta}{i\hbar}\langle\mathcal{G}|\hat{\psi}(\vec{x}, t), \hat{\psi}^\dagger(\vec{x'}, t'|\mathcal{G}\rangle \tag{12.29}$$

- The time-ordered Green's function has a systematic perturbation theory
- The retarded and advanced Green's functions have a nice analytical structure
- the retarded Green's function is the classical response function.
- the lesser and greater Green's function are directly linked to observables such as particle densities and currents.

(see `http://www.phys.nthu.edu.tw/classnews/96intr_phys/Lecture`)

They are converted as follows

$$G^R(\vec{x}, t, \vec{x'}, t') = \theta(t-t')\Big(G^>(\vec{x}, t, \vec{x'}, t') - G^<(\vec{x}, t, \vec{x'}, t')\Big) \tag{12.30}$$

$$G^A(\vec{x}, t, \vec{x'}, t') = \theta(t'-t)\Big(G^<(\vec{x}, t, \vec{x'}, t') - G^>(\vec{x}, t, \vec{x'}, t')\Big) \tag{12.31}$$

Chapter 13

Diagrammatic expansion of the Green's function

So far we have defined the Green's function and we have shown what it is good for. The problem is that we need to know the ground state $|\mathcal{G}\rangle$ of the interacting system to evaluate it. We can however only determine the Green's function of the non-interacting electron system. In the following we show the techniques that allow us to determine the Green's function of an interacting system from that of a non-interacting system.

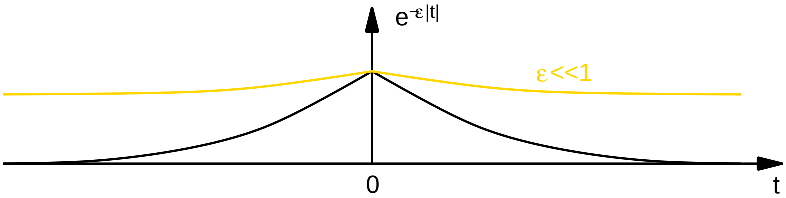
13.1 Adiabatic switching on

In order to construct the many-particle ground state of the interacting system, we start from a non-interacting system and switch the interaction on slowly, so that the initial state is transformed adiabatically into the ground state of the final system.

Our Hamiltonian for the non-interacting system is $\hat{H}_0$ and the interacting Hamiltonian is $\hat{H} = \hat{H}_0 + \hat{W}$. Now we formulate a time dependent problem with a Hamiltonian

$$\hat{H}(t) = \hat{H}_0 + \underbrace{\hat{W}\mathrm{e}^{-\zeta|t|}}_{\hat{W}_\zeta(t)}$$

where ζ is an infinitesimally small constant.



Note, that it is non-trivial to show that the system is indeed adiabatic. If the ground-state energy crosses with another another state some time during the switching on and off process, the former ground state could propagate into both of the crossing states, and the result will be undetermined. Thus, we should require that the ground state is never degenerate. Where could this be a problem? Clearly in a metal we have a continuum of states, implying that the ground state is arbitrarily close in energy to the first excited state. To avoid these problems, we can enclose the system into a very large, but still finite box. As a consequence, the spectrum becomes discrete. (Is this still true for an interacting system?) If the spectrum is discrete, the crossings are usually avoided: If we consider the two states, there is usually some interaction in the Hamiltonian that leads to a splitting. Crossings are not avoided if the matrix elements are zero by symmetry.

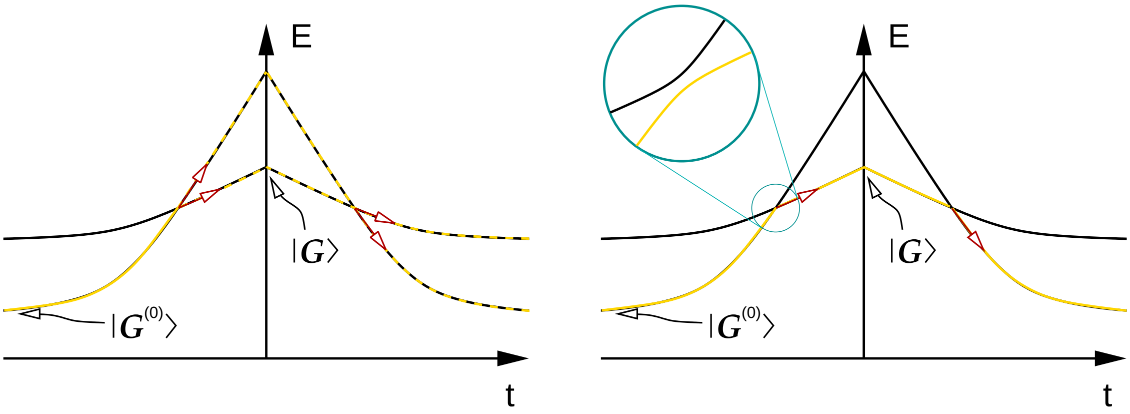


Fig. 13.1: At a band crossing the system can propagate from the ground state into the first excited state, violating the assumption of an adiabatic process. Normally the crossings are avoided, so that at least a small band gap remains. If the potential is weakly time dependent, an excitation across the band gap is no more possible.

The picture[1] behind is that a time dependent perturbation can induce excitations of the system. In order to excite the system from the ground state to the first excited state there must be an energy $\hbar\omega = E_1 - E_0$ transferred by the perturbation. The probability for such an excitation depends on the matrix element and the size of the Fourier component with wave number ω. As η approaches zero, the weight of Fourier components with finite frequency converges to zero. Thus, we conclude that transitions become impossible in the limit $\eta \to 0$, except when the excitation energy vanishes, that is when the ground state and the first excited state become degenerate.

Now we use time-dependent perturbation theory to obtain the state at $t = 0$. The propagator in the interaction picture has the form

PROPAGATOR...

$$\begin{aligned}\hat{U}_{\zeta,I}(t,t') &\overset{Eq.~11.15}{=} T\left[\mathrm{e}^{\frac{1}{i\hbar}\int_{t'}^{t} dt''\,\hat{W}_{\zeta,I}(t'')}\right] \\ &\overset{Eq.~11.15}{=} \sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{1}{i\hbar}\right)^n \int_{t'}^{t} dt_1 \cdots \int_{t'}^{t} dt_n\, T\left[\hat{W}_{\zeta,I}(t_1)\cdots\hat{W}_{\zeta,I}(t_n)\right]\end{aligned} \tag{13.1}$$

where

$$\hat{W}_{\zeta,I}(t) = \hat{U}^{(0)}(0,t)\hat{W}\mathrm{e}^{-\zeta|t|}\hat{U}^{(0)}(t,0)$$

and

$$\hat{U}_I(t,t') = \hat{U}^{(0)}(0,t)\hat{U}(t,t')\hat{U}^{(0)}(t',0)$$

and the propagator of the non-interacting system is

$$\hat{U}^{(0)}(t,t') = \mathrm{e}^{-\frac{i}{\hbar}\hat{H}_0(t-t')}$$

---

[1] Note that this is a motivation, and not a proof.

Now we want to determine the ground state of the interacting system from that of the non-interacting system by adiabatically switching on the interaction. For $t \to -\infty$ the full Hamiltonian approaches the unperturbed Hamiltonian. In that case the propagator becomes constant. This implies that the limes of $\hat{U}_{\zeta,I}(t,-\infty)$ is well defined.

$$|\Psi_I(0)\rangle\rangle = \hat{U}_{\zeta,I}(0,-\infty)|\mathcal{G}^{(0)}\rangle$$

At the origin of time, that is $t = 0$ the wave functions in the three pictures, Schrödinger-, Interaction- and Heisenberg Picture are identical according to Eq. 11.12 and Eq. 12.3, that is

$$|\Psi_I(0)\rangle = |\Psi_S(0)\rangle = |\Psi_H(0)\rangle$$

Thus, we obtain the ground state of the Heisenberg picture as

$$|\mathcal{G}\rangle = |\mathcal{G}_H\rangle = \hat{U}_{\zeta,I}(0,-\infty)|\mathcal{G}^{(0)}\rangle$$

Note that $|\mathcal{G}\rangle$ is the lowest eigenstate of the time independent, but interacting, Schrödinger equation. $|\mathcal{G}^{(0)}\rangle$ is defined analogously for the non-interacting Hamiltonian. None of them have an explicit time dependence.

The interesting point is that we have been able to start from the time-independent ground state of the non-interacting system and obtain the stationary ground state of the interacting system.

We can represent the same state differently in a form that will be required in the following:

$$\begin{aligned}|\mathcal{G}\rangle &= \hat{U}_{\zeta,I}(0,-\infty)|\mathcal{G}^{(0)}\rangle = \hat{U}_{\zeta,I}(0,\infty)\hat{U}_{\zeta,I}(\infty,-\infty)|\mathcal{G}^{(0)}\rangle \\ &= \hat{U}_{\zeta,I}(0,\infty)|\mathcal{G}^{(0)}\rangle\langle\mathcal{G}^{(0)}|\hat{U}_{\zeta,I}(\infty,-\infty)|\mathcal{G}^{(0)}\rangle\end{aligned}$$

Here, we assumed that switching the interaction on and off adiabatically leads back to the same state, even though the state may have a different phase. This is simply the requirement that the interaction does not lead to persisting excitations from the ground state.

The operator that propagates from minus infinity to plus infinity has a special name, the

S-MATRIX

$$\hat{S}_{\zeta,I} \overset{\text{def}}{=} \hat{U}_{\zeta,I}(\infty,-\infty) \tag{13.2}$$

The symbol S stands for Strahlung, Streumatrix or scattering. It relates a state in the infinite future to a state in the infinite past.

Editor: check the proof of Molinari[72]

Concluding, we can express the ground state of the interacting system as

$$|\mathcal{G}\rangle = \hat{U}_{\zeta,I}(0,-\infty)|\mathcal{G}^{(0)}\rangle \tag{13.3}$$

$$\langle\mathcal{G}| = \frac{1}{\langle\mathcal{G}^{(0)}|\hat{S}_{\zeta,I}|\mathcal{G}^{(0)}\rangle}\langle\mathcal{G}^{(0)}|\hat{U}_{\zeta,I}(\infty,0) \tag{13.4}$$

In the last step we have exploited that $\langle\mathcal{G}^{(0)}|\hat{S}_{\zeta,I}|\mathcal{G}^{(0)}\rangle$ is a simple phase factor so that $\langle\mathcal{G}^{(0)}|\hat{S}_{\zeta,I}|\mathcal{G}^{(0)}\rangle^* = \langle\mathcal{G}^{(0)}|\hat{S}_{\zeta,I}|\mathcal{G}^{(0)}\rangle^{-1}$. This is valid, if the system behaves indeed adiabatic and is lead back, up to a phase factor, to the same state after the perturbation is switched off again.

13.2 Interacting Green's function expressed by non-interacting ground states

Next we wish to determine the Green's function

$$\begin{aligned}\hat{G}(x,t,x',t') &\overset{Eq.\ 12.8}{=} \frac{1}{i\hbar}\frac{\langle\mathcal{G}|T[\hat\psi_H(\vec{x},t)\hat\psi^\dagger_H(\vec{x'},t')]|\mathcal{G}\rangle}{\langle\mathcal{G}|\mathcal{G}\rangle}\\ &\overset{Eq.\ 13.4}{=} \frac{1}{i\hbar}\frac{\langle\mathcal{G}^{(0)}|\hat U_{\zeta,I}(\infty,0)T[\hat\psi_H(\vec{x},t)\hat\psi^\dagger_H(\vec{x'},t')]\hat U_{\zeta,I}(0,-\infty)|\mathcal{G}^{(0)}\rangle}{\langle\mathcal{G}^{(0)}|\hat S_{\zeta,I}|\mathcal{G}^{(0)}\rangle}\end{aligned}$$

Let us now write down the time-ordered product in an explicit fashion

$$\begin{aligned}\hat{G}(x,t,x',t') =& \frac{1}{i\hbar}\theta(t-t')\frac{\langle\mathcal{G}^{(0)}|\hat U_{\zeta,I}(\infty,0)\hat\psi_H(\vec{x},t)\hat\psi^\dagger_H(\vec{x'},t')\hat U_{\zeta,I}(0,-\infty)|\mathcal{G}^{(0)}\rangle}{\langle\mathcal{G}^{(0)}|\hat S_{\zeta,I}|\mathcal{G}^{(0)}\rangle}\\ &-\frac{1}{i\hbar}\theta(t'-t)\frac{\langle\mathcal{G}^{(0)}|\hat U_{\zeta,I}(\infty,0)\hat\psi^\dagger_H(\vec{x'},t')\hat\psi_H(\vec{x},t)\hat U_{\zeta,I}(0,-\infty)|\mathcal{G}^{(0)}\rangle}{\langle\mathcal{G}^{(0)}|\hat S_{\zeta,I}|\mathcal{G}^{(0)}\rangle}\end{aligned}$$

Now we relate the Heisenberg operator to an interaction operator

$$\begin{aligned}\hat A_H(t) &= \hat U(0,t)\hat A_S(t)\hat U(t,0)\\ &= \underbrace{\overbrace{\hat U^{(0)}(0,0)}^{\hat 1}\hat U(0,t)\hat U^{(0)}(t,0)}_{\hat U_{\zeta,I}(0,t)}\hat A_I(t)\underbrace{\hat U^{(0)}(0,t)\hat U(t,0)\overbrace{\hat U^{(0)}(0,0)}^{\hat 1}}_{\hat U_{\zeta,I}(t,0)}\\ &= \hat U_I(0,t)\hat A_I(t)\hat U_I(t,0)\end{aligned}$$

Here, I want to alert the reader to a subtle point.The expression for the Green's function contains the propagator $\hat U_I$ and $\hat U_{\zeta,I}$. In the following we use in all cases $\hat U_{\zeta,I}$. Before, that is if we make the distinction, the limes $\zeta\rightarrow 0$ leads to the true interacting Green's function with an interaction that is switched on all the time. The Green's function propagates the particles with the complete interaction. Once we replace $\hat U_I$ by $\hat U_{\zeta,I}$ we determine the Green's function with a time-dependent interaction, that is switched off at $\pm\infty$. This is ok as long as the time arguments of the Green's function are finite, because the limes $\zeta\rightarrow 0$ makes the difference arbitrarily small. However, we must avoid any terms in which the time arguments approach $\pm\infty$.

Thus, we obtain

$$\begin{aligned}&\hat G(x,t,x',t')\\ &=\frac{\theta(t-t')}{i\hbar}\frac{\langle\mathcal{G}^{(0)}|\hat U_{\zeta,I}(\infty,0)\overbrace{\hat U_{\zeta,I}(0,t)\hat\psi_I(\vec{x},t)\hat U_{\zeta,I}(t,0)}^{\hat\psi_H(\vec{x},t)}\overbrace{\hat U_{\zeta,I}(0,t')\hat\psi^\dagger_I(\vec{x'},t')\hat U_{\zeta,I}(t',0)}^{\hat\psi^\dagger_H(\vec{x},t')}\hat U_{\zeta,I}(0,-\infty)|\mathcal{G}^{(0)}\rangle}{\langle\mathcal{G}^{(0)}|\hat S_{\zeta,I}|\mathcal{G}^{(0)}\rangle}\\ &-\frac{\theta(t'-t)}{i\hbar}\frac{\langle\mathcal{G}^{(0)}|\hat U_{\zeta,I}(\infty,0)\overbrace{\hat U_{\zeta,I}(0,t')\hat\psi^\dagger_I(\vec{x'},t')\hat U_{\zeta,I}(t',0)}^{\hat\psi^\dagger_H(\vec{x},t')}\overbrace{\hat U_{\zeta,I}(0,t)\hat\psi_I(\vec{x},t)\hat U_{\zeta,I}(t,0)}^{\hat\psi_H(\vec{x},t)}\hat U_{\zeta,I}(0,-\infty)|\mathcal{G}^{(0)}\rangle}{\langle\mathcal{G}^{(0)}|\hat S_{\zeta,I}|\mathcal{G}^{(0)}\rangle}\\ &=\frac{\theta(t-t')}{i\hbar}\frac{\langle\mathcal{G}^{(0)}|\hat U_{\zeta,I}(\infty,t)\hat\psi_I(\vec{x},t)\hat U_{\zeta,I}(t,t')\hat\psi^\dagger_I(\vec{x'},t')\hat U_{\zeta,I}(t',-\infty)|\mathcal{G}^{(0)}\rangle}{\langle\mathcal{G}^{(0)}|\hat S_{\zeta,I}|\mathcal{G}^{(0)}\rangle}\\ &-\frac{\theta(t'-t)}{i\hbar}\frac{\langle\mathcal{G}^{(0)}|\hat U_{\zeta,I}(\infty,t')\hat\psi^\dagger_I(\vec{x'},t')\hat U_{\zeta,I}(t',t)\hat\psi_I(\vec{x},t)\hat U_{\zeta,I}(t,-\infty)|\mathcal{G}^{(0)}\rangle}{\langle\mathcal{G}^{(0)}|\hat S_{\zeta,I}|\mathcal{G}^{(0)}\rangle}\end{aligned}$$

Now we realize from Eq. 13.1 that the propagator $\hat U_{\zeta,I}(t,t')$ contains only operators with times in the interval $[t,t']$, which shows that the product of operators is time ordered. Thus, we can permute the operators as we whish, as long as we take care of sign changes for the permutation of each Fermion operator. The time ordering operator will sort them into the correct position again. This allows us to shift the creation and annihilation operators past the interaction operators to the very right of the product. The propagator in the interaction picture contains only products of interaction operators. Because the interaction operator is made from an even number of creation and annihilation operators, the sign changes cancel. The only sign change is due to the relative order of the creator and the annihilator, which contain the arguments of the Green's function. When the creator and the annihilator are moved to the right, we can combine the propagators into the S-matrix.

Thus, we obtain our result

$$\boxed{\hat{G}(x,t,x',t') = \frac{1}{i\hbar}\frac{\langle\mathcal{G}^{(0)}|T[\hat{S}_{\zeta,I}\hat{\psi}_I(\vec{x},t)\hat{\psi}_I^\dagger(\vec{x'},t')]|\mathcal{G}^{(0)}\rangle}{\langle\mathcal{G}^{(0)}|\hat{S}_{\zeta,I}|\mathcal{G}^{(0)}\rangle}} \tag{13.5}$$

Thus, we have found a compact expression for the Green's function that does not resort directly to the ground state of the interacting electron system. Note that we use here the operators in the interaction picture and not the Heisenberg picture!

Fertig: WS06/07 23 Doppelstunde 12.Feb.07 Nachmittags

13.3 Wick's theorem

13.3.1 Prelude

In order to determine the expression Eq. 13.5 for the Green's function we decompose the S-matrix into products of creation and annihilation operators

$$\hat{S}_{\zeta,I} \overset{Eq.\ 13.2,13.1}{=} \sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{1}{i\hbar}\right)^n\int_{-\infty}^{\infty}dt_1\cdots\int_{-\infty}^{\infty}dt_n\; T[\hat{W}_{\zeta,I}(t_1)\cdots\hat{W}_{\zeta,I}(t_n)] \tag{13.6}$$

This shows that we need to work out terms of the form

$$\begin{gathered}\langle\mathcal{G}^{(0)}|T[\hat{W}_{\zeta,I}(t_1)\cdots\hat{W}_{\zeta,I}(t_n)\hat{\psi}_I(\vec{x},t)\hat{\psi}_I^\dagger(\vec{x'},t')]|\mathcal{G}^{(0)}\rangle\\ \langle\mathcal{G}^{(0)}|T[\hat{W}_{\zeta,I}(t_1)\cdots\hat{W}_{\zeta,I}(t_n)]|\mathcal{G}^{(0)}\rangle\end{gathered}$$

Each interaction operator is again composed of creation and annihilation operators.

$$\begin{aligned}\hat{W}_\zeta &= \sum_{i,j,k,l}W_{i,j,k,l}\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_l\hat{c}_k\\ \hat{W}_{\zeta,I}(t) &= \sum_{i,j,k,l}W_{i,j,k,l}\mathrm{e}^{-\zeta|t|}\hat{U}^{(0)}(0,t)\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_l\hat{c}_k\hat{U}^{(0)}(t,0)\\ &= \sum_{i,j,k,l}W_{i,j,k,l}\mathrm{e}^{-\zeta|t|}\\ &\quad\cdot\ \hat{U}^{(0)}(0,t)\hat{c}_i^\dagger\underbrace{\hat{U}^{(0)}(t,0)\hat{U}^{(0)}(0,t)}_{\hat{1}}\hat{c}_j^\dagger\underbrace{\hat{U}^{(0)}(t,0)\hat{U}^{(0)}(0,t)}_{\hat{1}}\hat{c}_l\underbrace{\hat{U}^{(0)}(t,0)\hat{U}^{(0)}(0,t)}_{\hat{1}}\hat{c}_k\hat{U}^{(0)}(t,0)\\ &= \sum_{i,j,k,l}W_{i,j,k,l}\mathrm{e}^{-\zeta|t|}\hat{c}_{i,I}^\dagger(t)\hat{c}_{j,I}^\dagger(t)\hat{c}_{l,I}(t)\hat{c}_{k,I}(t)\end{aligned}$$

This shows that we need a way to work out the expectation values of time-ordered products of field operators.

13.3.2 Some definitions

It will be helpful to decompose the field operators into contributions that annihilate the ground state of the non-interacting electron gas and those that do not. We can think of the annihilation part as the electron and hole annihilators of the non-interacting system and the creation part the electron and hole creation operators. This decomposition has the advantage that if we move the annihilators to the very right, the corresponding expectation values with the ground state vanish.

Using the creation and annihilation operators $\hat{c}_{i,I}(t)$ and $\hat{c}^\dagger_{i,I}(t)$ in the basis of eigenstates of the non-interacting Hamiltonian that defines the ground state of the non-interacting electron system, and the corresponding one-particle orbitals $\phi_i(x)$, we define

- the electron annihilation operator

$$\hat{\psi}_I^{(+)}(\vec{x},t)=\sum_i \theta(\epsilon_i-\mu)\phi_i(x)\hat{c}_{i,I}(t)$$

- the hole creation operator

$$\hat{\psi}_I^{(-)}(\vec{x},t)=\sum_i \theta(\mu-\epsilon_i)\phi_i(x)\hat{c}_{i,I}(t)$$

- the electron creation operator

$$\hat{\psi}_I^{(+),\dagger}(\vec{x},t)=\sum_i \theta(\epsilon_i-\mu)\phi^*_i(x)\hat{c}^\dagger_{i,I}(t)$$

- the hole annihilation operator

$$\hat{\psi}_I^{(-),\dagger}(\vec{x},t)=\sum_i \theta(\mu-\epsilon_i)\phi^*_i(x)\hat{c}^\dagger_{i,I}(t)$$

In all these definitions, the one-particle energies and the Fermi level refer to the non-interacting electron gas.

From the definition it becomes clear that

$$\hat{\psi}_I^{(+)}(t)|\mathcal{G}^{(0)}\rangle=\hat{\psi}_I^{(-)\dagger}(t)|\mathcal{G}^{(0)}\rangle=|0\rangle$$
$$\langle\mathcal{G}^{(0)}|\hat{\psi}_I^{(+)\dagger}(t)=\langle\mathcal{G}^{(0)}|\hat{\psi}_I^{(-)}(t)=\langle 0|$$

Furthermore we can express the original creation and annihilation operators as

$$\hat{\psi}_I(t)=\hat{\psi}_I^{(+)}(t)+\hat{\psi}_I^{(-)}(t)$$
$$\hat{\psi}^\dagger_I(t)=\hat{\psi}_I^{(+)\dagger}(t)+\hat{\psi}_I^{(-)\dagger}(t)$$

Normal ordering

The normal product is defined such that all annihilation operators, such as $\hat{\psi}^{(+)}$ and $\hat{\psi}^{(-),\dagger}$, are placed to the right of the creation operators, such as $\hat{\psi}^{(+),\dagger}$ and $\hat{\psi}^{(-)}$. For each interchange of two Fermion field operators the sign is changed.

$$N[\hat{\psi}_I^{(+)}(\vec{x},t)\hat{\psi}_I^{(-)}(\vec{x'},t')]=-\hat{\psi}_I^{(-)}(\vec{x'},t')\hat{\psi}_I^{(+)}(\vec{x},t)$$

The expectation value of every normal ordered product with the ground state of the non-interacting system vanishes.

$$\langle\mathcal{G}^{(0)}|N[\hat{A}\hat{B}\hat{C}\cdots]|\mathcal{G}^{(0)}\rangle=0$$

This is even true, if there are only creation operators because they annihilate the bra.

Contraction

The contraction of two operators is defined as the difference between the time ordered and the normal ordered product

$$C[\hat{A}\hat{B}]\overset{\text{def}}{=}T[\hat{A}\hat{B}]-N[\hat{A}\hat{B}]$$

These are the terms left over, when we convert a time ordered product into a normal product.

The contraction vanishes, if both operators remain in their original order or if both operators are interchanged. In that case we subtract two identical products. If only the operators are interchanged in only of the two products, we obtain an anticommutator due to the unbalanced sign change.

Let us now consider the anti-commutators. We have to make the assumption that the non-interacting Hamilton operator is time independent. With the help of Eqs. S.1 and S.2 shown in the appendix S on p. 319 we find

$$\begin{aligned}
.[\hat{c}_{n,I}(t),\hat{c}_{m,I}(t')]_+ &\overset{\text{Eq. S.1}}{=} \mathrm{e}^{-\frac{i}{\hbar}(\epsilon_n t+\epsilon_m t')}\underbrace{[\hat{c}_n,\hat{c}_m]_+}_{=0}=0\\
.[\hat{c}_{n,I}(t),\hat{c}^\dagger_{m,I}(t')]_+ &\overset{\text{Eq. S.2}}{=} \mathrm{e}^{-\frac{i}{\hbar}(\epsilon_n t-\epsilon_m t')}\underbrace{[\hat{c}_n,\hat{c}^\dagger_m]_+}_{\delta_{n,m}}=\delta_{n,m}\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')}\\
.[\hat{c}^\dagger_{n,I}(t),\hat{c}_{m,I}(t')]_+ &= [\hat{c}^\dagger_{m,I}(t'),\hat{c}_{n,I}(t)]_+=\delta_{n,m}\mathrm{e}^{+\frac{i}{\hbar}\epsilon_n(t-t')}\\
.[\hat{c}^\dagger_{n,I}(t),\hat{c}^\dagger_{m,I}(t')]_+ &= \mathrm{e}^{\frac{i}{\hbar}(\epsilon_n t+\epsilon_m t')}\underbrace{[\hat{c}^\dagger_n,\hat{c}^\dagger_m]_+}_{=0}=0
\end{aligned}$$

Interesting is that the anticommutators are all numbers and no operators any longer!

$$\begin{aligned}
C[\hat{c}_{n,I}(t)\hat{c}_{m,I}(t')]&=0\\
C[\hat{c}^\dagger_{n,I}(t)\hat{c}^\dagger_{m,I}(t')]&=0
\end{aligned}$$

$$\begin{aligned}
T[\hat{c}_{n,I}(t)\hat{c}^\dagger_{m,I}(t')] &= \theta(t-t')\hat{c}_{n,I}(t)\hat{c}^\dagger_{m,I}(t') - \theta(t'-t)\hat{c}^\dagger_{m,I}(t')\hat{c}_{n,I}(t) \\
&= \theta(t-t')\theta(\mu-\epsilon_n)\hat{c}_{n,I}(t)\hat{c}^\dagger_{m,I}(t') + \theta(t-t')\theta(\epsilon_n-\mu)\hat{c}_{n,I}(t)\hat{c}^\dagger_{m,I}(t') \\
&\quad - \theta(t'-t)\theta(\mu-\epsilon_n)\hat{c}^\dagger_{m,I}(t')\hat{c}_{n,I}(t) - \theta(t'-t)\theta(\epsilon_n-\mu)\hat{c}^\dagger_{m,I}(t')\hat{c}_{n,I}(t) \\
N[\hat{c}_{n,I}(t)\hat{c}^\dagger_{m,I}(t')] &= \theta(\mu-\epsilon_n)\hat{c}_{n,I}(t)\hat{c}^\dagger_{m,I}(t') - \theta(\epsilon_n-\mu)\hat{c}^\dagger_{m,I}(t')\hat{c}_{n,I}(t) \\
&= \theta(t-t')\theta(\mu-\epsilon_n)\hat{c}_{n,I}(t)\hat{c}^\dagger_{m,I}(t') + \theta(t'-t)\theta(\mu-\epsilon_n)\hat{c}_{n,I}(t)\hat{c}^\dagger_{m,I}(t') \\
&\quad - \theta(t-t')\theta(\epsilon_n-\mu)\hat{c}^\dagger_{m,I}(t')\hat{c}_{n,I}(t) - \theta(t'-t)\theta(\epsilon_n-\mu)\hat{c}^\dagger_{m,I}(t')\hat{c}_{n,I}(t) \\
C[\hat{c}_{n,I}(t)\hat{c}^\dagger_{m,I}(t')] &= \theta(t-t')\theta(\mu-\epsilon_n)\underbrace{\Big(\hat{c}_{n,I}(t)\hat{c}^\dagger_{m,I}(t') - \hat{c}_{n,I}(t)\hat{c}^\dagger_{m,I}(t')\Big)}_{=0} \\
&\quad + \theta(t-t')\theta(\epsilon_n-\mu)\underbrace{\Big(\hat{c}_{n,I}(t)\hat{c}^\dagger_{m,I}(t') + \hat{c}^\dagger_{m,I}(t')\hat{c}_{n,I}(t)\Big)}_{[\hat{c}_{n,I}(t),\hat{c}^\dagger_{m,I}(t')]_+} \\
&\quad + \theta(t'-t)\theta(\mu-\epsilon_n)\underbrace{\Big(-\hat{c}^\dagger_{m,I}(t')\hat{c}_{n,I}(t) - \hat{c}_{n,I}(t)\hat{c}^\dagger_{m,I}(t')\Big)}_{-[\hat{c}_{n,I}(t)\hat{c}^\dagger_{m,I}(t')]_+} \\
&\quad + \theta(t'-t)\theta(\epsilon_n-\mu)\underbrace{\Big(-\hat{c}^\dagger_{m,I}(t')\hat{c}_{n,I}(t) + \hat{c}^\dagger_{m,I}(t')\hat{c}_{n,I}(t)\Big)}_{=0} \\
&= \left(\theta(t-t')\theta(\epsilon_n-\mu) - \theta(t'-t)\theta(\mu-\epsilon_n)\right)[\hat{c}_{n,I}(t),\hat{c}^\dagger_{m,I}(t')]_+ \\
&= \left(\theta(t-t')\theta(\epsilon_n-\mu) - \theta(t'-t)\theta(\mu-\epsilon_n)\right)\delta_{n,m}\mathrm{e}^{-\frac{i}{\hbar}\epsilon_n(t-t')} \\
&= i\hbar G^{(0)}_{n,m}(t,t') \\
C[\hat{c}^\dagger_{n,I}(t)\hat{c}_{m,I}(t')] &= -C[\hat{c}_{m,I}(t')\hat{c}^\dagger_{n,I}(t)] \\
&= -i\hbar G^{(0)}_{m,n}(t',t)
\end{aligned}$$

We can bring this result into real space and summarize the result for the contractions:

$$\begin{aligned}
C[\hat{\psi}_I(\vec{x},t)\hat{\psi}_I(\vec{x'},t)] &= 0 \\
C[\hat{\psi}^\dagger_I(\vec{x},t)\hat{\psi}^\dagger_I(\vec{x'},t)] &= 0 \\
C[\hat{\psi}(\vec{x},t)\hat{\psi}^\dagger_I(\vec{x'},t)] &= i\hbar G^{(0)}(\vec{x},t,\vec{x'},t') \\
C[\hat{\psi}^\dagger(\vec{x},t)\hat{\psi}(\vec{x'},t)] &= -i\hbar G^{(0)}(\vec{x'},t',\vec{x},t)
\end{aligned}$$

It is still unclear what to do for equal times. Equal times occur because each interaction term introduces two creators and two annihilators with the same time argument. The convention of the time ordering operator is to leave the original order of the operators in the argument. That is the creator is located left of the corresponding annihilator. Thus, a contraction

$$\begin{aligned}
C[\hat{\psi}^\dagger(\vec{x},t)\hat{\psi}(\vec{x'},t)] &= -i\hbar G^{(0)}(\vec{x},t,\vec{x'},t^+) \\
C[\hat{\psi}(\vec{x},t)\hat{\psi}^\dagger(\vec{x'},t)] &= i\hbar G^{(0)}(\vec{x},t,\vec{x'},t^+)
\end{aligned}$$

Editor: This is not entirely clear!

13.3.3 Wick's Theorem

Wick's theorem[2] represents the time-ordered product of field operators by the sum of terms where each term consists of a normal ordered product and contractions. Since we later form the expectation

---

[2] Gian-Carlo Wick (1909-1992) was an Italian theoretical physicist. The Wick rotation, Wick's theorem, and the Wick product are named after him.

value of the time ordered product with the non-interacting ground state, every term that still contains a normal product vanishes, and we are left with terms that contain only products of contractions, that is Green's functions.

WICK'S THEOREM

Wick's[73] says that the time ordered product of an even number of field operators in the interaction picture is a sum of all possible combinations of a normal ordered product of an even subset of operators and contractions of the remaining operators. Each term in the sum contains a factor, which combines the sign changes that are required to bring the two operators in the contraction next to each other.

$$\begin{aligned}
T[\hat{A}_1\hat{A}_2\cdots\hat{A}_n] &= \underbrace{N[\hat{A}_1\hat{A}_2\cdots\hat{A}_n]}_{\text{normal product}}\\
&+\underbrace{(-1)^0C[\hat{A}_1\hat{A}_2]N[\hat{A}_3\hat{A}_4\cdots\hat{A}_n]+(-1)^1C[\hat{A}_1\hat{A}_3]N[\hat{A}_2\hat{A}_4\cdots\hat{A}_n]\ldots}_{\text{all possible terms with one contraction (singles)}}\\
&+\underbrace{(-1)^0C[\hat{A}_1\hat{A}_2]C[\hat{A}_3\hat{A}_4]N[\hat{A}_5\hat{A}_6\cdots\hat{A}_n]+(-1)^1C[\hat{A}_1\hat{A}_2]C[\hat{A}_3\hat{A}_5]N[\hat{A}_4\hat{A}_6\cdots\hat{A}_n]\ldots}_{\text{all possible terms with two contraction (doubles)}}\\
&\vdots\\
&+\underbrace{(-1)^0C[\hat{A}_1\hat{A}_2]\ldots C[\hat{A}_{n-1}\hat{A}_n]+(-1)^1C[\hat{A}_1\hat{A}_3]\ldots C[\hat{A}_{n-1}\hat{A}_n]\ldots}_{\text{all possible fully contracted terms}}
\end{aligned}$$

Lemma

In order to derive Wick's theorem we need to prove the following lemma:

The product of a normal product $N[\hat{A}_1\cdots\hat{A}_k]$ and field operator $\hat{A}_{k+1}(t_{k+1})$ with $t_1,\ldots t_k > t_{k+1}$ is a sum of time ordered products and and all contractions with the additional operator.

$$\begin{aligned}
N[\hat{A}_1\cdots\hat{A}_k]\hat{A}_{k+1} &= \sum_{j=1}^{k}(-1)^{k-j}N[\hat{A}_1\cdots\hat{A}_{j-1}\hat{A}_{j+1}\cdots\hat{A}_k]C[\hat{A}_j\hat{A}_{k+1}]\\
&+N[\hat{A}_1\cdots\hat{A}_k\hat{A}_{k+1}]
\end{aligned}\tag{13.7}$$

Case 1: $\hat{A}_{k+1}$ is annihilator

Let $\hat{A}_{k+1}$ be an annihilator of $|\mathcal{G}^{(0)}\rangle$ The product is already normal ordered and we obtain

$$N[\hat{A}_1\cdots\hat{A}_k]\hat{A}_{k+1} = N[\hat{A}_1\cdots\hat{A}_k\hat{A}_{k+1}]$$

All contractions between any operator in the normal ordered product and $\hat{A}_{k+1}$ vanish as shown below, so that this result is consistent with Eq. 13.7.

$$\begin{aligned}
C[\hat{A}_i\hat{A}_{k+1}] &= T[\hat{A}_i\hat{A}_{k+1}]-N[\hat{A}_i\hat{A}_{k+1}]\\
&\overset{t_{k+1}>t_i}{=} \hat{A}_i\hat{A}_{k+1}-\hat{A}_i\hat{A}_{k+1}=0
\end{aligned}$$

Case 2: $\hat{A}_{k+1}$ is a creator

Let us assume that the product $\hat{A}_1\cdots\hat{A}_k$ consists of m operators that create the non-interacting ground state and $k-m$ operators that annihilate the ground state.

$$\begin{aligned}
N[\hat{A}_{m+1}\cdots\hat{A}_k]\hat{A}_{k+1} &= \hat{A}_{m+1}\cdots\hat{A}_k\hat{A}_{k+1}\\
&\overset{t_k>t_{k+1}}{=} \hat{A}_{m+1}\cdots\hat{A}_{k-1}T[\hat{A}_k\hat{A}_{k+1}]\\
&\overset{T[]=C[]+N[]}{=} \hat{A}_{m+1}\cdots\hat{A}_{k-1}C[\hat{A}_k\hat{A}_{k+1}]+\hat{A}_{m+1}\cdots\hat{A}_{k-1}N[\hat{A}_k\hat{A}_{k+1}]\\
&= N[\hat{A}_{m+1}\cdots\hat{A}_{k-1}]C[\hat{A}_k\hat{A}_{k+1}]-\underbrace{\Big(N[\hat{A}_{m+1}\cdots\hat{A}_{k-1}]\hat{A}_{k+1}\Big)}_{X}\hat{A}_k
\end{aligned}$$

The term X in parenthesis has the same form as the our starting point. Thus, we can apply this result recursively and obtain

$$\begin{aligned}
N[\hat{A}_{m+1}\cdots\hat{A}_k]\hat{A}_{k+1} = &(-1)^0N[\hat{A}_{m+1}\cdots\hat{A}_{k-1}]C[\hat{A}_k\hat{A}_{k+1}]\\
&+(-1)^1N[\hat{A}_{m+1}\cdots\hat{A}_{k-1}]C[\hat{A}_{k-1}\hat{A}_{k+1}]\hat{A}_k\\
&+(-1)^2N[\hat{A}_{m+1}\cdots\hat{A}_{k-2}]C[\hat{A}_{k-2}\hat{A}_{k+1}]N[\hat{A}_{k-1}\hat{A}_k]\\
&\vdots\\
&+(-1)^{k-m-1}C[\hat{A}_{m+1}\hat{A}_{k+1}]N[\hat{A}_{m+1}\cdots\hat{A}_k]\\
&+(-1)^{k-m}A_{k+1}N[A_{m+1}\cdots\hat{A}_k]
\end{aligned}$$

The sign changes corresponds to the number of permutations require to place the two operators next to each other before they are contracted. Their number corresponds to the number of operators in between them

Since the contractions are complex numbers we can move them to the end, and combine the remaining operators into a single normal product.

$$\begin{aligned}
N[\hat{A}_{m+1}\cdots\hat{A}_k]\hat{A}_{k+1} = &\sum_{j=m+1}^{k}(-1)^{k-j}N[\hat{A}_{m+1}\cdots\hat{A}_{j-1}\hat{A}_{j+1}\cdots\hat{A}_k]C[\hat{A}_j\hat{A}_{k+1}]\\
&+(-1)^{k-m-1}\hat{A}_{k+1}N[\hat{A}_{m+1}\cdots\hat{A}_k]
\end{aligned}$$

Now we can consider also the creators of the normal ordered product.

$$\begin{aligned}
N[\hat{A}_1\cdots\hat{A}_k]\hat{A}_{k+1} &= \sum_{j=m+1}^{k}(-1)^{k-j}N[\hat{A}_{m+1}\cdots\hat{A}_{j-1}\hat{A}_{j+1}\cdots\hat{A}_k]C[\hat{A}_j\hat{A}_{k+1}]\\
&\quad+(-1)^{k-m-1}N[\hat{A}_1\cdots\hat{A}_m\hat{A}_{k+1}\hat{A}_{m+1}\cdots\hat{A}_k]\\
&= \sum_{j=m+1}^{k}(-1)^{k-j}N[\hat{A}_{m+1}\cdots\hat{A}_{j-1}\hat{A}_{j+1}\cdots\hat{A}_k]C[\hat{A}_j\hat{A}_{k+1}]\\
&\quad+N[\hat{A}_1\cdots\hat{A}_m\hat{A}_{m+1}\cdots\hat{A}_k\hat{A}_{k+1}]
\end{aligned}$$

In the last step we have moved the creator $\hat{A}_{k+1}$ to the end of the product.Because each permutation contributes a sign change, the sign changes cancel the prefactor of the normal product.

Because the contractions between two creators vanish, we can extend the sum also to the terms involving the creators.

$$\begin{aligned}
N[\hat{A}_1\cdots\hat{A}_k]\hat{A}_{k+1} = &\sum_{j=1}^{k}(-1)^{k-j}N[\hat{A}_1\cdots\hat{A}_{j-1}\hat{A}_{j+1}\cdots\hat{A}_k]C[\hat{A}_j\hat{A}_{k+1}]\\
&+N[\hat{A}_1\cdots\hat{A}_m\hat{A}_{m+1}\cdots\hat{A}_k\hat{A}_{k+1}]
\end{aligned}$$

Thus, we have proven our lemma.

Proof of Wick's theorem

The proof of Wick's theorem proceeds by induction.

For a product of two operators Wick's theorem is just the definition of the contraction

$$T[\hat{A}_1\hat{A}_2] = C[\hat{A}_1\hat{A}_2] + N[\hat{A}_1\hat{A}_2]$$

Now we assume that Wick's theorem holds for n operators and derive that it holds for $n+2$ operators as well. We assume again wlog[3] that the operators in the argument of the time-ordered product are already time ordered.

$$T[\hat{A}_1 \cdots \hat{A}_n\hat{A}_{n+1}\hat{A}_{n+2}] = \underbrace{T[\hat{A}_1 \cdots \hat{A}_n]}_{X}\hat{A}_{n+1}\hat{A}_{n+2}$$

According to Wick's theorem, which we assume to be valid up to a product with n operators, we can express the term X as a sum of all possible combinations of a normal product and contractions of the operators $\hat{A}_1, \ldots, \hat{A}_n$.

Using the lemma Eq. 13.7, each normal ordered product in the sum can again be written a sum of all possible products of normal products and contractions with the new operators. This is already the proof of Wick's theorem.

Fertig: WS06/07 24 Doppelstunde 13.Feb.07

---

[3] wlog=Without loss of generality

Chapter 14

Feynman diagrams

Now we use Wick's theorem to determine the Green's function Eq. 13.5 of the interacting system

$$\hat{G}(x,t,x',t')=\frac{1}{i\hbar}\frac{\langle\mathcal{G}^{(0)}|T[\hat{S}_{\zeta,I}\hat{\psi}_I(\vec{x},t)\hat{\psi}_I^\dagger(\vec{x'},t')]|\mathcal{G}^{(0)}\rangle}{\langle\mathcal{G}^{(0)}|\hat{S}_{\zeta,I}|\mathcal{G}^{(0)}\rangle} \tag{14.1}$$

Numerator and denominator can be expressed individually as a series expansion Eq. 13.6. The numerator has the form

$$\begin{aligned}&\langle\mathcal{G}^{(0)}|T[\hat{S}_{\zeta,I}\hat{\psi}_I(\vec{x},t)\hat{\psi}_I^\dagger(\vec{x'},t')]|\mathcal{G}^{(0)}\rangle\\ &=\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{1}{i\hbar}\right)^n\int_{-\infty}^{\infty}dt_1\cdots\int_{-\infty}^{\infty}dt_n\;\langle\mathcal{G}^{(0)}|T[\hat{W}_{\zeta,I}(t_1)\cdots\hat{W}_{\zeta,I}(t_n)\hat{\psi}_I(\vec{x},t)\hat{\psi}_I^\dagger(\vec{x'},t')]|\mathcal{G}^{(0)}\rangle\end{aligned} \tag{14.2}$$

and the denominator has the form

$$\begin{aligned}&\langle\mathcal{G}^{(0)}|\hat{S}_{\zeta,I}|\mathcal{G}^{(0)}\rangle\\ &=\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{1}{i\hbar}\right)^n\int_{-\infty}^{\infty}dt_1\cdots\int_{-\infty}^{\infty}dt_n\;\langle\mathcal{G}^{(0)}|T[\hat{W}_{\zeta,I}(t_1)\cdots\hat{W}_{\zeta,I}(t_n)]|\mathcal{G}^{(0)}\rangle\end{aligned} \tag{14.3}$$

where

$$\hat{W}_I(t)\overset{Eq.\ 8.15}{=}\frac{1}{2}\int d^4x\int d^4x'\int_{-\infty}^{\infty}dt'\;v_{int}(\vec{x},t,\vec{x'},t')\hat{\psi}_I^\dagger(x,t)\hat{\psi}_I^\dagger(x',t')\hat{\psi}_I(x',t')\hat{\psi}_I(x,t)$$

and

$$v_{int}(\vec{x},t,\vec{x'},t')=\frac{e^2\delta(t-t')}{4\pi\epsilon_0|\vec{r}-\vec{r'}|} \tag{14.4}$$

Basic diagrams

Our goal will be to work out numerator Eq. 14.2 and denominator Eq. 14.3 of the Green's function Eq. 14.1.

If we move all numbers out of the ground-state expectation value and leave only the time ordered products of creation and annihilation operators, we end up with expressions of the form

$$\begin{aligned}&X_n(\vec{x}_1,t_1,\ldots,\vec{x}_n,t_n,\vec{x'}_n,t'_n)\\ &\stackrel{\text{def}}{=}\langle\mathcal{G}^{(0)}|T\left[\hat{\psi}_I(\vec{x}_1,t_1)\hat{\psi}_I^\dagger(\vec{x}_1,t_1)\;\hat{\psi}_I(\vec{x}_2,t_2)\hat{\psi}_I^\dagger(\vec{x}_2,t_2)\cdots\hat{\psi}_I(\vec{x}_n,t_n)\hat{\psi}_I^\dagger(\vec{x'}_n,t'_n)\right]|\mathcal{G}^{(0)}\rangle\end{aligned} \tag{14.5}$$

Only for the last index n the spatial and time coordinates may be different for the annihilator and the creator. These arguments, namely $(\vec{x}_n, t_n)$ and $(\vec{x}'_n, t')$, are the arguments of the Green's function. For the matrix elements in the denominator, the two arguments are the same. We have placed the creators next to the annihilators. The number of sign changes, which are necessary to reorder the operators in the time ordered product, is even, so that they cancel.

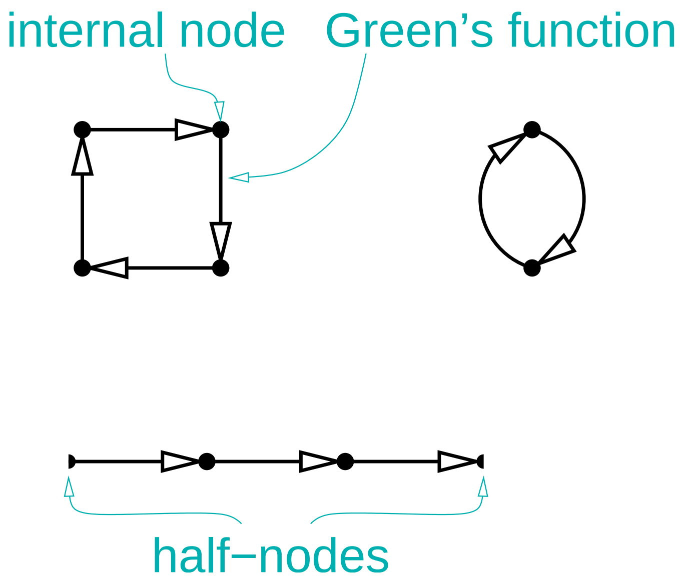
The expectation values X_n in Eq. 14.5 need to be worked out using Wick's theorem. According to Wick's theorem, the result consists of all possible contractions, that is Green's functions. Since there are many of them and each term corresponds to a complex expression, it is almost impossible to keep track of them, not to speak of developing some physical intuition for these terms.

A tremendous simplification is to represent these expressions graphically, which will lead us to Feynman diagrams. Let us describe the construction of the diagrams for one term X_n in Eq. . In our example we use $n = 9$.

Each space-time argument corresponds to a node, represented by a dot. Each contraction, that is, each Green's function connects two nodes in a directed manner, that is from creator to annihilator. Therefore we represent a Green's function by an arrow. In order to work out all possible contractions, we need to connect the nodes in all possible ways by contractions such that each node is connected to exactly one incoming and one outgoing Green's function.

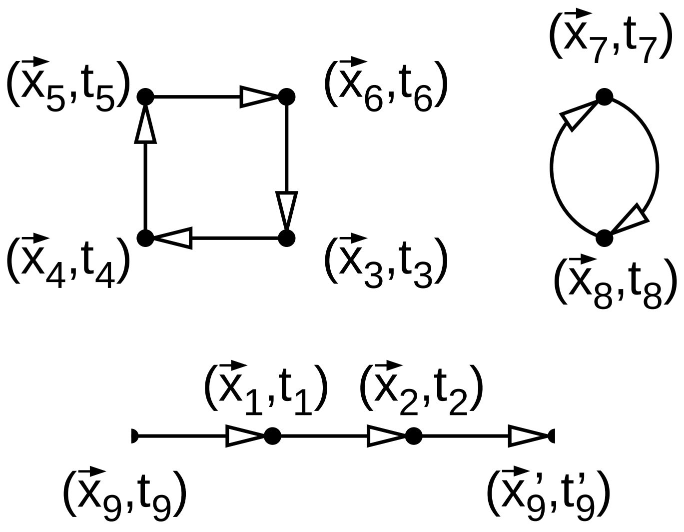
In addition to the nodes, which we will denote **internal nodes**, we also have one incoming and one outgoing **"half node"**. The half nodes are related to the last argument, that is split. They are not related to an interaction, but to the arguments of the Green's function.

Such a graph may look like this



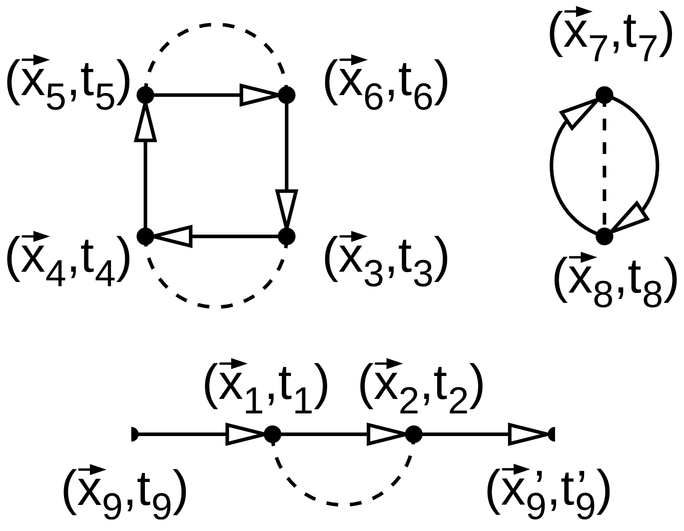
Next we attach the arguments to the nodes: For each diagram, consisting of one strand and a set of closed necklaces, we attach the arguments (x_i, t_i) in all possible ways to the internal nodes. For n internal nodes, there are $n!$ different arrangements. Finally we annotate the half nodes: The last argument is split into $\vec{x}_n, t_n$ and $\vec{x'}_n, t'_n$. The argument $\vec{x}_n, t_n$, which is connected to the annihilator is attached to the incoming half node and the argument $\vec{x'}_n, t'_n$ is connected to the outgoing half node.

A diagram may now look like this:



Interaction lines

As a second step, we introduce the interactions. An interaction connects two internal nodes. If the nodes are identified with defined indices, the arrangement of the interactions is already determined: There is one interaction from node 1 to node 2. Another one from node 2 to node 3. etc. We end up with diagrams like this



An interaction line has two vertices.

A **vertex**[1] is the point where an interaction line and two Green's functions meet.

Up to now we have determined the rules to represent all terms contributing to X_n in terms of diagrams. In the following we need to find out how to translate these diagrams into mathematical expression that can be evaluated.

Fermi Loops

We have ordered the operators such that the annihilator is left of the corresponding creator. However, we also have to satisfy the rule that for equal time arguments, the creator shall remain left of the annihilator. Thus, if a Green's function starts and ends at the same vertex, we must introduce an factor (-1), that corresponds from interchanging the annihilator and the creator. A Green's function that closes on itself is called a **Fermi loop**.

A Fermi loop corresponds to (-1) times the density of the non-interacting system for the given spin direction.

$$(-1)i\hbar G^{(0)}(\vec{x},t,\vec{x},t^-)=-\sum_n \theta(\mu-\epsilon_n)\phi_n(\vec{x})\phi_n^*(\vec{x})=-n^{(0)}(\vec{x})$$

Diagrammatic rules

1. For each order n in the perturbation theory draw $2n$ nodes. For the numerator add two half nodes.
2. Attribute an index to each node or half node.
3. Connect the nodes in all possible ways with arrows.
4. Connect subsequent nodes with interaction lines. Now we have formed all possible diagrams. Each full node is connected to two arrows and one interaction line.
5. for each propagating line, introduce a factor $i\hbar G(\vec{x},t,\vec{x'},t')$ where the arguments $(\vec{x'},t')$ is the starting node and $(\vec{x},t)$ is the final node.
6. for each non-propagating line, introduce a factor $-\langle\hat{n}(\vec{x})\rangle$. (**Editor: check if spins are treated correctly.**) A non-propagating line is one connecting the two ends of a vertex or that forms a closed loop with itself.
7. for each interaction line include a factor $\frac{1}{2}v_{int}(\vec{x},t,\vec{x'},t')$, where the arguments are taken from the two nodes bounding the vertex line.
8. Introduce the adiabatic factor $\mathrm{e}^{-\zeta|t|}$ for each node that is part of a vertex.
9. Integrate or sum over all coordinates (time,space,spin) of the full nodes The time integration goes over the entire time axis.
10. include a factor $\frac{1}{i\hbar}\frac{1}{n!}\left(\frac{1}{i\hbar}\right)^n$

Topologically different diagrams

Two diagrams are **topologically equivalent**, if they can be made equal by rearranging the diagram without disconnecting the Green's functions or interaction lines connected to a node. Interchanging two Green's functions or two interaction line does not create a topologically different diagram.

Let us now calculate the number of topological equivalent diagrams. This will allow us to avoid calculating a large number of equivalent diagrams that all lead to the same contribution.

Let us pick one out of the set of topologically different diagrams. For each such diagram, we can exchange the interaction lines among each other. This creates $n!$ equivalent diagrams.

The two nodes of an interaction line can be exchanged without creating a topologically different diagram. This gives another 2^n factor for the number of equivalent diagrams.

This simplifies the diagrammatic rules, which lead to the interacting Green's function.

1. determine all topologically different diagrams.
2. for each propagating line, introduce a factor $G(\vec{x},t,\vec{x'},t')$ where the arguments $(\vec{x'},t')$ is the starting node and $(\vec{x},t)$ is the final node. (the factor $i\hbar$ has been compensated with the factor below)

3. for each non-propagating line, introduce a factor $-\frac{1}{i\hbar}\langle\hat{n}(\vec{r})\rangle$ instead of the Green's function. A non-propagating line is one connecting the two ends of a vertex or that forms a closed loop with itself.

4. for each interaction line include a factor $v_{int}(\vec{x},t,\vec{x'},t')$, where the arguments are taken from the two nodes bounding the vertex line. (the factor $\frac{1}{2}$ has been canceled by limiting to the topologically different diagrams.

5. Introduce the adiabatic factor $\mathrm{e}^{-\zeta|t|}$ for each node that is part of a vertex.

6. Integrate or sum over all coordinates (time,space,spin) of the full nodes The time integration goes over the entire time axis.

14.1 Linked Cluster theorem

Each diagram in the numerator can be expressed as a product of a diagram that connects the two arguments of the Green's function, the linked cluster, and all diagrams that are unconnected to the former. On the other hand the denominator is exactly composed of all unconnected diagrams.

It turns out that the unconnected diagrams can be factored out and cancel the denominator. This is the **linked cluster theorem**[74].

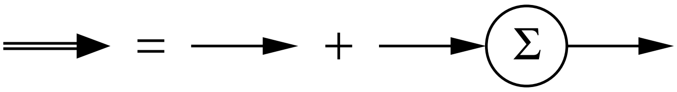
As a result, we can limit our diagrams to connected parts and we can drop the denominator.

14.2 Dyson's equation

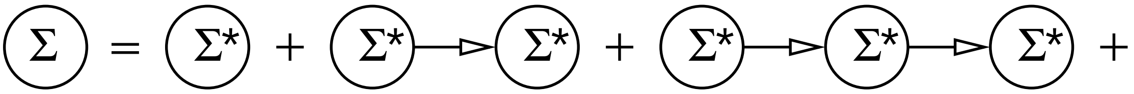
14.2.1 Self energy

We can further combine diagrams. Let us define the **self energy** .

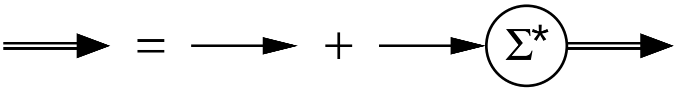
Each connected diagram can be divided into a starting Green's function, a middle part and a leaving Green's function. The sum over all middle parts is called the self energy $\Sigma(\vec{x},t,\vec{x'},t')$.



Next we define the proper self energy, which consists of all self-energy diagrams that cannot be divided into two parts by cutting through a single Green's function.



With the proper self energy we obtain the Green's function via



Equation of motion of the Green's function

The last equation results in an equation of motion of the Green's function:

$$G = G_0 + G_0\Sigma^* G$$

Written in a slightly more explicit form the equation reads as

$$\hat{G}(t,t') = \hat{G}_0(t,t') + \int dt'' \int dt''' \, \hat{G}_0(t,t'')\hat{\Sigma}^*(t''.t''')\hat{G}(t''',t')$$

we multiply this equation from the left with G_0^{-1} and from the right with G^{-1}.

$$G_0^{-1} = G^{-1} + \Sigma^* \qquad \Rightarrow \qquad \left(G_0^{-1} - \Sigma^*\right) G = 1$$

The inverse of the unperturbed Green's function is a differential operator

$$\left(i\hbar\partial_t - \hat{H}_0\right)\hat{G}_0(t,t') = \hat{1}\delta(t-t')$$

Thus, we obtain the

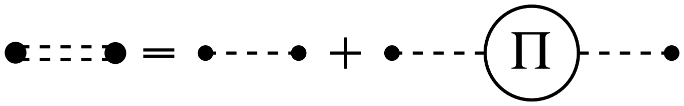
Differential equation for the interacting Green's function

$$\int dt'' \, \left[\delta(t-t'')\left(i\hbar\partial_{t''} - \hat{H}_0\right) - \hat{\Sigma}^*(t,t'')\right] G(t'',t') = \hat{1}\delta(t-t')$$

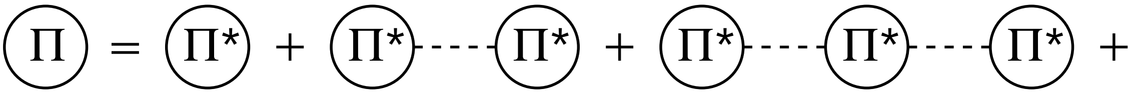
This equation is almost identical to the Differential equation for the non-interacting Green's function Eq. 12.23 on p.200. The difference is that the many-particle effects via the proper self-energy. The self-energy acts like a non-local, retarded[2] potential.

14.2.2 Polarization

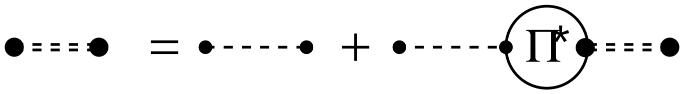
Next we define the **polarization**. We proceed as in the definition of the self energy. The polarization consist of all diagrams, between two nodes, that can each accept and donate a Green's function.



We define the **proper polarization** $\Pi^*(\vec{x},t,\vec{x'},t')$ as the sum of all diagrams, that cannot be divided into two parts by cutting through an interaction line.



With the proper polarization we obtain the full interaction line as



Fertig: WS06/07 25 Doppelstunde 14.Feb.07

---

[2] retarded means that it is non-local in the time arguments

Chapter 15

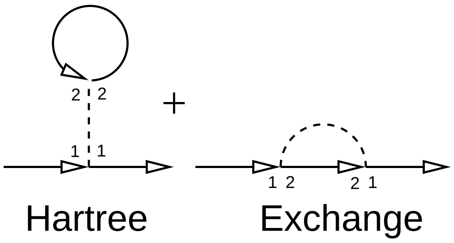
Some typical approximations

Everything from here on is very sketchy and may be wrong!

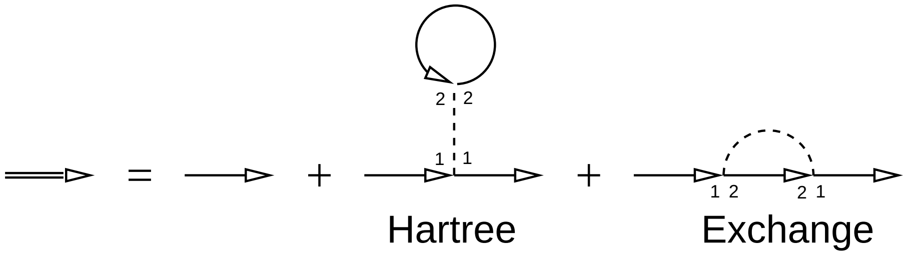
The diagrammatic technique allows to add up infinite sums of diagrams. Nevertheless there is also an infinite sum of diagrams that will be ignored. Furthermore not all partial sums converge. Thus, one chooses diagrams on the basis of physical arguments.

15.1 Hartree-Fock

The most simple diagrams for the Green's function are the terms to first order in the interaction. It contains all self-energy diagrams with exactly one interaction line, which can be identified by the Hartree term and the exchange term.



This leads to a Dyson equation for the Green's function



Let us work out the equation:

$$G(0,0') = G_0(0,0') + \int d1 \int d2\; \mathrm{e}^{-\zeta(|t_1|-|t_2|} G_0(0,1) v_{int}(1,2) \underbrace{\left(\frac{-n(2)}{i\hbar}\right)}_{FermiLoop} G_0(1,0')$$

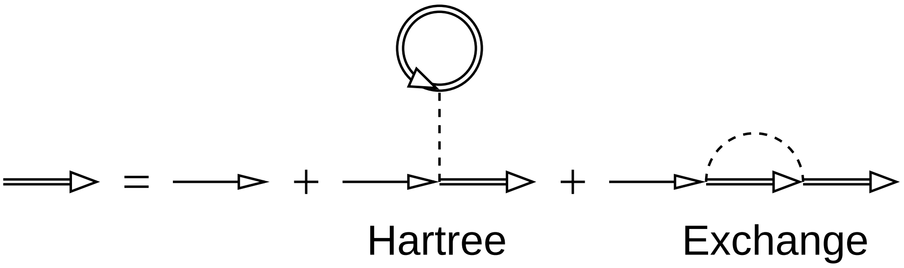
$$+ \int d1 \int d2\; \mathrm{e}^{-\zeta(|t_1|-|t_2|} G_0(0,1) v_{int}(1,2) G_0(1,2) G_0(2,0')$$

Let us look at the improper self energy

$$-\Sigma(1,1') = \delta(t_1 - t_1')\int d2\; v_{int}(1,2)\frac{n(2)}{i\hbar} - v_{int}(1,1')G_0(1,1')$$

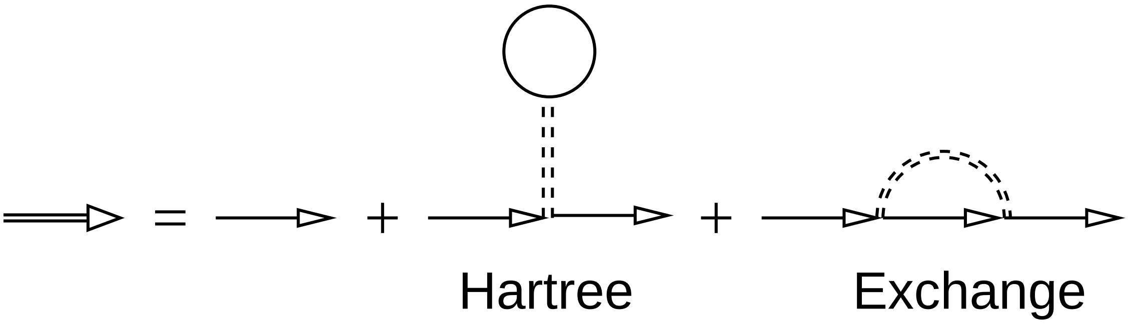
We insert the non-interacting Green's function $G_0(\vec{x},t,\vec{x'},t')$ from Eq. 12.22 on p. 199 and the interaction $v_{int}(\vec{x},t,\vec{x'},t')$ from Eq. 14.4 on p. 215

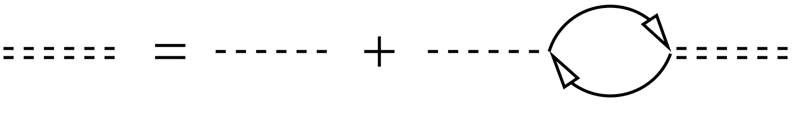
$$\begin{aligned}-i\hbar\Sigma(x,t,x',t') &= \delta(t-t')\delta(\vec{x}-\vec{x'})\int d^4x''\;\frac{e^2 n(x'')}{4\pi\epsilon_0|\vec{r}-\vec{r''}|}\\ &\quad-\theta(t-t')\sum_{n\in occ}\frac{e^2\phi(\vec{x})\phi(\vec{x'})}{4\pi\epsilon_0|\vec{r}-\vec{r''}|}+\theta(t'-t)\sum_{n\in unocc}\frac{e^2\phi(\vec{x})\phi(\vec{x'})}{4\pi\epsilon_0|\vec{r}-\vec{r''}|}\end{aligned}$$

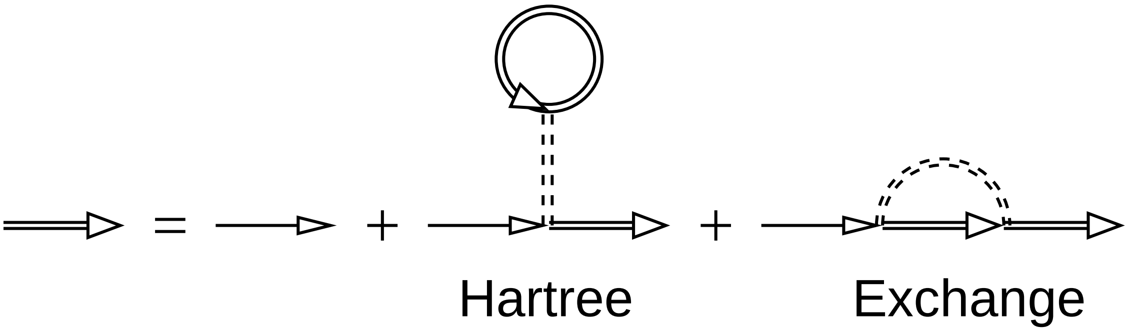


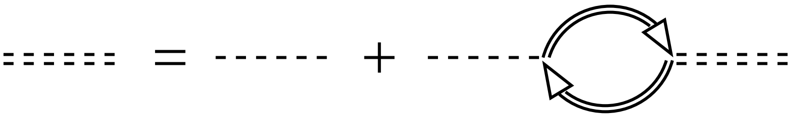
15.2 GW Approximation

The GW-approximation has been invented by Lars Hedin in 1965[75]. The name GW is an abbreviation of names as one may suspect, but the combination of the mathematical Symbols G for Green's function and W for the interaction operator from the German word “Wechselwirkung”.









======

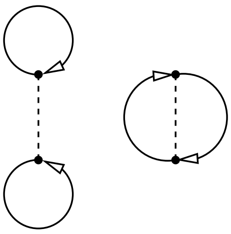
Each non-interacting Green's function is represented by an arrow. The arrow starts at the annihilator and it ends at the creator.

Consider the expectation value of an interaction term:

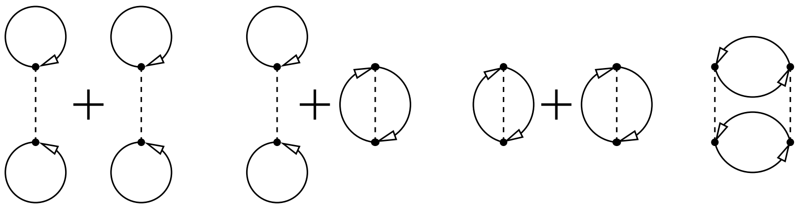
$$
\begin{aligned}
&W(\vec{x},t,\vec{x'},t')\langle\mathcal{G}^{(0)}|T[\hat{\psi}^\dagger(\vec{x},t)\hat{\psi}^\dagger(\vec{x'},t')\hat{\psi}(\vec{x},t)\hat{\psi}(\vec{x'},t')]|\mathcal{G}^{(0)}\rangle\\
=&W(\vec{x},t,\vec{x'},t')i\hbar G_0(\vec{x},t,\vec{x'},t')i\hbar G_0(\vec{x'},t',\vec{x},t)\\
-&W(\vec{x},t,\vec{x'},t')i\hbar G_0(\vec{x},t,\vec{x},t)i\hbar G_0(\vec{x'},t',\vec{x'},t')\\
=&(i\hbar)^2W(\vec{x},t,\vec{x'},t')\left[G_0(\vec{x},t,\vec{x'},t')G_0(\vec{x'},t',\vec{x},t)-G_0(\vec{x},t,\vec{x},t)G_0(\vec{x'},t',\vec{x'},t')\right]
\end{aligned}
$$

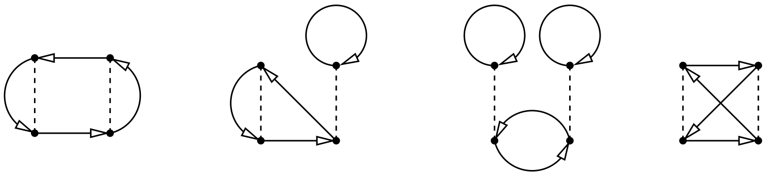
in consists of two creators and two annihilators. Thus, the time ordered product has two nodes. At each node a greens function starts and another one ends. Thus, we represent an interaction term by a dashed line and the corresponding Green's functions.

For example for the time ordered product of one interaction term we obtain terms of the type



For example for the time ordered product of two interaction terms we obtain terms of the type





We arrange the field operators such that annihilators and creators appear in alternating order. Then we form pairs between neighboring annihilators and creators. Each pair corresponds to a Green's function. Next we form all possible permutations of the creators among each other. A term with n Green's functions will have $n!$ such permutations. In the example of 4 Green's functions there are $4! = 24$ such terms. Each permutation corresponds to a Feynman diagram.

1	(1,2)(3,4)(5,6)(7,8)
2	(1,2)(3,4)(5,8)(7,6)
3	(1,2)(3,6)(5,4)(7,8)
4	(1,2)(3,6)(5,8)(7,4)
5	(1,2)(3,8)(5,4)(7,6)
6	(1,2)(3,8)(5,6)(7,4)
7	(1,4)(3,2)(5,6)(7,8)
8	(1,4)(3,2)(5,8)(7,6)
9	(1,4)(3,6)(5,2)(7,8)
10	(1,4)(3,6)(5,8)(7,2)
11	(1,4)(3,8)(5,2)(7,6)
12	(1,4)(3,8)(5,6)(7,2)
13	(1,6)(3,2)(5,4)(7,8)
14	(1,6)(3,2)(5,8)(7,4)
15	(1,6)(3,4)(5,6)(7,8)
16	(1,6)(3,4)(5,8)(7,6)
17	(1,6)(3,8)(5,2)(7,4)
18	(1,6)(3,8)(5,4)(7,2)
19	(1,8)(3,2)(5,4)(7,6)
20	(1,8)(3,2)(5,6)(7,4)
21	(1,8)(3,4)(5,2)(7,6)
22	(1,8)(3,4)(5,6)(7,2)
23	(1,8)(3,6)(5,2)(7,4)
24	(1,8)(3,6)(5,4)(7,2)

15.3 Random Phase approximation (RPA)

The **random phase approximation**(**RPA**) has been first introduced by Bohm and Pines[54] in a heuristic manner. The model was that most of the electrostatic interaction between electrons will be absorbed in a plasma-mode of collective oscillation, while the remaining modes can be considered as weakly interacting electrons with a screened interaction potential. Gell-Mann and Brueckner derived the random phase approximation by summing leading order chain Feynman diagrams[76]. A more systematic treatment that that of Bohm and Pines has been given by Nozieres and Pines[**?**](see [77]).

Chapter 16

Models in many-particle theory

16.1 Dominant interactions

$$\hat{W} \overset{\text{Eq. 8.9}}{=} \frac{1}{2}\sum_{i,j,k,l} W_{a,b,c,d}\hat{c}_a^\dagger \hat{c}_b^\dagger \hat{c}_d \hat{c}_c$$

$$W_{a,b,c,d} \overset{\text{Eq. 8.3}}{=} \int d^4x \int d^4x'\ \phi_a^*(\vec{x})\phi_b^*(\vec{x'})\frac{e^2}{4\pi|\vec{r}-\vec{r'}|}\phi_d(\vec{x})\phi_c(\vec{x'})$$

We choose spin orbitals, which are eigenstates of s_z, i.e.

$$\langle\vec{r},\sigma|\phi_{n_a,\sigma_a}\rangle = \delta_{\sigma,\sigma_a}\phi_n(\vec{r})$$

This already limits the number of non-vanishing matrix elements

$$\hat{W} \overset{\text{Eq. 8.9}}{=} \frac{1}{2}\sum_{i,j,k,l} W_{a,b,c,d}\hat{c}_a^\dagger \hat{c}_b^\dagger \hat{c}_d \hat{c}_c$$

$$W_{a,b,c,d} \overset{\text{Eq. 8.3}}{=} \delta_{\sigma_a,\sigma_d}\delta_{\sigma_b,\sigma_c}\int d^3r \int d^3r'\ \phi_{n_a}^*(\vec{r})\phi_{n_d}(\vec{r})\frac{e^2}{4\pi|\vec{r}-\vec{r'}|}\phi_{n_b}^*(\vec{r'})\phi_{n_c}(\vec{r'})$$

Now we pick out only the dominant matrix elements. These are the Coulomb integrals $K_{i,j}$ and the exchange integrals $J_{i,j}$. (Note that chemists use the symbols in the opposite way.)

$$K_{i,j} = \frac{1}{2}\int d^3r \int d^3r'\ \phi_i^*(\vec{r})\phi_i(\vec{r})\frac{e^2}{4\pi|\vec{r}-\vec{r'}|}\phi_j^*(\vec{r'})\phi_j(\vec{r'})$$

$$J_{i,j} = \frac{1}{2}\int d^3r \int d^3r'\ \phi_i^*(\vec{r})\phi_j(\vec{r})\frac{e^2}{4\pi|\vec{r}-\vec{r'}|}\phi_j^*(\vec{r'})\phi_i(\vec{r'})$$

The rationale for the selection is that the matrix element is the Coulomb interaction of two densities. If the orbitals $|\phi(\vec{r})\rangle$ are orthonormal, each of the density is charge neutral unless the two orbitals are identical. Thus, the dominant contribution is $K_{i,j}$ where the densities are not neutral. The Matrix elements $K_{i,j}$ are long-ranged, but can be approximated as

$$K_{i,j} \overset{|R_i-R_j|\to\infty}{\to} \frac{e^2}{4\pi|\vec{R}_i-\vec{R}_j|}$$

where R_i and $\vec{R}_j$ are centers of the respective orbitals.

The exchange matrix elements $J_{i,j}$ describe the electrostatic interaction of neutral densities. Thus, they contribute effectively only if the two densities overlap. Considering local orbitals it makes sense to consider only the interaction of identical pairs. Note, however, that this limitation becomes invalid as we take more and more orbitals.

If we limit the interaction matrix elements only to those terms, we obtain

$$W_{a,b,c,d} = \delta_{\sigma_a,\sigma_d}\delta_{\sigma_b,\sigma_c}\Big(\delta_{n_a,n_d}\delta_{n_b,n_c}K_{n_a,n_b} + \delta_{n_a,n_c}\delta_{n_b,n_d}(1-\delta_{n_a,n_b})J_{n_a,n_b}\Big)$$

The factor $(1-\delta_{n_a,n_b})$ takes into account that $K_{i,i} = J_{i,i}$, in order to avoid double counting.

Thus, the dominant part of the interaction has the form[1]

$$\begin{aligned}
\hat{W} &= \frac{1}{2}\sum_{\sigma,\sigma'}\sum_{i,j} K_{i,j}\hat{c}^\dagger_{i,\sigma}\hat{c}^\dagger_{j,\sigma'}\hat{c}_{j,\sigma'}\hat{c}_{i,\sigma} + \frac{1}{2}\sum_{\sigma,\sigma'}\sum_{i\neq j} J_{i,j}\hat{c}^\dagger_{i,\sigma}\hat{c}^\dagger_{j,\sigma'}\hat{c}_{i,\sigma'}\hat{c}_{j,\sigma} \\
&= \frac{1}{2}\sum_{\sigma,\sigma'}\sum_{i,j} K_{i,j}\hat{c}^\dagger_{i,\sigma}\hat{c}_{i,\sigma}\hat{c}^\dagger_{j,\sigma'}\hat{c}_{j,\sigma'} - \frac{1}{2}\sum_{\sigma,\sigma'}\sum_{i\neq j} J_{i,j}\hat{c}^\dagger_{i,\sigma}\hat{c}_{i,\sigma'}\hat{c}^\dagger_{j,\sigma'}\hat{c}_{j,\sigma} \\
&- \frac{1}{2}\sum_{\sigma}\sum_{i} K_{i,i}\hat{c}^\dagger_{i,\sigma}\hat{c}_{i,\sigma} \\
&= \frac{1}{2}\sum_{\sigma,\sigma'}\sum_{i\neq j} K_{i,j}\hat{c}^\dagger_{i,\sigma}\hat{c}_{i,\sigma}\hat{c}^\dagger_{j,\sigma'}\hat{c}_{j,\sigma'} - \frac{1}{2}\sum_{\sigma,\sigma'}\sum_{i\neq j} J_{i,j}\hat{c}^\dagger_{i,\sigma}\hat{c}_{i,\sigma'}\hat{c}^\dagger_{j,\sigma'}\hat{c}_{j,\sigma} \\
&+ \sum_{i} K_{i,i}\hat{c}^\dagger_{i,\uparrow}\hat{c}_{i,\uparrow}\hat{c}^\dagger_{i,\downarrow}\hat{c}_{i,\downarrow}
\end{aligned}$$

The parameter $K_{i,i}$ is the so-called **U-parameter**

$$U \stackrel{\text{def}}{=} K_{i,i}$$

It will be helpful to replace the creation and annihilation operators by orbital occupation and spin operators.

$$\begin{aligned}
\hat{n}_i &= \hat{c}^\dagger_{i,\uparrow}\hat{c}_{i,\uparrow} + \hat{c}^\dagger_{i,\downarrow}\hat{c}_{i,\downarrow} \\
\hat{\vec{S}}_i &= \frac{\hbar}{2}\begin{pmatrix}\hat{c}^\dagger_{i,\uparrow} \\ \hat{c}^\dagger_{i,\downarrow}\end{pmatrix}\boldsymbol{\vec{\sigma}}\begin{pmatrix}\hat{c}^\dagger_{i,\uparrow} \\ \hat{c}^\dagger_{i,\downarrow}\end{pmatrix} \\
\hat{S}_{i,x} &= \frac{\hbar}{2}\left(\hat{c}^\dagger_{i,\uparrow}\hat{c}_{i,\downarrow} + \hat{c}^\dagger_{i,\downarrow}\hat{c}_{i,\uparrow}\right) \\
\hat{S}_{i,y} &= -i\frac{\hbar}{2}\left(\hat{c}^\dagger_{i,\uparrow}\hat{c}_{i,\downarrow} - \hat{c}^\dagger_{i,\downarrow}\hat{c}_{i,\uparrow}\right) \\
\hat{S}_{i,z} &= \frac{\hbar}{2}\left(\hat{c}^\dagger_{i,\uparrow}\hat{c}_{i,\uparrow} - \hat{c}^\dagger_{i,\downarrow}\hat{c}_{i,\downarrow}\right) \\
\hat{S}_{i,+} &= \hat{S}_{i,x} + i\hat{S}_{i,y} = \hbar\hat{c}^\dagger_{i,\uparrow}\hat{c}_{i,\downarrow} \\
\hat{S}_{i,-} &= \hat{S}_{i,x} - i\hat{S}_{i,y} = \hbar\hat{c}^\dagger_{i,\downarrow}\hat{c}_{i,\uparrow}
\end{aligned}$$

---

[1]

$$\begin{aligned}
\hat{c}^\dagger_{i,\sigma}\hat{c}^\dagger_{j,\sigma'}\hat{c}_{j,\sigma'}\hat{c}_{i,\sigma} &= -\hat{c}^\dagger_{i,\sigma}\hat{c}^\dagger_{j,\sigma'}\hat{c}_{i,\sigma}\hat{c}_{j,\sigma'} = -\delta_{i,j}\delta_{\sigma,\sigma'}\hat{c}^\dagger_{i,\sigma}\hat{c}_{j,\sigma'} + \hat{c}^\dagger_{i,\sigma}\hat{c}_{i,\sigma}\hat{c}^\dagger_{j,\sigma'}\hat{c}_{j,\sigma'} \\
\hat{c}^\dagger_{i,\sigma}\hat{c}^\dagger_{j,\sigma'}\hat{c}_{i,\sigma'}\hat{c}_{j,\sigma} &= \hat{c}^\dagger_{i,\sigma}\delta_{i,j}\hat{c}_{j,\sigma}\hat{c}^\dagger_{i,\sigma}\hat{c}_{i,\sigma'}\hat{c}^\dagger_{j,\sigma'}\hat{c}_{j,\sigma}
\end{aligned}$$

Thus, we obtain

$$
\begin{aligned}
\sum_{\sigma,\sigma'} \hat{c}^\dagger_{i,\sigma}\hat{c}_{i,\sigma'}\hat{c}^\dagger_{j,\sigma'}\hat{c}_{j,\sigma} &= \Big(\frac{1}{2}\hat{n}_i+\frac{1}{\hbar}\hat{S}_{i,z}\Big)\Big(\frac{1}{2}\hat{n}_j+\frac{1}{\hbar}\hat{S}_{j,z}\Big)+\frac{1}{\hbar^2}\hat{S}_{i,+}\hat{S}_{j,-}+\frac{1}{\hbar^2}\hat{S}_{i,-}\hat{S}_{j,+}+\Big(\frac{1}{2}\hat{n}_i-\frac{1}{\hbar}\hat{S}_{i,z}\Big)\Big(\frac{1}{2}\hat{n}_j-\frac{1}{\hbar}\hat{S}_{j,z}\Big)\\
&= \frac{1}{2}\hat{n}_i\hat{n}_j+\frac{1}{\hbar^2}\big(\hat{S}_{i,x}+i\hat{S}_{i,y}\big)\big(\hat{S}_{j,x}-i\hat{S}_{j,y}\big)+\frac{1}{\hbar^2}\big(\hat{S}_{i,x}-i\hat{S}_{i,y}\big)\big(\hat{S}_{j,x}+i\hat{S}_{j,y}\big)-\frac{2}{\hbar^2}\hat{S}_{i,z}\hat{S}_{j,z}\\
&= \frac{1}{2}\hat{n}_i\hat{n}_j+\frac{2}{\hbar^2}\hat{S}_{i,x}\hat{S}_{j,x}+\frac{2}{\hbar^2}\hat{S}_{i,y}\hat{S}_{j,y}+\frac{2}{\hbar^2}\hat{S}_{i,z}\hat{S}_{j,z}\\
&= \frac{1}{2}\hat{n}_i\hat{n}_j+\frac{2}{\hbar^2}\hat{\vec{S}}_i\hat{\vec{S}}_j
\end{aligned}
$$

Thus, we obtain

$$
\begin{aligned}
\hat{W} &= \frac{1}{2}\sum_{i\neq j}K_{i,j}\hat{n}_i\hat{n}_j-\frac{1}{2}\sum_{i\neq j}J_{i,j}\Big(\frac{1}{2}\hat{n}_i\hat{n}_j+\frac{2}{\hbar^2}\hat{\vec{S}}_i\hat{\vec{S}}_j\Big)+\sum_i K_{i,i}\hat{c}^\dagger_{i,\uparrow}\hat{c}_{i,\uparrow}\hat{c}^\dagger_{i,\downarrow}\hat{c}_{i,\downarrow}\\
&= \frac{1}{2}\sum_{i\neq j}\Big(K_{i,j}-J_{i,j}\Big)\hat{n}_i\hat{n}_j-\frac{1}{\hbar^2}\sum_{i\neq j}J_{i,j}\hat{\vec{S}}_i\hat{\vec{S}}_j+\sum_i K_{i,i}\hat{c}^\dagger_{i,\uparrow}\hat{c}_{i,\uparrow}\hat{c}^\dagger_{i,\downarrow}\hat{c}_{i,\downarrow}
\end{aligned}
$$

The last term is the **Hubbard interaction**

HUBBARD INTERACTION

$$
\hat{W}_{Hubbard}\stackrel{\text{def}}{=}\sum_i K_{i,i}\hat{c}^\dagger_{i,\uparrow}\hat{c}_{i,\uparrow}\hat{c}^\dagger_{i,\downarrow}\hat{c}_{i,\downarrow}
$$

The Hubbard interaction describes the Coulomb repulsion of two electrons in the same spatial orbital.

HUND COUPLING

The interaction term proportional to the exchange integrals J is called Hund coupling.

$$
\hat{W}_{Hund}\stackrel{\text{def}}{=}-\frac{1}{2}\sum_{i\neq j}J_{i,j}\hat{n}_i\hat{n}_j-\frac{1}{\hbar^2}\sum_{i\neq j}J_{i,j}\hat{\vec{S}}_i\hat{\vec{S}}_j
$$

Hund's coupling causes spins to align parallel.

16.2 Hubbard model

The original Hubbard model was suggested in 1963 independently by Hubbard[**?**], Kanamori and Gutzwiller[78]. It describes a lattice of atoms with a single orbital on each atom, which has a large Coulomb repulsion on each site.

$$
\hat{H}=-t\sum_i\sum_\sigma\Big(\hat{c}^\dagger_{i,\sigma}\hat{c}_{i+1,\sigma}\hat{c}^\dagger_{i+1,\sigma}\hat{c}_{i+1,\sigma}\Big)+U\sum_i n_{i,\uparrow}n_{i,\downarrow}+V\sum_i n_i n_{i,i+1}
$$

t is the hopping integral and U is the Coulomb parameter. V is the Coulomb repulsion of electrons on two neighboring sites.

The one-dimensional Hubbard model has been solved by Lieb and Wu[79]. For half-filling the Hubbard model describes a Mott insulator.

16.3 t-J model

`http://www-public.tu-bs.de:8080/~honecker/rgp/q-mod/`

Nehmen wir an dass die Hubbard Wechselwirkung so gross ist, dass Zustaende mit einer Doppelbesetzung eines Gitterplatzes nicht mehr beitragen, dann erhält man aus dem Allgemeinen Modell das t-J Modell.

$$H = -tP\sum_i\sum_\sigma\left(\hat{c}^\dagger_{i,\sigma}\hat{c}_{i+1,\sigma}\hat{c}^\dagger_{i+1,\sigma}\hat{c}_{i+1,\sigma}\right)P + J\sum_i \vec{S}_i\vec{S}_{i+1}$$

where P projects out doubly occupied sites.

- For half filling the t-J model reduces to the heisenberg model
- The hubbard model for U»t is equivalent to the t-J model. In second order perturbation theory one finds $J = \frac{4t^2}{U}$

16.4 Heisenberg model

$$\hat{H} = J\sum_i \vec{S}_i\vec{S}_{i+1}$$

16.5 Anderson impurity model

The Anderson model[80], named after its inventor P.W. Anderson[2], describes an impurity of large Coulomb repulsion in a normal metal. This model explains the formation of a local magnetic moment on the impurity atom.

The Anderson Hamiltonian has the form

$$H = \sum_{\vec{k},\sigma}\epsilon_k\hat{n}_{\vec{k},\sigma} + \sum_\sigma \epsilon_d\hat{n}_{d,\sigma} - \Delta\sum_{\vec{k}}\left(\hat{c}^\dagger_{k,\sigma}\hat{d}_\sigma + \hat{d}^\dagger_\sigma\hat{c}_{k,\sigma}\right) + U\hat{n}_{d,\uparrow}\hat{n}_{d,\downarrow}$$

We used the creation and annihilation operators $\hat{d}^\dagger_\sigma$ and $\hat{d}_\sigma$ in the localized impurity level and its number operator $\hat{n}_{d,\sigma} = \hat{d}^\dagger_\sigma\hat{d}_\sigma$. The creation and annihilation operators in the Bloch states of the normal metal are $\hat{c}^\dagger_{\vec{k},\sigma}$ and $\hat{c}_{\vec{k},\sigma}$, and the corresponding number operator is $\hat{n}_{\vec{k},\sigma} := \hat{c}^\dagger_{\vec{k},\sigma}\hat{c}_{\vec{k},\sigma}$. The index σ denotes the spin quantum number, which can have values ↑ or ↓.

The first term describes the one-particle energies of the normal metal and the second describes the one-particle levels of the localized impurity. The third term describes the coupling between the impurity and the band states. So far all terms are simple one-electron terms. The many-particle problem enters with the last term that describes the Coulomb repulsion of the electrons in the localized orbital.

We could as well write the Hamiltonian as

$$\begin{aligned} H = &\sum_{i,\sigma}\epsilon_i\hat{c}^\dagger_{i,\sigma}\hat{c}_{i,\sigma} + \sum_{i,j,\sigma}t_{i,j}\left(\hat{c}^\dagger_{i,\sigma}\hat{c}_{j,\sigma} + \hat{c}_{i,\sigma}\hat{c}^\dagger_{j,\sigma}\right) + \sum_\sigma \epsilon_d\hat{d}^\dagger_\sigma\hat{d}_\sigma \\ &-\Delta\sum_{i,\sigma}\left(\hat{c}^\dagger_{i,\sigma}\hat{d}_\sigma + \hat{d}^\dagger_\sigma\hat{c}_{i,\sigma}\right) + U\hat{n}_{d,\uparrow}\hat{n}_{d,\downarrow} \end{aligned}$$

Here, we consider the Wannier functions localized on the sites of the normal metal, which make up the Bloch states. The creation and annihilation operators are denoted by $\hat{c}^\dagger_{i,\sigma}$ and $\hat{c}_{i,\sigma}$.

The Anderson model includes normal metal states also at the site of the impurity. This makes sense because also a transition metal has s- and p-electrons, that however lie at different energies.

---

[2] P.W. Anderson, American physicist born 1923. Obtained the Nobel price in physics 1977. Anderson localization describes how electron states become localized in an disordered material. more...

16.5.1 The interaction energy

The U-parameter is defined as[80]

$$U = \frac{e^2}{4\pi\epsilon_0} \int d^3r \int d^3r' \frac{\phi_d^*(\vec{r})\phi_d(\vec{r})\phi_d^*(\vec{r'})\phi_d(\vec{r'})}{|\vec{r}-\vec{r'}|}$$

where $\phi_d(\vec{r})$ is the wave function of a localized orbital.

According to Eq. 8.9 the interaction energy has the form

$$\hat{W} = \frac{1}{2}\sum_{i,j,k,l} W_{i,j,k,l}\hat{c}_i^\dagger \hat{c}_j^\dagger \hat{c}_l \hat{c}_k$$

where $W_{i,j,k,l}$ is given,according to Eq. 8.3, by

$$W_{i,j,k,l} = \frac{e^2}{4\pi\epsilon_0} \int d^4x \int d^4x' \ \frac{\phi_i^*(\vec{x})\phi_j^*(\vec{x'})\phi_k(\vec{x})\phi_l^*(\vec{x'})}{|\vec{r}-\vec{r'}|}$$

In our case we just have two states, one for the spin up and one for spin down electron. Thus, the orbital indices can assume two values $i =\uparrow$ and $i =\downarrow$. The spinor wave functions are

$$\phi_\sigma(\vec{r},\sigma') = \phi_d(\vec{r})\delta_{\sigma,\sigma'}$$

where $\phi_d(\vec{r})$ is simply the spatial part of the orbital. The product of two spinors, that are spin-eigenstates, is only non-zero, when both wave functions have the same spin.

$$\begin{aligned}
W_{\sigma_1,\sigma_2,\sigma_3,\sigma_4} &= \frac{e^2}{4\pi\epsilon_0}\sum_{\sigma,\sigma'} \int d^4r \int d^4r' \ \frac{\overbrace{\phi_d^*(\vec{r})\delta(\sigma_1,\sigma)}^{\phi^*_{\sigma_1}(\vec{x})}\overbrace{\phi_d^*(\vec{r'})\delta(\sigma_2,\sigma')}^{\phi^*_{\sigma_2}(\vec{x'})}\overbrace{\phi_d(\vec{r})\delta(\sigma_3,\sigma)}^{\phi_{\sigma_3}(\vec{x})}\overbrace{\phi_l^*(\vec{r'})\delta(\sigma_4,\sigma')}^{\phi_{\sigma_4}(\vec{x'})}}{|\vec{r}-\vec{r'}|} \\
&= \frac{e^2}{4\pi\epsilon_0}\sum_{\sigma,\sigma'}\delta(\sigma_1,\sigma)\delta(\sigma_2,\sigma')\delta(\sigma_3,\sigma)\delta(\sigma_4,\sigma') \int d^4r \int d^4r' \ \frac{\phi_d^*(\vec{r})\phi_d^*(\vec{r'})\phi_d(\vec{r})\phi_l^*(\vec{r'})}{|\vec{r}-\vec{r'}|} \\
&= U\delta(\sigma_1,\sigma_3)\delta(\sigma_2,\sigma_4) \\
\hat{W} &= \frac{U}{2}\sum_{\sigma,\sigma'}\hat{d}_\sigma^\dagger\hat{d}_{\sigma'}^\dagger\hat{d}_{\sigma'}\hat{d}_\sigma = \frac{U}{2}\sum_{\sigma,\sigma'}\left[-\hat{d}_\sigma^\dagger\delta_{\sigma',\sigma}\hat{d}_{\sigma'} + \hat{d}_\sigma^\dagger\hat{d}_\sigma\hat{d}_{\sigma'}^\dagger\hat{d}_{\sigma'}\right] \\
&= -\frac{U}{2}\left(\hat{n}_{d,\uparrow}+\hat{n}_{d,\downarrow}\right) + \frac{U}{2}\sum_{\sigma,\sigma'}\hat{n}_{d,\sigma}\hat{n}_{d,\sigma'} \\
&= -\frac{U}{2}\left(\hat{n}_{d,\uparrow}+\hat{n}_{d,\downarrow}\right) + \frac{U}{2}\left(\hat{n}_{d,\uparrow}^2+\hat{n}_{d,\uparrow}\hat{n}_{d,\downarrow}+\hat{n}_{d,\downarrow}\hat{n}_{d,\uparrow}+\hat{n}_{d,\downarrow}^2\right) \\
&= U\hat{n}_{d,\downarrow}\hat{n}_{d,\uparrow}
\end{aligned}$$

We used that $\hat{n}_{d,\sigma}^2 = \hat{n}_{d,\sigma}$ and that $\hat{n}_{d,\uparrow}\hat{n}_{d,\downarrow} = \hat{n}_{d,\downarrow}\hat{n}_{d,\uparrow}$.

16.5.2 Mean field approximation

Editor: Not ready yet

We can develop a mean field solution by replacing

$$\hat{n}_{d,\sigma} = \langle\hat{n}_{d,\sigma}\rangle + \left(\hat{n}_{d,\sigma} - \langle\hat{n}_{d,\sigma}\rangle\right)$$

Chapter 17

Zoo of Concepts from Solid State Physics

This is nothing than a collection of material for myself.

- Bloch electron (<-> Fermi liquid?), hole
- phonon
- photon,
- magnon, spin wave, Bloch[81]
- polaron, Zener polaron (phonon+electron)
- Polariton (light-phonon, light-plasmon)
- Soliton
- plasmon
- exciton (electron-hole)
- First order/second order phase transition
- Jahn-Teller effect
- Peierls distortion
- Nesting
- Orbital ordering
- Fermi liquid
- Luttinger liquid
- Mott insulator/ Band insulator
- Heavy Fermion
- Superconductivity
- Kondo Effect
- Stoner Magnetism

- Super exchange
- Double exchange
- Multiferroics (See article by Khomskii[82])
- Pseudogap
- Crystal Field splitting, 10Dt (M. Lezaic, in *Electronic Oxides*. Lecture notes of the 41th IFF Spring school 2010, Stefan Blüegel et al, Eds., Schriften des Forschungszentrums Jülich)
- Topological insulator
- Anderson Localization
- Conductance quantization
- Quantum Hall effect
- Pomeranchuk instability
- van Hove Singularity
- quantum critical phase: Popular introduction

Models

- Ising model
- t-J-model
- Heisenberg model
- Hubbard model

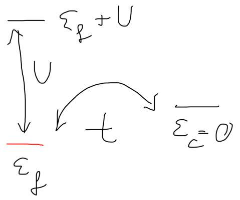
Methods

- Mean Field approximation
- saddle point approximation
- DCA
- Coherent Potential approximation
- Virtual Potential approximation
- Lanczos method, exact diagonalization
- NRG
- Dynamical matrix renormalization theory (DMRG): [83]. See also K.H. Marti and M. Reiher, in Progress in Physical Chemistry Vol. 3, *Modern and Universal First-principles Methods for Many-electron Systems in Chemistry and Physics*, (Oldenburg Verlag, p293)
- Dynamical mean-field theory (DMFT)
- reduced-density-matrix functional theory (rDMFT)
- Quantum Monte Carlo

17.1 Kondo effect

This is not finished. It only contains notes after a talk.

The model for the Kondo effect is a two site model, where one represents a impurity orbital with strong Coulomb interaction and the second orbital represents the bath.



Let us construct first the basis states in terms of eigenstates for particle number and spin, which are conserved quantities in thus model.

$$\begin{array}{rc}
|\mathcal{O}\rangle & N=0,\ S=0,\ S_z=0,\ E=0 \\
\hat{f}^\dagger_\uparrow|\mathcal{O}\rangle & N=1,\ S=\frac{1}{2},\ S_z=\frac{1}{2} \\
\hat{f}^\dagger_\downarrow|\mathcal{O}\rangle & N=1,\ S=\frac{1}{2},\ S_z=-\frac{1}{2} \\
\hat{c}^\dagger_\uparrow|\mathcal{O}\rangle & N=1,\ S=\frac{1}{2},\ S_z=\frac{1}{2} \\
\hat{c}^\dagger_\downarrow|\mathcal{O}\rangle & N=1,\ S=\frac{1}{2},\ S_z=-\frac{1}{2} \\
\hat{f}^\dagger_\uparrow\hat{f}^\dagger_\downarrow|\mathcal{O}\rangle & N=2,\ S=0,\ S_z=0 \\
\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow|\mathcal{O}\rangle & N=2,\ S=0,\ S_z=0 \\
\hat{f}^\dagger_\uparrow\hat{c}^\dagger_\uparrow|\mathcal{O}\rangle & N=2,\ S=1,\ S_z=1,\ E=\epsilon_f+\epsilon_c \\
\hat{f}^\dagger_\downarrow\hat{c}^\dagger_\downarrow|\mathcal{O}\rangle & N=2,\ S=1,\ S_z=-1,\ E=\epsilon_f+\epsilon_c \\
\frac{1}{\sqrt{2}}\left(\hat{f}^\dagger_\uparrow\hat{c}^\dagger_\downarrow+\hat{f}^\dagger_\downarrow\hat{c}^\dagger_\uparrow\right)|\mathcal{O}\rangle & N=2, S=1,\ S_z=0 \\
\frac{1}{\sqrt{2}}\left(\hat{f}^\dagger_\uparrow\hat{c}^\dagger_\downarrow-\hat{f}^\dagger_\downarrow\hat{c}^\dagger_\uparrow\right)|\mathcal{O}\rangle & N=2,\ S=0,\ S_z=0 \\
\hat{f}^\dagger_\uparrow\hat{f}^\dagger_\downarrow\hat{c}^\dagger_\uparrow|\mathcal{O}\rangle & N=3,\ S=\frac{1}{2},\ S_z=\frac{1}{2} \\
\hat{f}^\dagger_\uparrow\hat{f}^\dagger_\downarrow\hat{c}^\dagger_\downarrow|\mathcal{O}\rangle & N=3,\ S=\frac{1}{2},\ S_z=-\frac{1}{2} \\
\hat{f}^\dagger_\uparrow\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow|\mathcal{O}\rangle & N=3,\ S=\frac{1}{2},\ S_z=+\frac{1}{2} \\
\hat{f}^\dagger_\downarrow\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow|\mathcal{O}\rangle & N=3,\ S=\frac{1}{2},\ S_z=-\frac{1}{2} \\
\hat{f}^\dagger_\uparrow\hat{f}^\dagger_\downarrow\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow|\mathcal{O}\rangle & N=4,\ S=0,\ S_z=0\ E=2\epsilon_f+2\epsilon_c+U
\end{array}$$

A check of this analysis is that there are again $2^4=16$ states.

In the one-particle channel the resulting states are the bonding and antibonding one-particle states.

$$\left(\hat{f}^\dagger_\sigma a+\hat{c}^\dagger_\sigma b\right)|\mathcal{O}\rangle$$

In the two-particle channel, there are three singlet states

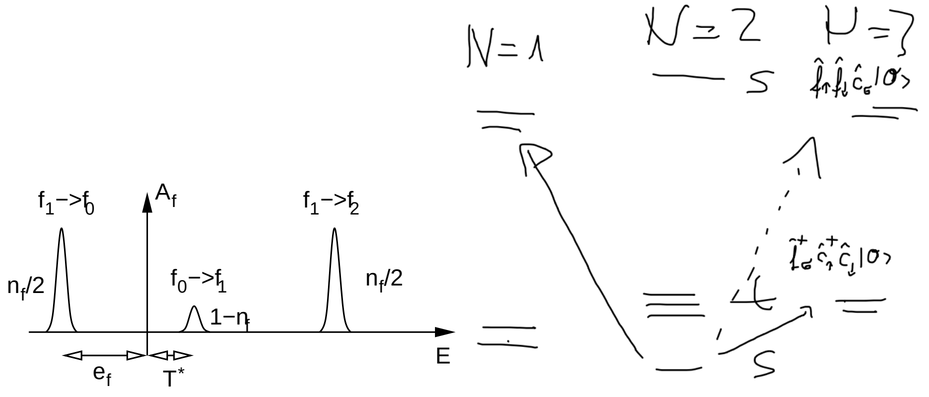
$$\left[\hat{f}^\dagger_\uparrow\hat{f}^\dagger_\downarrow a+\hat{c}^\dagger_\uparrow\hat{c}^\dagger_\downarrow b+\frac{1}{\sqrt{2}}\left(\hat{f}^\dagger_\uparrow\hat{c}^\dagger_\downarrow-\hat{f}^\dagger_\downarrow\hat{c}^\dagger_\uparrow\right)c\right]|\mathcal{O}\rangle$$

In the three-particle channel we obtain again bonding and antibonding orbitals

$$\hat f_\uparrow \hat c_\uparrow \left(\hat f^\dagger_\downarrow a + \hat c^\dagger_\downarrow b\right)|\mathcal{O}\rangle$$
$$\hat f_\downarrow \hat c_\downarrow \left(\hat f^\dagger_\uparrow a + \hat c^\dagger_\uparrow b\right)|\mathcal{O}\rangle$$

We see that the singlet two-particle state lies below the triplet state, because only the singlet state can hybridize with higher lying states, when t is switched on. A simplification the infinite-U limit, in which case we may exclude any states with two electrons on the impurity. The stabilization is mainly due to the delocalization of the electron in one spin direction between impurity and bath. The Coulomb repulsion is a secondary effect.

When we investigate the spectra for adding and removing electrons, we will find one transition to



Increasing the temperature reduces the singlet occupation which results in larger resistance.

Kondo Resonance=Abrikosov-Suhl resonance.

17.1.1 Kondo model

Not finished

$$\hat H = \underbrace{\sum_{k,\sigma} \epsilon_{\vec k} \hat c^\dagger_{\vec k,\sigma} \hat c_{\vec k,\sigma}}_{\text{bath}} + \underbrace{\sum_{k,\sigma} V_{\vec k}\left(\hat d^\dagger_\sigma \hat c_{\vec k,\sigma} + \hat c^\dagger_{\vec k,\sigma} \hat d_\sigma\right)}_{\text{hybridization}} + \underbrace{\sum_{\sigma} \epsilon_d \hat d^\dagger_\sigma \hat d_\sigma + U \hat d^\dagger_\uparrow \hat d_\uparrow \hat d^\dagger_\downarrow \hat d_\downarrow}_{\text{impurity}}$$

This model can be solved by dividing the Hilbert state into subsets with 0,1,2, electrons on the impurity.

The projection operators onto states with zero, one or two particles on the impurity is

$$\hat P_0 = \left(\hat 1 - \hat d^\dagger_\uparrow \hat d_\uparrow\right)\left(\hat 1 - \hat d^\dagger_\downarrow \hat d_\downarrow\right)$$
$$\hat P_1 = \left(\hat 1 - \hat d^\dagger_\uparrow \hat d_\uparrow\right)\hat d^\dagger_\downarrow \hat d_\downarrow + \hat d^\dagger_\uparrow \hat d_\uparrow\left(\hat 1 - \hat d^\dagger_\downarrow \hat d_\downarrow\right)$$
$$\hat P_2 = \hat d^\dagger_\uparrow \hat d_\uparrow \hat d^\dagger_\downarrow \hat d_\downarrow$$

The projection operators can be rewritten in the form

$$\begin{aligned}
\hat{P}_0 &= \hat{1} - \sum_\sigma \hat{d}^\dagger_\sigma \hat{d}_\sigma + \hat{P}_2 \\
\hat{P}_1 &= \sum_\sigma \hat{d}^\dagger_\sigma \hat{d}_\sigma - 2\hat{P}_2 \\
\hat{P}_2 &= \hat{d}^\dagger_\uparrow \hat{d}_\uparrow \hat{d}^\dagger_\downarrow \hat{d}_\downarrow
\end{aligned}$$

One can easily show that the three projection operators add up to the unity operator.

$$\sum_{j=0}^{2} \hat{P}_j = \hat{1}$$

This allows us to decompose the Hamiltonian according to

$$\hat{H} = \sum_{i,j=0}^{2} \underbrace{\hat{P}_i \hat{H} \hat{P}_j}_{\hat{H}_{i,j}} = \sum_{i,j=0}^{2} \hat{H}_{i,j}$$

The individual blocks are

$$\begin{aligned}
\hat{H}_{00} &= \sum_{k,\sigma} \epsilon_{\vec{k}} \hat{c}^\dagger_{\vec{k},\sigma} \hat{c}_{\vec{k},\sigma} \hat{P}_0 = \hat{H}_{\text{bath}} \hat{P}_0 \\
\hat{H}_{11} &= \left(\sum_{k,\sigma} \epsilon_{\vec{k}} \hat{c}^\dagger_{\vec{k},\sigma} \hat{c}_{\vec{k},\sigma} + \epsilon_d \right) \hat{P}_1 = (\hat{H}_{\text{bath}} + \epsilon_d) \hat{P}_1 \\
\hat{H}_{22} &= \left(\sum_{k,\sigma} \epsilon_{\vec{k}} \hat{c}^\dagger_{\vec{k},\sigma} \hat{c}_{\vec{k},\sigma} + 2\epsilon_d + U \right) \hat{P}_2 = (\hat{H}_{\text{bath}} + 2\epsilon_d + U) \hat{P}_2 \\
\hat{H}_{01} &= \sum_{\vec{k}} V_{\vec{k}} \hat{c}^\dagger_{\vec{k},\sigma} \hat{d}_\sigma \hat{P}_1 \\
\hat{H}_{12} &= \sum_{\vec{k}} V_{\vec{k}} \hat{c}^\dagger_{\vec{k},\sigma} \hat{d}_\sigma \hat{P}_2 \\
\hat{H}_{02} &= \hat{0}
\end{aligned}$$

The annihilation operator $\hat{d}_\sigma$ in $\hat{H}_{01}$ is needed to specify the spin of the electron to be deleted and because it may induce a sign change.

All energy contributions except the hybridization term conserve the number of electrons on the impurity. Therefore these terms are diagonal in a representation with fixed number of electrons on the impurity. The hybridization term is then the coupling between states with different numbers of impurity electrons.

Formally Eliminate the zero and two-particle contribution from the Schrödinger equation

Next we determine the wave function, which may contain zero, one and two electrons on the impurity as

$$|\Psi\rangle = |\Psi_0\rangle + |\Psi_1\rangle + |\Psi_2\rangle$$

$$\begin{aligned}
(\hat{H}_{00} - E)|\Psi_0\rangle + \hat{H}_{01}|\Psi_1\rangle &= 0 \\
\hat{H}_{10}|\Psi_0\rangle + (\hat{H}_{11} - E)|\Psi_1\rangle + \hat{H}_{12}|\Psi_2\rangle &= 0 \\
\hat{H}_{21}|\Psi_1\rangle + (\hat{H}_{22} - E)|\Psi_2\rangle &= 0
\end{aligned}$$

The first and the last equations relate the contribution with zero and two particles on the impurity to that with one particle on the impurity

$$
\begin{aligned}
|\Psi_0\rangle &= -(\hat{H}_{00}-E)^{-1}\hat{H}_{01}|\Psi_1\rangle \\
|\Psi_2\rangle &= -(\hat{H}_{22}-E)^{-1}\hat{H}_{21}|\Psi_1\rangle
\end{aligned}
$$

Insertion into the middle equation yields

$$
\Big(\hat{H}_{11}-E-\hat{H}_{10}(\hat{H}_{00}-E)^{-1}\hat{H}_{01}-\hat{H}_{12}(\hat{H}_{22}-E)^{-1}\hat{H}_{21}\Big)|\Psi_1\rangle = 0
$$

Note that $|\Psi_1\rangle$ is only part of the state and therefore it is not normalized.

Substitute the Hamiltonian

Let us evaluate

$$
\begin{aligned}
\Big(E-\hat{H}_{22}\Big)\underbrace{\Big(\hat{P}_0+\hat{P}_1+\hat{P}_2\Big)}_{\hat{1}} &= E\hat{P}_0+E\hat{P}_1+\Big(E-\hat{H}_{\text{bath}}-2\epsilon_d-U\Big)\hat{P}_2 \\
\Big(\hat{H}_{22}-E\Big)^{-1} &= \frac{1}{E}\hat{P}_0+\frac{1}{E}\hat{P}_1+\Big(E-\hat{H}_{\text{bath}}-2\epsilon_d-U\Big)^{-1}\hat{P}_2 \\
\hat{H}_{12}\Big(\hat{H}_{22}-E\Big)^{-1}\hat{H}_{21} &= \sum_{\vec{k},\sigma}V^*_{\vec{k}}\hat{P}_1\hat{c}^\dagger_{\vec{k},\sigma}\hat{d}_\sigma\hat{P}_2\Big(E-\hat{H}_{\text{bath}}-2\epsilon_d-U\Big)^{-1}\hat{P}_2\sum_{\vec{k}',\sigma'}V_{\vec{k}'}\hat{d}^\dagger_{\sigma'}\hat{c}_{\vec{k}',\sigma'}\hat{P}_1 \\
&= \sum_{\vec{k},\sigma}\sum_{\vec{k}',\sigma'}V^*_{\vec{k}}V_{\vec{k}'}\hat{P}_1\hat{c}^\dagger_{\vec{k},\sigma}\hat{d}_\sigma\hat{P}_2\Big(E-\hat{H}_{\text{bath}}-2\epsilon_d-U\Big)^{-1}\hat{P}_2\hat{d}^\dagger_{\sigma'}\hat{c}_{\vec{k}',\sigma'}\hat{P}_1 \\
&= \sum_{\vec{k},\sigma}\sum_{\vec{k}',\sigma'}V^*_{\vec{k}}V_{\vec{k}'}\hat{c}^\dagger_{\vec{k},\sigma}\Big(E-\hat{H}_{\text{bath}}-2\epsilon_d-U\Big)^{-1}\hat{c}_{\vec{k}',\sigma'}\hat{P}_1\hat{d}_\sigma\hat{P}_2\hat{P}_2\hat{d}^\dagger_{\sigma'}\hat{P}_1
\end{aligned}
$$

All bath operators $\hat{c}^\dagger_{\vec{k},\sigma}$, $\hat{c}_{\vec{k},\sigma}$ anticommutate with the impurity operators. The projection operators and the bath Hamiltonian contain even products of creation and annihilation operators and therefore commutate with the creation operators of bath and impurity respectively. When we shift the impurity operators to the right we experience two sign changes, which do not have a net effect.

$$
\begin{aligned}
\hat{P}_1\hat{d}_\sigma\hat{P}_2\hat{P}_2\hat{d}^\dagger_{\sigma'}\hat{P}_1 &= \Big(\sum_{\sigma''}\hat{d}^\dagger_{\sigma''}\hat{d}_{\sigma''}-2\hat{P}_2\Big)\hat{d}_\sigma\hat{P}_2\hat{d}^\dagger_{\sigma'}\Big(\sum_{\sigma'''}\hat{d}^\dagger_{\sigma''}\hat{d}_{\sigma'''}-2\hat{P}_2\Big) \\
&= \Big(\sum_{\sigma''}\hat{d}^\dagger_{\sigma''}\hat{d}_{\sigma''}\Big)\hat{d}_\sigma\Big(\hat{d}^\dagger_\uparrow\hat{d}_\uparrow\hat{d}^\dagger_\downarrow\hat{d}_\downarrow\Big)\hat{d}^\dagger_{\sigma'}\Big(\sum_{\sigma'''}\hat{d}^\dagger_{\sigma'''}\hat{d}_{\sigma'''}\Big) \\
&= \Big(\hat{d}^\dagger_{-\sigma}\hat{d}_{-\sigma}\Big)\hat{d}_\sigma\Big(\hat{d}^\dagger_\uparrow\hat{d}_\uparrow\hat{d}^\dagger_\downarrow\hat{d}_\downarrow\Big)\hat{d}^\dagger_{\sigma'}\Big(\hat{d}^\dagger_{-\sigma'}\hat{d}_{-\sigma'}\Big) \\
&= \hat{d}_\sigma\Big(\hat{d}^\dagger_{-\sigma}\hat{d}_{-\sigma}\Big)\Big(\hat{d}^\dagger_\uparrow\hat{d}_\uparrow\hat{d}^\dagger_\downarrow\hat{d}_\downarrow\Big)\Big(\hat{d}^\dagger_{-\sigma'}\hat{d}_{-\sigma'}\Big)\hat{d}^\dagger_{\sigma'} \\
&= \hat{d}_\sigma\Big(\hat{d}^\dagger_\uparrow\hat{d}_\uparrow\hat{d}^\dagger_\downarrow\hat{d}_\downarrow\Big)\hat{d}^\dagger_{\sigma'}
\end{aligned}
$$

$$
\begin{aligned}
\hat{d}_\uparrow\Big(\hat{d}^\dagger_\uparrow\hat{d}_\uparrow\hat{d}^\dagger_\downarrow\hat{d}_\downarrow\Big)\hat{d}^\dagger_\uparrow &= \Big(1-\hat{d}^\dagger_\uparrow\hat{d}_\uparrow\Big)\hat{d}_\uparrow\hat{d}^\dagger_\downarrow\hat{d}_\downarrow\hat{d}^\dagger_\uparrow = -\hat{d}^\dagger_\uparrow\hat{d}_\uparrow\hat{d}^\dagger_\downarrow\hat{d}_\downarrow \\
\hat{d}_\uparrow\Big(\hat{d}^\dagger_\uparrow\hat{d}_\uparrow\hat{d}^\dagger_\downarrow\hat{d}_\downarrow\Big)\hat{d}^\dagger_\downarrow &= \Big(1-\hat{d}^\dagger_\uparrow\hat{d}_\uparrow\Big)\hat{d}_\uparrow\hat{d}^\dagger_\downarrow\Big(1-\hat{d}^\dagger_\downarrow\hat{d}_\downarrow\Big) = \hat{d}_\uparrow\hat{d}^\dagger_\downarrow \\
\hat{d}_\downarrow\Big(\hat{d}^\dagger_\uparrow\hat{d}_\uparrow\hat{d}^\dagger_\downarrow\hat{d}_\downarrow\Big)\hat{d}^\dagger_\uparrow &= \hat{d}_\downarrow\Big(\hat{d}^\dagger_\downarrow\hat{d}_\downarrow\hat{d}^\dagger_\uparrow\hat{d}_\uparrow\Big)\hat{d}^\dagger_\uparrow = \hat{d}_\downarrow\hat{d}^\dagger_\uparrow
\end{aligned}
$$

Thus, we obtain

$$
\hat{d}_\sigma\Big(\hat{d}^\dagger_\uparrow\hat{d}_\uparrow\hat{d}^\dagger_\downarrow\hat{d}_\downarrow\Big)\hat{d}^\dagger_{\sigma'} = -\delta_{\sigma,\sigma'}\hat{P}_2 - \delta_{\sigma,-\sigma'}\hat{d}^\dagger_{-\sigma}\hat{d}_\sigma
$$

Because the term will be applied to a state with exactly one electron on the impurity, the first term can be ignored and we obtain

$$
\begin{aligned}
\hat{H}_{12}\Big(E-\hat{H}_{22}\Big)^{-1}\hat{H}_{21} = &-\sum_{\vec{k},\vec{k'},\sigma} V^*_{\vec{k}}V_{\vec{k'}}\hat{c}^\dagger_{\vec{k},\sigma}\Big(E-\hat{H}_{\text{bath}}-2\epsilon_d-U\Big)^{-1}\hat{c}_{\vec{k'},\sigma}\hat{d}^\dagger_\uparrow\hat{d}_\uparrow\hat{d}^\dagger_\downarrow\hat{d}_\downarrow \\
&-\sum_{\vec{k},\vec{k'},\sigma} V^*_{\vec{k}}V_{\vec{k'}}\hat{c}^\dagger_{\vec{k},\sigma}\Big(E-\hat{H}_{\text{bath}}-2\epsilon_d-U\Big)^{-1}\hat{c}_{\vec{k'},-\sigma}\hat{d}^\dagger_{-\sigma}\hat{d}_\sigma
\end{aligned}
$$

We observe here that the mixing with the states with two particles on the impurity introduces a term in the effective Hamiltonian, that flips the spin on the impurity and in the bath.

Appendix A

Notation

A.1 Fermions and Bosons

Many expressions for Fermions and for Bosons are similar but differ by a sign. We indicate this sign by the symbol ζ. It is the sign of the wave function under particle permutation, that is $\psi(\vec{x}_1, \vec{x}_2) = \zeta\psi(\vec{x}_2, \vec{x}_1)$. The commutator or anticommutator relation then read

$$\left[\hat{\psi}^\dagger(\vec{x}), \hat{\psi}(\vec{x'})\right]_{-\zeta} = \delta(\vec{x} - \vec{x'})$$

Similarly the Fermi or Bose distribution functions are

$$f_{T,\mu}(\epsilon) = \frac{1}{\mathrm{e}^{\beta(\epsilon-\mu)} - \zeta}$$

A.2 Vectors, matrices, operators, functions, etc.

The **Heaviside step function** is $\theta(x)$.

Dirac's Delta function is $\delta(x)$. For vectorial arguments it is the product of the one-dimensional delta functions for the cartesian coordinates.

Operators are indicated by a hat such as $\hat{A}$.

Vectors are indicated by a vector arrow such as $\vec{x}$. Matrices are made bold face such as $\boldsymbol{A}$. I have purposely combined two different notation, because this makes the difference optically more evident. I also used the psychological analogy of using a arrow for a one-dimensionally extended object, and "fat symbol" for a two-dimensionally extended object. On the blackboard I use two underscores for matrices.

Functionals are indicated by an argument in brackets such as $F[y]$, where $y(x)$ is the function in the argument. If function arguments are mixed with normal arguments they are combined as in $F([y], x)$.

Commutators and anticommutators are indicated by brackets. The commutator has no a minus sign as subscript or no subscript. An anticommutator always carries a plus sign as subscript.

A.3 Comparison with Fetter Walecka

The Green's function G^{FW} in the Book of Fetter Walecka is related to our definition by

$$G = \frac{1}{\hbar} G^{FW}$$

see Eq. FW-7.1.

Appendix B

Derivations of mathematical expressions

Laplace expansion theorem (not needed)

Let D be the determinant of an $N \times N$ matrix $\boldsymbol{A}$, and let $D^{i,j}_{N-1}$ be the determinant of the matrix obtained from $\boldsymbol{A}$ by deleting the i-th line and the j-th column. Then

$$D = \sum_{i=1}^{N} (-1)^{i+j} A_{i,j} D^{i,j}_{N-1}$$

The Laplace expansion theorem can be applied twice, leading to

$$D = \sum_{i=1}^{N} \sum_{k=1}^{N} (-1)^{i+j+k+l} A_{i,j} A_{k,l} D^{ik,jl}_{N-2}$$

where $D^{ik,jl}$ is the determinant of the matrix $\boldsymbol{A}$ with the i-th and k-th line, and the j-th and l-th column deleted.

B.1 Fourier transform of the step function

Here, we derive the relation

$$\int_{-\infty}^{\infty} dt\; \theta(t) \mathrm{e}^{\frac{i}{\hbar}(\epsilon+i\eta)t} = \frac{i\hbar}{\epsilon + i\eta} \tag{B.1}$$

$$\int_{-\infty}^{\infty} dt\; \theta(-t) \mathrm{e}^{\frac{i}{\hbar}(\epsilon-i\eta)t} = \frac{-i\hbar}{\epsilon - i\eta} \tag{B.2}$$

where $\eta > 0$.

Derivation:

$$\begin{aligned}
\int_{-\infty}^{\infty} dt\; \theta(t) \mathrm{e}^{\frac{i}{\hbar}(\epsilon+i\eta)t} &= \int_{0}^{\infty} dt\; \mathrm{e}^{\frac{i}{\hbar}(\epsilon+i\eta)t} = \frac{1}{\frac{i}{\hbar}(\epsilon+i\eta)} \mathrm{e}^{\frac{i}{\hbar}(\epsilon+i\eta)t}\Big|_{0}^{\infty} \\
&\overset{\eta>0}{=} \frac{-1}{\frac{i}{\hbar}(\epsilon+i\eta)} = \frac{i\hbar}{\epsilon+i\eta}
\end{aligned}$$

$$\begin{aligned}\int_{-\infty}^{0} dt\, \theta(-t) \mathrm{e}^{\frac{i}{\hbar}(\epsilon - i\eta)t} &= \int_{-\infty}^{0} dt\, \mathrm{e}^{\frac{i}{\hbar}(\epsilon - i\eta)t} = \frac{1}{\frac{i}{\hbar}(\epsilon - i\eta)} \mathrm{e}^{\frac{i}{\hbar}(\epsilon - i\eta)t}\Big|_{-\infty}^{0} \\ &\overset{\eta \geq 0}{=} \frac{1}{\frac{i}{\hbar}(\epsilon - i\eta)} = \frac{-i\hbar}{\epsilon - i\eta}\end{aligned}$$

The function is a simple pole at $\epsilon = -i\eta$ or $\epsilon = +i\eta$ respectively.

Appendix C

Adiabatic decoupling in the Born-Oppenheimer approximation

C.1 Justification for the Born-Oppenheimer approximation

In the following, I will present an argument, that I **believe** is close to the original proof of Born and Oppenheimer. We will investigate an artificial model that we can solve exactly and compare the results of the Born-Oppenheimer approximation with the exact result.

Imagine a Hamiltonian in which all interactions are harmonic instead of Coulombic. Secondly, consider only a one-dimensional electron and a one-dimensional nucleus. The electron coordinate is denoted by x and the nuclear coordinate by R. The corresponding momenta are p_{el} for the electron coordinate and P_{nuc} for the nuclear coordinate. Thus, the full Hamiltonian is

$$\begin{aligned}\hat{H} &= \frac{\hat{P}^2_{nuc}}{2M} + \frac{\hat{p}^2_{el}}{2m} + \frac{1}{2}a\hat{x}^2 + b\hat{x}\hat{R} + \frac{1}{2}c\hat{R}^2 \\ &= \frac{m}{M}\frac{\hat{P}^2_{nuc}}{2m} + \frac{\hat{p}^2_{el}}{2m} + \frac{1}{2}a\hat{x}^2 + b\hat{x}\hat{R} + \frac{1}{2}c\hat{R}^2 \end{aligned} \tag{C.1}$$

Exact solution of the model system

From the description of the multidimensional harmonic oscillator (see ΦSX:Quantum Physics and ΦSX:Klassische Mechanik), we know that we need to know only the two classical eigenfrequencies ω_+ and ω_- to specify the spectrum of the two-dimensional quantum mechanical harmonic oscillator.

$$E_{j,n} = \hbar\omega_+\left(n + \frac{1}{2}\right) + \hbar\omega_-\left(j + \frac{1}{2}\right) \tag{C.2}$$

In order to obtain the classical frequencies $\omega_\pm$, let us consider the classical harmonic oscillator, which has the Hamilton function

$$H(P_{nuc}, p_{el}, R, x) \overset{\text{Eq. C.1}}{=} \frac{P^2_{nuc}}{2M} + \frac{p^2_{el}}{2m} + \frac{1}{2}ax^2 + bxR + \frac{1}{2}cR^2$$

Hamilton's second equation yield

$$\begin{pmatrix} \dot{p}_{el} \\ \dot{P}_{nuc} \end{pmatrix} = \begin{pmatrix} m\ddot{x} \\ M\ddot{R} \end{pmatrix} = -\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} x \\ R \end{pmatrix} = \begin{pmatrix} F_{el} \\ F_{nuc} \end{pmatrix}$$

By diagonalizing the dynamical matrix $\boldsymbol{D}$

$$\boldsymbol{D} = \frac{1}{m}\begin{pmatrix} a & \sqrt{\frac{m}{M}}b \\ \sqrt{\frac{m}{M}}b & \frac{m}{M}c \end{pmatrix}$$

we obtain the eigenvalues ω^2 of the dynamical matrix

$$m\omega^2=\frac{a+\frac{m}{M}c}{2}\pm\sqrt{\left(\frac{a-\frac{m}{M}c}{2}\right)^2+\frac{m}{M}b^2}$$

Let us investigate first two terms in an expansion in orders of $\frac{m}{M}$.

$$\begin{aligned}
m\omega_+^2&=a+\frac{m}{M}\frac{b^2}{a}+O\left(\frac{m}{M}\right)^2\\
m\omega_-^2&=\frac{m}{M}\left(c-\frac{b^2}{a}\right)-\left(\frac{m}{M}\right)^2\left(\frac{c^2}{2a}+\frac{a}{2}\left(\frac{2b^2-ac}{2a^2}\right)^2\right)+O\left(\frac{m}{M}\right)^4
\end{aligned}$$

The higher frequency ω_+ can be attributed to electronic excitation energies $\Delta E=\hbar\omega_+$. Thus, $\hbar\omega_+$ is similar size as excitation energies, that is typically in the range of few eV. The lower frequency ω_- can be attributed to the atomic motion. The energy $\hbar\omega_-$ is comparable to vibrational excitation energies, that is in the range of about 10 meV.

Thus, the eigenvalue spectrum is

$$\begin{aligned}
E_{j,n}&\overset{\text{Eq. C.2}}{=}\hbar\sqrt{\frac{c-\frac{b^2}{a}-\frac{m}{M}\left[\frac{2c^2}{a}+\frac{a}{2}\left(\frac{2b^2-ac}{2a^2}\right)^2\right]}{M}}\left(j+\frac{1}{2}\right)\\
&+\quad\hbar\sqrt{\frac{a+\frac{m}{M}\frac{b^2}{a}}{m}}\left(n+\frac{1}{2}\right)+O\left(\left(\frac{m}{M}\right)^2\right)
\end{aligned}\tag{C.3}$$

Born-Oppenheimer solution of the model system

We can now apply the Born-Oppenheimer approximation in order to see how it deviates from the exact result:

$$\begin{aligned}
\hat{H}^{BO}&=\frac{\hat{p}_{el}^2}{2m}+\frac{1}{2}ax^2+bRx+\frac{1}{2}cR^2\\
&=\frac{\hat{p}_{el}^2}{2m}+\frac{1}{2}a(x+\frac{b}{a}R)^2+\frac{1}{2}\left(c-\frac{b^2}{a}\right)R^2
\end{aligned}$$

The force constant for the electrons is a, so that we obtain a frequency $\omega_+^{BO}=\sqrt{\frac{a}{m}}$. Knowing that the energy eigenvalues of the harmonic oscillator are $\hbar\omega_+^{BO}(n+\frac{1}{2})$ we can immediately determine the Born-Oppenheimer energy surfaces.

$$E_n^{BO}(R)=\frac{1}{2}\left(c-\frac{b^2}{a}\right)R^2+\hbar\sqrt{\frac{a}{m}}\left(n+\frac{1}{2}\right)$$

Now we can determine the eigenvalues of the nuclear Hamiltonian

$$H^{nuc}=\frac{P_{nuc}^2}{2M}+E_n^{BO}(R)$$

which has eigenvalues

$$E_{j,n}=\hbar\sqrt{\frac{c-\frac{b^2}{a}}{M}}\left(j+\frac{1}{2}\right)+\hbar\sqrt{\frac{a}{m}}\left(n+\frac{1}{2}\right)\tag{C.4}$$

We see that the spectrum obtained from the Born-Oppenheimer approximation Eq. C.4 provides the correct result, Eq. C.3, if $\frac{m}{M}$ vanishes.

The corrections to the Born-Oppenheimer approximations, that is the difference between Eq. C.3 and Eq. C.4, are of order $\frac{m}{M}$.

This ratio, namely $\frac{m}{M}$, is typically in of order 10^{-4} and definitely smaller than 10^{-3}, because the mass of a single nucleon (proton or neutron) is about 1823 times the mass of an electron. Therefore we can assume that the Born-Oppenheimer corrections can be ignored in most cases.

The underlying reason for the validity of the Born-Oppenheimer approximation is that oscillators do not affect each other strongly, if they are out of resonance. In a solid, the oscillators are the electronic excitations and the phonons, that is lattice vibrations. Electronic excitation energies are larger than the optical band gap of the material, which is typically in the range of 1 eV. Lattice vibrations have wavenumbers below than 800 cm^{-1} and thus their energies lie below 10 meV. Except for metals or molecules with a nearly degenerate ground state this separation of frequencies is real.

Illustration of the out-of-resonance condition

To illustrate this effect let us consider the classical model of two coupled harmonic oscillators shown in Fig. C.1. The large pendulum represents the nuclear vibrations, while the small pendulum represents the electronic degrees of freedom. For the electrons, the nuclei seem static. The dynamics of the nuclei is perturbed little, because the forces from the electrons are almost averaged out. Two effects must be considered though: Only the center of gravity of electrons and nuclei performs a smooth oscillation, that is unperturbed by the electrons. The nucleus has in addition a small high-frequency oscillation with small amplitude. The long oscillation period is not given by the mass of the nucleus alone, but by the total mass of electrons and nucleus.

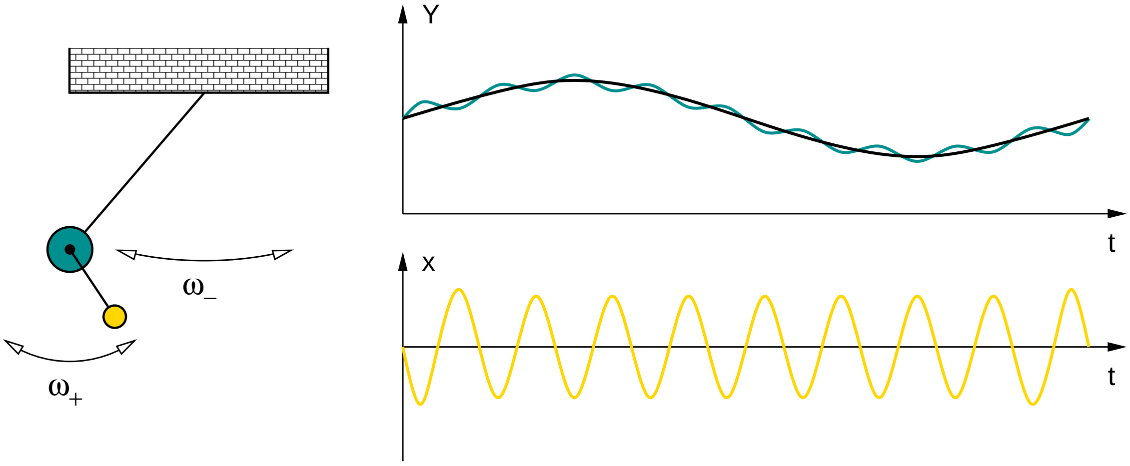


Fig. C.1: Schematic demonstration of the classical analogon for the Born-Oppenheimer approximation

Appendix D

Non-adiabatic effects

This section contains some derivations to which we refer in the section on non-adiabatic effects.

D.1 The off-diagonal terms of the first-derivative couplings $\vec{A}_{n,m,j}$

The non-Born-Oppenheimer terms in the Schrödinger equation for the nuclei Eq. 2.20 are related to the derivative couplings defined in Eq. 2.22 on p. 26

$$\vec{A}_{m,n} \overset{\text{Eq. 2.22}}{:=} \langle\Psi^{BO}_m|\frac{\hbar}{i}\vec{\nabla}_R|\Psi^{BO}_n\rangle = \langle\Psi^{BO}_m|\hat{P}|\Psi^{BO}_n\rangle \tag{D.1}$$

Below, we will relate the first-derivative couplings to the derivatives of the Born-Oppenheimer Hamiltonian and the Born-Oppenheimer surfaces.

FIRST-DERIVATIVE COUPLINGS

The first-derivative couplings can be expressed as

$$\vec{A}_{m,n,j}(\vec{R}) \overset{\text{Eq. D.3}}{=} \frac{\left\langle\Psi^{BO}_m(\vec{R})\middle|\big[\hat{P}_j,\hat{H}^{BO}(\vec{R})\big]_-\middle|\Psi^{BO}_n(\vec{R})\right\rangle}{\hat{E}^{BO}_n(\vec{R})-E^{BO}_m(\vec{R})} \qquad \text{for } E_n\neq E_m \tag{D.2}$$

where $\hat{P}_j=\frac{\hbar}{i}\vec{\nabla}_{R_j}$ is the momentum operator for the j-th nucleus.

Eq. D.2 makes it evident that non-adiabatic effects are dominant, when two Born-Oppenheimer surfaces $E^{BO}_n(\vec{R})$ become degenerate. Furthermore, these expressions, Eq. D.2, can be calculated without the need to evaluate the derivative of Born-Oppenheimer wave functions with respect to atomic positions. We only need the expectation values of the derivatives of the Born-Oppenheimer Hamiltonian, and that of the Born-Oppenheimer surfaces.

Derivation of Eq. D.2 The off-diagonal elements of $\vec{A}_{n,m,j}(\vec{r})$ of the first-derivative couplings are obtained as follows: We begin with the Schrödinger equation for the Born-Oppenheimer wave function

$$\Big(\hat{H}^{BO}(\vec{R})-E^{BO}_n(\vec{R})\Big)|\Psi^{BO}_n(\vec{R})\rangle \overset{\text{Eq. 2.4}}{=} 0$$

Because this equation is valid for all atomic positions, the gradient of the above expression vanishes.

$$\begin{aligned}
0 &= \vec{\nabla}_R\Big(\hat{H}^{BO}(\vec{R})-E^{BO}_n(\vec{R})\Big)\left|\Psi^{BO}_n(\vec{R})\right\rangle\\
&= \Big[\vec{\nabla}_R,\hat{H}^{BO}(\vec{R})\Big]_-\left|\Psi^{BO}_n(\vec{R})\right\rangle - \left|\Psi^{BO}_n(\vec{R})\right\rangle\vec{\nabla}_R E^{BO}_n(\vec{R}) + \Big(\hat{H}^{BO}(\vec{R})-E^{BO}_n(\vec{R})\Big)\vec{\nabla}_R\left|\Psi^{BO}_n(\vec{R})\right\rangle
\end{aligned}$$

Now, we multiply from the left with the bra $\langle\Psi_m^{BO}|$ and form the scalar products in the electronic Hilbert space.[1]

$$
\begin{aligned}
0 &= \langle\Psi_m^{BO}|\left[\vec{\nabla}_R, \hat{H}^{BO}(\vec{R})\right]_-|\Psi_n^{BO}\rangle - \langle\Psi_m^{BO}|\Psi_n^{BO}\rangle\vec{\nabla}_R E_n^{BO}(\vec{R}) + \langle\Psi_m^{BO}|\left(\hat{H}^{BO}(\vec{R}) - E_n^{BO}(\vec{R})\right)\vec{\nabla}_R|\Psi_n^{BO}\rangle \\
&= \langle\Psi_m^{BO}|\left[\vec{\nabla}_R, \hat{H}^{BO}(\vec{R})\right]_-|\Psi_n^{BO}\rangle - \delta_{m,n}\vec{\nabla}_R E_n^{BO}(\vec{R}) + \left(\hat{E}_m^{BO}(\vec{R}) - E_n^{BO}(\vec{R})\right)\langle\Psi_m^{BO}|\vec{\nabla}_R|\Psi_n^{BO}\rangle
\end{aligned}
$$

Thus, we obtain

$$
\begin{aligned}
\vec{A}_{m,n} \overset{\text{Eq. D.1}}{=} \langle\Psi_m^{BO}|\frac{\hbar}{i}\vec{\nabla}_R|\Psi_n^{BO}\rangle &= \frac{\langle\Psi_m^{BO}|\left[\frac{\hbar}{i}\vec{\nabla}_R, \hat{H}^{BO}(\vec{R})\right]_-|\Psi_n^{BO}\rangle}{\hat{E}_n^{BO}(\vec{R}) - E_m^{BO}(\vec{R})} \\
&= \frac{\langle\Psi_m^{BO}|\left[\hat{P}, \hat{H}^{BO}(\vec{R})\right]_-|\Psi_n^{BO}\rangle}{\hat{E}_n^{BO}(\vec{R}) - E_m^{BO}(\vec{R})} \qquad \text{for } E_n \neq E_m \qquad \text{(D.3)}
\end{aligned}
$$

This is the proof for Eq. D.2 given above. Unfortunately, no information can be extracted for the diagonal elements of the first-derivative couplings.

D.2 Diagonal terms of the derivative couplings

As we have seen, the non-adiabatic effects are determined entirely by the derivative couplings $\vec{A}_{m,n,j}(\vec{R})$ defined in Eq. 2.22 on p. 26, where m and n refer to the particular sheet of the Born-Oppenheimer surfaces, j refers to a particular atom and $\vec{R}$ is the 3N-dimensional vector describing the nuclear coordinates.

The most simple approximation including some of the non-adiabatic effects is to include only the diagonal terms of the first derivative couplings. In this approximation, the nuclear motion on the different Born-Oppenheimer surfaces are decoupled as in the Born-Oppenheimer approximation: This can be seen by inspecting Eq. 2.22: The derivative couplings are the only non-diagonal terms.

Within this diagonal approximation, we can do again the classical approximation and see that the derivative couplings act like a magnetic field.

$$
H_n(\vec{P}, \vec{R}) = \sum_{j=1}^{3M} \frac{1}{2M_j}\left(\vec{P}_j + \vec{A}_{n,n,j}(\vec{R})\right)^2 + E_n^{BO}(\vec{R}) \qquad \text{(D.4)}
$$

For convenience, we use $3M$ dimensional vectors for the momenta and coordinates of the nuclei. Similarly, we attribute a matrix M_j to each of the $3M$ nuclear coordinates. The index n refers to a particular sheet of the Born-Oppenheimer surfaces.

Hamilton's equations for this system have the form

$$
\dot{\vec{P}}_j = -\vec{\nabla}_{\vec{R}_j} H(\vec{R}, \vec{P}) = -\sum_k \left(\vec{\nabla}_j \otimes \vec{A}_{n,n,k}\right) \underbrace{\frac{1}{M_k}\left(\vec{P}_k - \vec{A}_{n,n,k}\right)}_{\dot{\vec{R}}_k} - \vec{\nabla} E_n^{BO} \qquad \text{(D.5)}
$$

$$
\dot{\vec{R}}_j = \vec{\nabla}_{\vec{P}_j} H(\vec{R}, \vec{P}) = \frac{1}{M_j}\left(\vec{P}_j + \vec{A}_{n,n,j}\right) \qquad \text{(D.6)}
$$

With $\otimes$ I denote the dyadic or outer product defined by $(\vec{a}\otimes\vec{b})_{i,j} = a_i b_j$. Note here that the gradient Eq. D.5 does not act beyond the closed parenthesis, that is, it does not act on the vector potential standing on the right.

---

[1] The Born-Oppenheimer wave functions depend on electronic and nuclear coordinates. The nuclear coordinates, however, play a different role, because they are treated like external parameters.

The two Hamilton's equations Eq. D.5 and Eq. D.6 can be combined into the form of Newton's equations

$$
\begin{aligned}
M_j\ddot{\vec{R}}_j &\overset{\text{Eq. D.6}}{=} \dot{\vec{P}}_j+\sum_k\Big(\dot{\vec{R}}_k\vec{\nabla}_{\vec{R}_k}\Big)\vec{A}_{n,n,j}(\vec{R})\\
&\overset{\text{Eq. D.5}}{=} \overbrace{-\sum_k\Big(\vec{\nabla}_j\otimes\vec{A}_{n,n,k}\Big)\underbrace{\frac{1}{M_k}\Big(\vec{P}_k+\vec{A}_{n,n,k}\Big)}_{\dot{\vec{R}}_k}-\vec{\nabla}_jE_n^{BO}}^{\dot{\vec{P}}_j}+\sum_k\Big(\dot{\vec{R}}_k\vec{\nabla}_k\Big)\vec{A}_{n,n,j}(\vec{R})\\
&\overset{\text{Eq. D.6}}{=} \sum_k\Big(\dot{\vec{R}}_k\vec{\nabla}_k\Big)\vec{A}_{n,n,j}(\vec{R})-\sum_k\Big(\vec{\nabla}_j\otimes\vec{A}_{n,n,k}\Big)\dot{\vec{R}}_k-\vec{\nabla}_jE_n^{BO}\\
&\overset{\text{Eq. D.6}}{=} \sum_k\dot{\vec{R}}_k\boldsymbol{F}_{k,j}-\vec{\nabla}_jE_n^{BO}
\end{aligned}
\tag{D.7}
$$

where $\boldsymbol{F}_{k,j}$ is something like a generalized field-strength tensor, which captures the non-adiabatic effects

$$
\boldsymbol{F}_{k,j}=\Big(\vec{\nabla}_k\otimes\vec{A}_{n,n,j}\Big)-\Big(\vec{\nabla}_j\otimes\vec{A}_{n,n,k}\Big)^{\dagger}
\tag{D.8}
$$

Like the Hamilton function, also the equation of motion has great similarities with equation of motion of a charged particle (with charge $q=-1$) in an electromagnetic field. There, is however a major difference, namely the velocity of one particle contributes to the force acting on another. There would be generalized magnetic field for each pair of atoms. The field strength tensor is not a three-by-three matrix, but a $3M\times3M$-matrix, where M is the number of nuclei.

Reminder of electrodynamics The equation of motion of a charged particle in a magnetic field is

$$M_j\ddot{\vec{R}}_j = q\Big(\vec{E}(\vec{R}) + \dot{\vec{R}} \times \vec{B}(\vec{R})\Big) = q\Big(\vec{E} + \dot{\vec{R}}\boldsymbol{F}\Big) \tag{D.9}$$

The magnetic field $\vec{B}$ and the electric field $\vec{E}$ are connected to the vector potential $\vec{A}$ and the electric potential ϕ by

$$\begin{aligned}\vec{B} &= \vec{\nabla} \times \vec{A}\\ \vec{E} &= -\vec{\nabla}\Phi - \dot{\vec{A}}\end{aligned} \tag{D.10}$$

The electric field-strength tensor $\boldsymbol{F}$ is

$$\boldsymbol{F} = \Big(\vec{\nabla} \otimes \vec{A}\Big) - \Big(\vec{\nabla} \otimes \vec{A}\Big)^{\dagger} = \begin{pmatrix} 0 & B_z & -B_y \\ B_y & 0 & B_x \\ B_y & -B_y & 0 \end{pmatrix} \tag{D.11}$$

The components the field-strength tensor are

$$F_{i,j} = \partial_i A_j - \partial_j A_i \tag{D.12}$$

and those of the magnetic field are

$$B_i = \sum_{j,k} \epsilon_{i,j,k}\partial_j A_k \tag{D.13}$$

where $\epsilon_{i,j,k}$ is the Levi-Civita symbol or fully antisymmetric tensor. The Lorentz force can be written in terms of the Field strength tensor using

$$\dot{\vec{R}}\boldsymbol{F} = \dot{\vec{R}} \times \vec{B} \tag{D.14}$$

We can work out the analogue of the field-strength tensor using the notation with $3M$ dimensional vectors that combine cartesian coordinate with the particle index.

$$\begin{aligned}\partial_i A_{n,n,j} - \partial_j A_{n,n,i} &= \partial_{R_i}\langle\Psi_n^{BO}|\frac{\hbar}{i}\partial_{R_j}|\Psi_n^{BO}\rangle - \partial_{R_j}\langle\Psi_n^{BO}|\frac{\hbar}{i}\partial_{R_i}|\Psi_n^{BO}\rangle\\ &= \frac{\hbar}{i}\Big(\langle\partial_{R_i}\Psi_n^{BO}|\partial_{R_j}\Psi_n^{BO}\rangle - \langle\partial_{R_j}\Psi_n^{BO}|\partial_{R_i}\Psi_n^{BO}\rangle\Big)\\ &= 2\hbar\,\mathrm{Im}\Big(\langle\partial_{R_i}\Psi_n^{BO}|\partial_{R_j}\Psi_n^{BO}\rangle\Big)\end{aligned}$$

Using again the vector-matrix notation for cartesian coordinates and indices for the atoms, we obtain

$$\boldsymbol{F}_{k,j} = 2\hbar\,\mathrm{Im}\,\Big\langle\vec{\nabla}_k\Psi_n^{BO}\Big|\otimes\Big|\vec{\nabla}_j\Psi_n^{BO}\Big\rangle \tag{D.15}$$

This result seems to indicate that the non-adiabatic effects can be made to vanish, if the Born-Oppenheimer wave functions are chosen to be real. This is, however, not so: While real wave functions can be chosen in the absence of magnetic fields as a consequence of time-inversion symmetry[2], there are cases where such a choice unavoidable produces a step in the wave function. This is related to the existence of a **geometrical phase**. Key words related to the geometrical phase is the **Jahn-Teller effect** and the **Berry phase**.

D.3 The diabatic picture

`Editor: This section is not yet completed. It may be used alongside Baer's paper.`

---

[2] See appendix L

When the Born-Oppenheimer surfaces approach each other, the derivative couplings diverge. Therefore not only the Born-Oppenheimer approximation fails in those regions of configuration space, the formulation becomes including the non-adiabatic corrections becomes untractable. In order to avoid this problem, the diabatic representation of the nuclear wave functions has been developed, which avoids these problems. Unfortunately, the construction of the diabatic states is non-trivial and it is a non-local operation. The diabatic representation has been described well in the paper by Baer [84], which we follow here:

We start from the nuclear Hamiltonian in the adiabatic picture, which we have derived earlier

$$\begin{aligned} i\hbar\partial_t\phi^{adia}_m(\vec{R},t) \overset{\text{Eq. 2.20}}{=} \sum_n\Big[\sum_{j=1}^{M}\frac{1}{2M_j}\sum_k\Big(\delta_{m,k}\frac{\hbar}{i}\vec{\nabla}_{R_j}+\vec{A}_{m,k,j}(\vec{R})\Big)\Big(\delta_{k,n}\frac{\hbar}{i}\vec{\nabla}_{R_j}+\vec{A}_{k,n,j}(\vec{R})\Big)& \\ +\delta_{m,n}E^{BO}_m(\vec{R})\Big]\phi^{adia}_n(\vec{R},t)& \qquad \text{(D.16)} \end{aligned}$$

Now we introduce a position-dependent, unitary transformation of the nuclear wave function $\boldsymbol{S}(\vec{R})$ that depends on the atomic positions.

$$\phi^{adia}_n(\vec{R},t)=\sum_m \phi^{dia}_m(\vec{R},t)S_{m,n}(\vec{R})$$

We want to insert it into Eq. D.16. Therefore, we first investigate the action of the kinetic energy term on the transformation. We obtain

$$\begin{aligned} \sum_m\Big(\delta_{n,m}\hat{\vec{P}}_j+\vec{A}_{n,m,j}\Big)S_{m,o}(\vec{R}) &= \frac{\hbar}{i}\vec{\nabla}_j S_{n,o}(\vec{R})+\sum_m S_{n,m}(\vec{R})\delta_{m,o}\hat{\vec{P}}_j+\sum_m\vec{A}_{n,m,j}S_{m,o}(\vec{R}) \\ &= \sum_m S_{n,m}(\vec{R})\delta_{m,o}\hat{\vec{P}}_j\ , \qquad \text{(D.17)} \end{aligned}$$

if we can find a transformation $S_{n,m}$ so that

$$\frac{\hbar}{i}\vec{\nabla}_j S_{n,o}(\vec{R})+\sum_m\vec{A}_{n,m,j}S_{m,o}(\vec{R})=0 \qquad \text{(D.18)}$$

Now we insert this transformation into Eq. D.16

$$\begin{aligned} i\hbar\partial_t\sum_n S_{m,n}\phi^{dia}_n(\vec{R},t) &= \sum_{n,o}\Big[\sum_{j=1}^{M}\frac{1}{2M_j}\sum_k\Big(\delta_{m,k}\frac{\hbar}{i}\vec{\nabla}_{R_j}+\vec{A}_{m,k,j}(\vec{R})\Big)\Big(\delta_{k,n}\frac{\hbar}{i}\vec{\nabla}_{R_j}+\vec{A}_{k,n,j}(\vec{R})\Big) \\ &\qquad +\delta_{m,n}E^{BO}_m(\vec{R})\Big]S_{n,o}(\vec{R})\phi^{dia}_o(\vec{R},t) \\ &= \sum_{n,o}S_{m,n}(\vec{R})\Big[\sum_{j=1}^{M}\frac{1}{2M_j}\sum_k\Big(\delta_{n,k}\frac{\hbar}{i}\vec{\nabla}_{R_j}\Big)\Big(\delta_{k,o}\frac{\hbar}{i}\vec{\nabla}_{R_j}\Big) \\ &\qquad +\sum_k S^{\dagger}_{n,k}(\vec{R})E^{BO}_k(\vec{R})S_{k,o}(\vec{R})\Big]\phi^{dia}_o(\vec{R},t) \\ i\hbar\partial_t\phi^{dia}_n(\vec{R},t) &= \sum_o\Big[\sum_{j=1}^{M}\frac{1}{2M_j}\delta_{n,o}\hat{\vec{P}}^2_j+\sum_k S^{\dagger}_{n,k}(\vec{R})E^{BO}_k(\vec{R})S_{k,o}(\vec{R})\Big]\phi^{dia}_o(\vec{R},t) \qquad \text{(D.19)} \end{aligned}$$

That is, in the diabatic picture the nuclear Schrödinger equation Eq. 2.20 has the simple form

$$i\hbar\partial_t\phi^{dia}_n(\vec{R},t)=\Big[\sum_{j=1}^{M}\frac{-\hbar^2}{2M_j}\vec{\nabla}^2_{R_j}+E^{BO}_n(\vec{R})\Big]\phi_n(\vec{R},t)+\sum_m W_{n,m}(\vec{R})\phi_m(\vec{R},t) \qquad \text{(D.20)}$$

$$\text{where}\qquad W_{n,m}(\vec{R})=\sum_k S^{\dagger}_{n,k}(\vec{R})E^{BO}_k(\vec{R})S_{k,o}(\vec{R})-E^{BO}_n(\vec{R})\delta_{n,o} \qquad \text{(D.21)}$$

Instead of a vector potential, the non-adiabatic effects are capture in a shifts of the total energy surfaces and coupling terms between them.

The adiabatic-to-diabatic transformation can be obtained

$$S_{m,n}(\vec{R}) = \sum_{o} \left[\exp\left(-\frac{i}{\hbar} \sum_{j=1}^{M} \int_{\vec{R}_0}^{\vec{R}} d\vec{R}'_j \; \vec{\boldsymbol{A}}_j(\vec{R}) \right) \right]_{m,o} S_{o,n}(\vec{R}_0) \tag{D.22}$$

if the integral is independent of the path taken. The integral is performed along a path conecting $\vec{R}_0$ and $\vec{R}$.

The requirement that the integral in Eq. D.22 is independent of the path is violated at conical intersections. Thus, in the presence of a conical intersection the transformation to a diabatic representation can only be approximated.

From here on, I recommend to follow the paper by Baer [84], and for references to newer developments the review by Worth and Robb[9].

Appendix E

Non-crossing rule

Here we provide the proof of the non-crossing rule discussed in section 2.3.2 on p. 2.3.2.

NON-CROSSING RULE

Consider a Hamiltonian $\hat{H}(Q_1, \dots, Q_m)$ which depends on m parameters $Q_1, \dots, Q_m$. For each accidental crossing of two eigenvalues, there is a three-dimensional subspace of the parameter space in which the degeneracy is lifted except for a single point.
If the hamiltonian is real for all parameters, the subspace where the degeneracy is lifted has two dimensions.

The **non-crossing rule** has been derived by von Neumann and Wigner[15].

E.1 Proof of the non-crossing rule

Consider a Hamiltonian $\boldsymbol{H}(\vec{Q})$, which depends on a set of coordinates $\vec{Q}$. Let us explore the changes of the Hamiltonian in the neighborhood of $\vec{Q}_0$.

$$\boldsymbol{H}(\vec{Q}) = \boldsymbol{H}(\vec{Q}_0) + \left(\vec{Q} - \vec{Q}_0\right) \vec{\nabla}_{\vec{Q}}\big|_{\vec{Q}_0} \boldsymbol{H} + O\left(|\vec{Q} - \vec{Q}_0|^2\right) \tag{E.1}$$

1. A gradient of a scalar function of $\vec{Q}$ defines a single direction in the coordinate space. Any displacement perpendicular to this direction will leave the scalar function unchanged to first order.
2. For a double valued function, the gradients of the first and the second component span a two-dimensional space.
3. The number of free parameters of an object determines the maximum dimension of the space in which it can be varied without affecting its defining properties.

This argument is generalized to a Hamiltonian, once with a doubly degenerate state and once with the degeneracy split into two non-degenerate states. Below, we will see that the space spanned without the degeneracy has a higher dimension than that with the degeneracy. The number of additional dimensions is 2 for real hamiltonians and 3 for complex hamiltonians. Thus, there are two, respectively three directions, along which the degeneracy is lifted. Displacements orthogonal to these two or three distortions leave the degeneracy intact.

Thus, one can identify a two-dimensional, respectively three-dimensional space of displacements that are relevant for the intersection. Within this subspace, the intersection is zero-dimensional, that is, a point. This is why one calls this intersection a **conical intersection**. Orthogonal to this two, respectively three-dimensional subspace, are the “irrelevant” directions, which may affect the position or value of the conical intersection, but they do not lift the degeneracy.

Free parameters of a hamiltonian with a specified multiplett structure: What is left, is to determine the number of independent parameters that change a hamiltonian with a given multiplet structure without, however, lifting any degeneracies.

Every Hamiltonian can be constructed from its eigenvalues E_k and eigenvectors. The eigenvectors form the unitary matrix $\boldsymbol{U}$

$$H_{i,j}=\sum_k U^\dagger_{i,k}E_kU_{k,j} \tag{E.2}$$

Below, we will determine the number $\kappa(n)$ of free parameters for a unitary matrix of dimension n. Note that $\kappa(n)$ differs for real and complex unitary matrices.

Let us assume that the eigenvalues form ℓ multiplets with degeneracies g_j for $j=1,\ldots,\ell$.

If we sum the number ℓ of free parameters for the energies E_k and the number $\kappa(n)$ of parameters of a unitary transformation $\boldsymbol{U}$, we overcount the number of free parameters, because a number of unitary matrices $\boldsymbol{U}$ do not affect the Hamiltonian at all.

The overcounting can be quantified as the number of independent unitary transformations that leave the Hamiltonian invariant. These are the unitary transformations $\boldsymbol{V}$ with

$$\sum_k V^\dagger_{m,i}H_{i,j}V_{j,n}=H_{i,j}\,. \tag{E.3}$$

Thus, thus the desired unitary matrices $\boldsymbol{V}$ commutate with the Hamiltonian, i.e.

$$[\boldsymbol{H},\boldsymbol{V}]_-=0\,. \tag{E.4}$$

This requirement can be transformed into

$$0=[\boldsymbol{H},\boldsymbol{V}]_-=\left[\boldsymbol{U}\boldsymbol{H}\boldsymbol{U}^\dagger,\boldsymbol{U}\boldsymbol{V}\boldsymbol{U}^\dagger\right]_-=\left[\boldsymbol{E},\boldsymbol{V}'\right]_- \tag{E.5}$$

where $(\boldsymbol{E})_{i,j}\stackrel{\text{def}}{=}E_i\delta_{i,j}$ and $\boldsymbol{V}'\stackrel{\text{def}}{=}\boldsymbol{U}\boldsymbol{V}\boldsymbol{U}^\dagger$.

Any unitary transformation that transforms only states within the same multiplet leaves the matrix $\boldsymbol{E}$ invariant.[1] Any unitary transformation between different multipletts mixes energies and thus changes the matrix $\boldsymbol{E}$. Hence, the overcounting corresponds to the free parameters of all unitary transformations that act only within the multiplets. Thus we need to subtract $\sum_{j=1}^{\ell}\kappa(g_j)$.

Thus, the number of free parameters of a Hamiltonian with ℓ multiplets with degeneracies g_j is

$$D(g_1,\ldots,g_\ell)=\kappa\left(\underbrace{\sum_{j=1}^{\ell}g_j}_{n}\right)+\ell-\sum_{j=1}^{\ell}\kappa(g_j) \tag{E.6}$$

Degrees of freedom for a real unitary matrix Let me now work out the number $\kappa(n)$ of a real unitary matrix with dimension n.

1. a general real matrix of dimension n has n^2 matrix elements and thus n^2 degrees of freedom.
2. the requirement of orthonormal vectors accounts for $\frac{n(n-1)}{2}$ real-valued conditions $\sum_k U^*_{k,i}U_{k,j}=0$ with $j>k$. This corresponds to $\frac{1}{2}n(n-1)$ degrees of freedom. Thus the number of degrees of freedom is reduced to $n^2-\frac{1}{2}n(n-1)$.
3. the normalization of the vectors, i.e. $\sum_k U^*_{k,i}U_{k,i}=1$ with $i=1,\ldots,n$, contributes n real-valued conditions. This reduces the number of degrees of freedom further to $n^2-\frac{1}{2}n(n-1)-n=\frac{1}{2}n(n-1)$.

---

[1] Consider one multiplett. The energies for all states in the multiplett are the same, that is $\boldsymbol{E}=E\boldsymbol{1}$, where E is the energy of the multiplett. $\boldsymbol{U}^\dagger\boldsymbol{E}\boldsymbol{U}=E\boldsymbol{U}^\dagger\boldsymbol{U}=E\boldsymbol{1}=\boldsymbol{E}$.

Thus, the number of degrees of freedom to fix a general unitary matrix is

$$\kappa^{\text{real}}(n) = \frac{n(n-1)}{2} \tag{E.7}$$

Degrees of freedom for a complex unitary matrix Let me now work out the number $\kappa^{\text{complex}}(n)$ of a compex unitary matrix with dimension n.

1. a general complex matrix of dimension n has n^2 complex matrix elements and thus $2n^2$ real degrees of freedom.
2. the requirement of orthonormal vectors accounts for $\frac{n(n-1)}{2}$ complex-valued conditions $\sum_k U^*_{k,i}U_{k,j} = 0$ with $j > k$. This corresponds to $n(n-1)$ degrees of freedom. Thus the number of degrees of freedom is reduced to $2n^2 - n(n-1)$.
3. the normalization of the vectors, i.e. $\sum_k U^*_{k,i}U_{k,i} = 1$ with $i = 1, \dots, n$, contributes n real-valued conditions. This reduces the number of degrees of freedom further to $2n^2 - n(n-1) - n = n^2$.

Thus, the number of degrees of freedom to fix a general unitary matrix is

$$\kappa^{\text{complex}}(n) = n^2 \tag{E.8}$$

Compare parameters with and without degeneracy: Now we need to understand how the number of free parameters changes when a double degenerate states splits into two nondegenerate levels. In the presence of the double degeneracy, there shall be p muliplets plus one doublet. When the degeneracy is lifted, the doublet splits into two singlets. The dimension of both hamiltonians is $n = 2 + \sum_{j=1}^{p} g_j$.

$$\begin{aligned}
D(g_1,\dots,g_p,1,1) - D(g_1,\dots,g_p,2) &\overset{\text{Eq. E.6}}{=} \underbrace{\Big[\kappa(n) + (p+2) - \sum_{j=1}^{p}\kappa(g_j) - 2\kappa(1)\Big]}_{\text{split}} \\
&- \underbrace{\Big[\kappa(n) + (p+1) - \sum_{j=1}^{p}\kappa(g_j) - \kappa(2)\Big]}_{\text{degenerate}} \\
&= 1 - 2\kappa(1) + \kappa(2) \\
&\overset{\text{Eqs. E.7, E.8}}{=} \begin{cases} 2 & \text{for a real hamiltonian} \\ 3 & \text{for a complex hamiltonian} \end{cases}
\end{aligned} \tag{E.9}$$

Appendix F

Landau-Zener Formula

not finished!!!

The transition between two adiabatic total-energy surfaces has been analyzed by Landau and Zener[1] [86**?** , 85]. for a one-dimensional model. Here, we follow the analysis of Wittig[85].

The model of Landau and Zener considers two crossing Born-Oppenheimer surfaces in only one dimension X.

$$\boldsymbol{H}^{BO}(X)=\begin{pmatrix}-F_1X & 0\\ 0 & -F_2X\end{pmatrix}$$

The Born-Oppenheimer surfaces are assumed to depend linearly on the atomic coordinate X.

The non-adiabatic effects are approximated by non-diagonal elements of the Hamiltonian, that are considered independent of the atomic position. Thus, the full Hamiltonian has the form

$$\boldsymbol{H}(X)=\begin{pmatrix}-F_1X & H_{12}\\ H^*_{12} & -F_2X\end{pmatrix}$$

The nuclear dynamics is considered classical, and with a constant velocity $\partial_t X(t)=V$, that is $X(t)=Vt$. This turns the position-dependent Hamiltonian into an effectively time-dependent Hamiltonian.

$$\boldsymbol{H}(t)=\begin{pmatrix}-F_1Vt & H_{12}\\ H^*_{12} & -F_2Vt\end{pmatrix}$$

The time dependent electronic Hamiltonian is described by the following Ansatz

$$\begin{pmatrix}\Phi_1\\ \Phi_2\end{pmatrix}=\begin{pmatrix}A(t)\mathrm{e}^{-\frac{i}{\hbar}\int dt\;E_1^{BO}(X(t))}\\ B(t)\mathrm{e}^{-\frac{i}{\hbar}\int dt\;E_2^{BO}(X(t))}\end{pmatrix}=\begin{pmatrix}A(t)\mathrm{e}^{+\frac{i}{\hbar}\int dt\;F_1Vt}\\ B(t)\mathrm{e}^{+\frac{i}{\hbar}\int dt\;F_2Vt}\end{pmatrix}=\begin{pmatrix}A(t)\mathrm{e}^{+\frac{i}{\hbar}\frac{F_1V}{2}t^2}\\ B(t)\mathrm{e}^{+\frac{i}{\hbar}\frac{F_2V}{2}t^2}\end{pmatrix}$$

The lower bound of the time integral has been ignored because they are a multiplicative factor, that can be absorbed in the initial conditions.

Insertion into the time dependent Schrödinger equation yield

$$i\hbar\partial_t\begin{pmatrix}A(t)\mathrm{e}^{iF_1Vt^2/(2\hbar)}\\ B(t)\mathrm{e}^{iF_2Vt^2/(2\hbar)}\end{pmatrix}=\begin{pmatrix}-F_1Vt & H_{12}\\ H^*_{12} & -F_2Vt\end{pmatrix}\begin{pmatrix}A(t)\mathrm{e}^{iF_1Vt^2/(2\hbar)}\\ B(t)\mathrm{e}^{iF_2Vt^2/(2\hbar)}\end{pmatrix}$$

$$\begin{pmatrix}\partial_t A(t)\\ \partial_t B(t)\end{pmatrix}=\begin{pmatrix}-\frac{i}{\hbar}H_{12}B(t)\mathrm{e}^{i(F_2-F_1)Vt^2/(2\hbar)}\\ -\frac{i}{\hbar}H^*_{12}A(t)\mathrm{e}^{-i(F_2-F_1)Vt^2/(2\hbar)}\end{pmatrix}$$

---

[1] According to Wittig[85] the result of Landau[**?**] has an error of 2π

thus

$$
\begin{aligned}
\partial_t A(t) &= -\frac{i}{\hbar} H_{12} B(t) \mathrm{e}^{i(F_2-F_1)Vt^2/(2\hbar)} \\
\partial_t B(t) &= -\frac{i}{\hbar} H_{12}^* A(t) \mathrm{e}^{-i(F_2-F_1)Vt^2/(2\hbar)} \qquad \Rightarrow \qquad A(t) = -\frac{\hbar}{i} \underbrace{\frac{H_{12}}{|H_{12}|^2}}_{1/H_{12}^*} \mathrm{e}^{i(F_2-F_1)Vt^2/(2\hbar)} \partial_t B(t)
\end{aligned}
$$

Insertion of the second equation into the first yields a differential equation for $B(t)$.

$$
\begin{aligned}
\partial_t \underbrace{\left(-\frac{\hbar}{i} \frac{H_{12}}{|H_{12}|^2} \mathrm{e}^{i(F_2-F_1)Vt^2/(2\hbar)} \partial_t B(t) \right)}_{A(t)} &= -\frac{i}{\hbar} H_{12} B(t) \mathrm{e}^{i(F_2-F_1)Vt^2/(2\hbar)} \\
-\frac{H_{12}}{|H_{12}|^2} (F_2-F_1) V t \mathrm{e}^{i(F_2-F_1)Vt^2/(2\hbar)} \partial_t B(t) + \left(-\frac{\hbar}{i} \frac{H_{12}}{|H_{12}|^2} \mathrm{e}^{i(F_2-F_1)Vt^2/(2\hbar)} \partial_t^2 B(t) \right) &= -\frac{i}{\hbar} H_{12} B(t) \mathrm{e}^{i\frac{(F_2-F_1)V}{2\hbar}t^2} \\
-\frac{1}{|H_{12}|^2} (F_2-F_1) V t \partial_t B(t) + \left(-\frac{\hbar}{i} \frac{1}{|H_{12}|^2} \partial_t^2 B(t) \right) &= -\frac{i}{\hbar} B(t) \\
\frac{i}{\hbar} (F_2-F_1) V t \, \partial_t B(t) + \partial_t^2 B(t) &= -\frac{1}{\hbar^2} |H_{12}|^2 B(t) \\
\partial_t^2 B(t) + \frac{i}{\hbar} (F_2-F_1) V t \, \partial_t B(t) + \frac{1}{\hbar^2} |H_{12}|^2 B(t) &= 0
\end{aligned}
$$

Thus, we have a differential equation for the probability amplitude on the second sheet. We now look for a solution of this differential equation, respectively its value $B(t=+\infty)$ with the initial condition $B(-\infty) = \partial_t B(-\infty) = 0$.

This is a differential equation of the form

$$
\ddot{x} + i\alpha t \, \dot{x} + \beta x = 0
$$

So far, the derivation is standard. Wittig showed a solution to the problem using contour integrals.

$$
\begin{aligned}
\frac{1}{t} \frac{\ddot{x}}{x} + i\alpha \frac{\dot{x}}{x} + \frac{\beta}{t} &= 0 \\
i\alpha \int_{t_i}^{t_f} dt \, \partial_t \ln[x(t)] &= -\beta \int_{t_i}^{t_f} dt \, \frac{1}{t} - \int_{t_i}^{t_f} dt \, \frac{1}{t} \frac{\ddot{x}}{x} \\
i\alpha \ln[\frac{x_f}{x_i}] &= -\beta \int_{-\infty}^{\infty} dt \, \frac{1}{t} - \int_{-\infty}^{\infty} dt \, \frac{1}{t} \frac{\ddot{x}}{x}
\end{aligned}
$$

The integral of the first term on the right hand side yields

$$
\int_{-\infty}^{\infty} dt \, \frac{1}{t} = \pm i\pi
$$

depending on whether the integration is performed in the upper or lower half plane.

Thus, we obtain

$$
\begin{aligned}
\ln[x_f] &= \ln[x_i] \mp \pi \frac{\beta}{\alpha} + \frac{i}{\alpha} \int_{-\infty}^{\infty} dt \, \frac{1}{t} \frac{\ddot{x}}{x} \\
x_f &= x_i \mathrm{e}^{\mp\pi\beta/\alpha} \exp\left(\frac{i}{\alpha} \int_{-\infty}^{\infty} dt \, \frac{1}{t} \frac{\ddot{x}}{x} \right)
\end{aligned}
$$

If the transition probability is small, that is $x_f - x_i << 1$, we can ignore the last term and obtain the estimate

$$
x_f \approx x_i \mathrm{e}^{\mp\pi\beta/\alpha}
$$

When we insert the parameters $\alpha = (F_2 - F_1)V/\hbar$ and $\beta = |H_{12}|^2/\hbar^2$ we obtain the Landau-Zener estimate.

LANDAU ZENER ESTIMATE FOR THE TRANSITION PROBABILITY AT A BAND CROSSING

The transition probability is (Eq. 16 of Wittig)

$$P = \exp\left(2\pi \underbrace{\frac{|H_{12}|}{\hbar}}_{\text{Rabi frequency}} \underbrace{\frac{|H_{12}|}{V|F_1 - F_2|}}_{\tau}\right)$$

τ is of order of the time it takes to cross the region where the Born-Oppenheimer surfaces are closer than the avoided crossing surface.

================================

The result of their analysis is the **Landau-Zener formula**, which gives the probability for a transition for a one-dimensional model

$$P = \mathrm{e}^{-2\pi\omega_{12}\tau} \qquad \text{with } \omega_{12} = \frac{H_{12}}{\hbar} \text{ and } \tau = \frac{|H_{12}|}{v|F_1 - F_2|} \tag{F.1}$$

where v is the velocity and H_{12} is the coupling between the two crossing total-energy surfaces. ω_{12} is the **Rabi frequency** at the crossing point and τ is a measure for the duration of the interaction between the surfaces. The model hamiltonian underlying the Landau-Zener formula has the form

Appendix G

Supplementary material for the Jahn-Teller model

This appendix extends the description of the Jahn-Teller model described in section 2.3.4 on p. 31.

The stationary wave functions of the Jahn-Teller problem have been given in Eq. 2.62 on p. 40

$$|\Phi_{j,m}\rangle \overset{\text{Eq. 2.62}}{=} \int dX \int dZ \, \frac{1}{2} \Big(\, |a,X,Z\rangle \Big[(\mathrm{e}^{i\alpha}+1)\mathcal{R}_{+,j,m}(R) + i(\mathrm{e}^{i\alpha}-1)\mathcal{R}_{-,j,m}(R)\Big] + |b,X,Z\rangle \Big[-i(\mathrm{e}^{i\alpha}-1)\mathcal{R}_{+,j,m}(R) + (\mathrm{e}^{i\alpha}+1)\mathcal{R}_{-,j,m}(R)\Big]\Big) \mathrm{e}^{im\alpha} \tag{G.1}$$

where the radial nuclear wave functions obey the one-dimensional, two-component radial Schrödinger equation Eq. 2.63 for the generalized nuclear coordinates (X, Z) of the Jahn-Teller model,

$$0 \overset{\text{Eq. 2.63}}{=} \Big\{ \underbrace{\frac{-\hbar^2\left(R^2\partial_R^2 + R\partial_R - m^2\right)}{2MR^2}\mathbf{1}}_{-\hbar^2\vec{\nabla}^2/(2M)} + \underbrace{gR\boldsymbol{\sigma}_z + wR^2\mathbf{1}}_{E^{BO}} + \underbrace{\frac{\hbar^2\left(m+\frac{1}{2}\right)}{2MR^2}(\mathbf{1}-\boldsymbol{\sigma}_x)}_{\text{nonBO}} - \mathcal{E}_{j,m}\mathbf{1}\Big\} \begin{pmatrix} i\mathcal{R}_{+,j,m}(R) \\ \mathcal{R}_{-,j,m}(R) \end{pmatrix} \tag{G.2}$$

G.1 Relation of stationary wave functions for m and $-m-1$

In Eq. 2.64 and Eq. 2.65 on p. 41 we mentioned that the electron-nuclear wave functions $\Phi_{j,m}(\vec{x}, X, Z)$ and $\Phi_{j,-m-1}(\vec{x}, X, Z)$ are complex conjugates of each other.[1]

$$\Phi_{j,-m-1}(\vec{x}, X, Z) = \Phi^*_{j,m}(\vec{x}, X, Z) \tag{G.3}$$

and that the same holds for the radial functions

$$\Big(\mathcal{R}_{\pm,j,-m-1}(R)\Big)^* = \mathcal{R}_{\pm,j,m}(R) \qquad \Leftrightarrow \qquad \begin{cases} \mathcal{R}_{-,j,-m-1}(R) = -\mathcal{R}_{-,j,m}(R) \\ \mathcal{R}_{+,j,-m-1}(R) = +\mathcal{R}_{+,j,m}(R) \end{cases} \tag{G.4}$$

Here, we provide the detailed derivation.

---

[1] More precisely, they can be chosen to be complex conjugates of each other.

Complex conjugation of the radial nuclear wave function: We start from the equation Eq. 2.63 for $\mathcal{R}_{\pm,j,-m-1}$

$$\begin{aligned}
&\Biggl\{\underbrace{\frac{-\hbar^2\left(R^2\partial_R^2+R\partial_R-(-m-1)^2\right)}{2MR^2}\mathbf{1}}_{-\hbar^2\vec{\nabla}^2/(2M)}+\underbrace{gR\boldsymbol{\sigma}_z+wR^2\mathbf{1}}_{E^{BO}}\\
&\qquad+\underbrace{\frac{\hbar^2\left((-m-1)+\frac{1}{2}\right)}{2MR^2}(\mathbf{1}-\boldsymbol{\sigma}_x)}_{\text{nonBO}}-\mathcal{E}_{j,m}\mathbf{1}\Biggr\}\begin{pmatrix}i\mathcal{R}_{+,j,-m-1,}(R)\\ \mathcal{R}_{-,j,-m-1}(R)\end{pmatrix}\overset{\text{Eq. 2.63}}{=}0\\
\Rightarrow&\Biggl\{\frac{-\hbar^2\left(R^2\partial_R^2+R\partial_R-m^2\right)+\overbrace{2\hbar^2(m+\frac{1}{2})}^{2A}}{2MR^2}\mathbf{1}+gR\boldsymbol{\sigma}_z+wR^2\mathbf{1}\\
&\qquad+\frac{\hbar^2\left(m+\frac{1}{2}\right)}{2MR^2}(\underbrace{-\mathbf{1}}_{-A}+\underbrace{\boldsymbol{\sigma}_x}_{B})-\mathcal{E}_{j,m}\mathbf{1}\Biggr\}\begin{pmatrix}i\mathcal{R}_{+,j,-m-1,}(R)\\ \mathcal{R}_{-,j,-m-1}(R)\end{pmatrix}=0
\end{aligned}\tag{G.5}$$

The term denoted $2A$ combines with the term $-A$. The term B is the only coupling term between the two components of the wave function. A sign change of this term can be absorbed by a sign change of one component relative to the other. Thus we arrive at

$$\begin{aligned}
&\Biggl\{\frac{-\hbar^2\left(R^2\partial_R^2+R\partial_R-m^2\right)}{2MR^2}\mathbf{1}+gR\boldsymbol{\sigma}_z+wR^2\mathbf{1}\\
&\qquad+\frac{\hbar^2\left(m+\frac{1}{2}\right)}{2MR^2}(\mathbf{1}-\boldsymbol{\sigma}_x)-\mathcal{E}_{j,m}\mathbf{1}\Biggr\}\begin{pmatrix}-i\mathcal{R}_{+,j,-m-1,}(R)\\ \mathcal{R}_{-,j,-m-1}(R)\end{pmatrix}=0
\end{aligned}\tag{G.6}$$

This is the equation for the index m. Thus, if the solution for m is known, we immediately know the solution for $-m-1$ by the transformation Eq. 2.65, which completes the proof of Eq. G.4.

Complex conjugation of the electron-nuclear wave function: Next, we need to show that also the electron nuclear wave function Eq. G.1 obey

$$\langle\vec{x},X,Z|\Phi_{j,m}\rangle=\langle\vec{x},X,Z|\Phi_{j,-m-1}\rangle^*\tag{G.7}$$

We start with Eq. G.1 with m replaced by $-m-1$. We use that $\mathcal{R}_{+,j,m}$ is purely imaginary and that $\mathcal{R}_{-,j,m}$ is purely real.

$$\begin{aligned}
\langle\vec{x},X,Z|\Phi_{j,-m-1}\rangle&\overset{\text{Eq. G.1}}{=}\frac{1}{2}\Biggl(\langle\vec{x}|a\rangle\Bigl[(\mathrm{e}^{i\alpha}+1)\mathcal{R}_{+,j,-m-1}(R)+i(\mathrm{e}^{i\alpha}-1)\mathcal{R}_{-,j,-m-1}(R)\Bigr]\\
&\qquad+\langle\vec{x}|b\rangle\Bigl[-i(\mathrm{e}^{i\alpha}-1)\underbrace{\mathcal{R}_{+,j,-m-1}(R)}_{\mathcal{R}^*_{+,j,m}(R)}+(\mathrm{e}^{i\alpha}+1)\underbrace{\mathcal{R}_{+,j,-m-1}(R)}_{\mathcal{R}^*_{+,j,m}(R)}\Bigr]\Biggr)\mathrm{e}^{i(-m-1)\alpha}\\
&\overset{\text{Eq. G.4}}{=}\Biggl[\frac{1}{2}\Biggl(\langle\vec{x}|a\rangle^*\Bigl[(\mathrm{e}^{-i\alpha}+1)\mathcal{R}_{+,j,m}(R)-i(\mathrm{e}^{-i\alpha}-1)\mathcal{R}_{-,j,m}(R)\Bigr]\\
&\qquad+\langle\vec{x}|b\rangle^*\Bigl[i(\mathrm{e}^{-i\alpha}-1)\mathcal{R}_{+,j,-m-1}(R)+(\mathrm{e}^{-i\alpha}+1)\mathcal{R}_{-,j,-m-1}(R)\Bigr]\Biggr)\mathrm{e}^{i(m+1)\alpha}\Biggr]^*\\
&=\Biggl[\frac{1}{2}\Biggl(\langle\vec{x}|a\rangle^*\Bigl[(\mathrm{e}^{i\alpha}+1)\mathcal{R}_{+,j,m}(R)+i(\mathrm{e}^{i\alpha}-1)\mathcal{R}_{-,j,m}(R)\Bigr]\\
&\qquad-\langle\vec{x}|b\rangle^*\Bigl[+i(\mathrm{e}^{i\alpha}-1)\mathcal{R}_{+,j,m}(R)+(\mathrm{e}^{i\alpha}+1)\mathcal{R}_{-,j,m}(R)\Bigr]\Biggr)\mathrm{e}^{im\alpha}\Biggr]^*\\
&\overset{\text{Eq. G.1}}{=}\langle\vec{x},X,Z|\Phi_{j,m}\rangle^*
\end{aligned}\tag{G.8}$$

which completes the proof that the electron-nuclear wave functions for m and $-m-1$ are complex conjugates of each other.

G.2 Behavior of the radial nuclear wave function at the origin

In order to obtain a solution of a radial Schrödinger equation such as Eq. 2.63, it is always helpful to know the behavior at the origin, which will be shown here.

At the origin the derivative couplings and the kinetic energy dominate the behavior of the wave function. Let me analyze this behavior. We start with a Taylor expansion:

$$\mathcal{R}_\pm = \sum_n a_{\pm,n} R^n \tag{G.9}$$

Insertion of the power-series ansatz into Eq. 2.63 yields

$$\begin{aligned} 0 &= \left\{ \frac{-\hbar^2\left(n(n-1)+n-m^2\right)}{2M}\mathbf{1} + \underbrace{\frac{\hbar^2\left(m+\frac{1}{2}\right)}{2M}(\mathbf{1}-\boldsymbol{\sigma}_x)}_{\text{nonBO}} \right\} \begin{pmatrix} ia_{+,n} \\ a_{-,n} \end{pmatrix} \\ &\quad + g\boldsymbol{\sigma}_z \begin{pmatrix} ia_{+,n-3} \\ a_{-,n-3} \end{pmatrix} + w \begin{pmatrix} ia_{+,n-4} \\ a_{-,n-4} \end{pmatrix} - \mathcal{E} \begin{pmatrix} ia_{+,n-2} \\ a_{-,n-2} \end{pmatrix} \\ &= \frac{\hbar^2}{2M} \left\{ \left(-n^2+m^2\right)\mathbf{1} + \underbrace{\left(m+\frac{1}{2}\right)(\mathbf{1}-\boldsymbol{\sigma}_x)}_{\text{nonBO}} \right\} \begin{pmatrix} ia_{+,n} \\ a_{-,n} \end{pmatrix} \\ &\quad + g\boldsymbol{\sigma}_z \begin{pmatrix} ia_{+,n-3} \\ a_{-,n-3} \end{pmatrix} + w \begin{pmatrix} ia_{+,n-4} \\ a_{-,n-4} \end{pmatrix} - \mathcal{E} \begin{pmatrix} ia_{+,n-2} \\ a_{-,n-2} \end{pmatrix} \end{aligned} \tag{G.10}$$

This equation can be written as a recursive equation for the coefficients with increasing order

$$\begin{aligned} \begin{pmatrix} ia_{+,n} \\ a_{-,n} \end{pmatrix} &= \frac{2M}{\hbar^2} \left\{ \left(-n^2+m^2\right)\mathbf{1} + \underbrace{\left(m+\frac{1}{2}\right)(\mathbf{1}-\boldsymbol{\sigma}_x)}_{\text{nonBO}} \right\}^{-1} \\ &\quad \times \left\{ w \begin{pmatrix} ia_{+,n-4} \\ a_{-,n-4} \end{pmatrix} + g\boldsymbol{\sigma}_z \begin{pmatrix} ia_{+,n-3} \\ a_{-,n-3} \end{pmatrix} - \mathcal{E} \begin{pmatrix} ia_{+,n-2} \\ a_{-,n-2} \end{pmatrix} \right\} \end{aligned} \tag{G.11}$$

We start the expansion with zero coefficients, i.e. $a_{\pm,n} = 0$ for all n below a minimum. The iteration produces non-zero elements only if the inverted matrix is singular. Thus we need to look for the indices where an eigenvalue of $\left(-n^2+m^2\right)\mathbf{1} + \left(m+\frac{1}{2}\right)(\mathbf{1}-\boldsymbol{\sigma}_x)$ vanishes.

Thus the leading order of the nuclear wave function is determined by the zero of the determinant of the above matrix.

$$\begin{aligned} \det \begin{vmatrix} -n^2+m^2+m+\frac{1}{2} & -\left(m+\frac{1}{2}\right) \\ -\left(m+\frac{1}{2}\right) & -n^2+m^2+m+\frac{1}{2} \end{vmatrix} &= 0 \\ \Rightarrow \quad \left(-n^2+m^2+m+\frac{1}{2}\right)^2 - \left(m+\frac{1}{2}\right)^2 &= 0 \\ \Rightarrow \quad n^2 = m^2+m+\frac{1}{2} \pm \left(m+\frac{1}{2}\right) & \\ \Rightarrow \quad n = \pm(m+1) \qquad \text{or} \qquad n = \pm m & \end{aligned} \tag{G.12}$$

Thus we obtain four solutions for the leading order n, of which two correspond to regular solutions of the differential equation and two further ones correspond to the irregular solutions.

m	$n_{reg,1}$	$n_{reg,2}$	$n_{irr,1}$	$n_{irr,2}$
$m \geq 0$	m	$m+1$	$-m$	$-m+1$
2	2	3	-2	-3
1	1	2	-1	-2
0	0	1	0	-1
-1	0	1	0	-1
-2	1	2	-1	-2
-3	2	3	-2	-3
-|m|	|m|-1	|m|	-|m|+1	-|m|

What is the meaning of having two orders where the power series may start? They correspond to two independent solutions. Two independent solutions rather than one occur because the nuclear Schrödinger equation of the Jahn-Teller problem is a two-component equation. Giving the leading order n_0 the next higher power determined by the recursion Eq. G.11 is n_0+2. An independent power series expansion can be started with n_0+1.

Next we need to determine the null vectors. We distinguish the cases with $m \geq 0$ and with $m < 0$. (Note that the eigenvectors of the inverse are the eigenvectors of the original matrix.)

- first we consider $m \geq 0$:

$$\begin{aligned} 0 &= \left(-n^2+m^2\right)\mathbf{1} + \underbrace{\left(m+\frac{1}{2}\right)(\mathbf{1}-\boldsymbol{\sigma}_x)}_{\text{nonBO}}\Big\} \begin{pmatrix} ia_{+,n} \\ a_{-,n} \end{pmatrix} \\ &\overset{n=m}{=} \left(m+\frac{1}{2}\right)\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} ia_{+,m} \\ a_{-,m} \end{pmatrix} \qquad \Rightarrow \qquad \begin{pmatrix} ia_{+,m} \\ a_{-,m} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} A_m \end{aligned} \tag{G.13}$$

where A_m is an arbitrary coefficient. Now we consider the second regular solution

$$\begin{aligned} 0 &= \left(-n^2+m^2\right)\mathbf{1} + \underbrace{\left(m+\frac{1}{2}\right)(\mathbf{1}-\boldsymbol{\sigma}_x)}_{\text{nonBO}}\Big\} \begin{pmatrix} ia_{+,n} \\ a_{-,n} \end{pmatrix} \\ &\overset{n=m+1}{=} -\left(m+\frac{1}{2}\right)\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} ia_{+,m} \\ a_{-,m} \end{pmatrix} \qquad \Rightarrow \qquad \begin{pmatrix} ia_{+,m} \\ a_{-,m} \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} B_m \cdot A_m \end{aligned} \tag{G.14}$$

where B is another arbitrary coefficient.

Thus the wave function for $m \geq 0$ has the form

$$\begin{pmatrix} i\mathcal{R}_+(R) \\ \mathcal{R}_-(R) \end{pmatrix} = AR^m \begin{pmatrix} 1 - BR + O(R^2) \\ 1 + BR + O(R^2) \end{pmatrix} \tag{G.15}$$

- now we consider $m < 0$:

$$\begin{aligned} 0 &= \left(-n^2+m^2\right)\mathbf{1} + \underbrace{\left(m+\frac{1}{2}\right)(\mathbf{1}-\boldsymbol{\sigma}_x)}_{\text{nonBO}}\Big\} \begin{pmatrix} ia_{+,n} \\ a_{-,n} \end{pmatrix} \\ &\overset{n=-m-1}{=} -\left(m+\frac{1}{2}\right)\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} ia_{+,m} \\ a_{-,m} \end{pmatrix} \qquad \Rightarrow \qquad \begin{pmatrix} ia_{+,m} \\ a_{-,m} \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} A_m \end{aligned} \tag{G.16}$$

where A_m is an arbitrary coefficient.

$$
\begin{aligned}
0 &= \left.\left(-n^2+m^2\right)\mathbf{1} + \underbrace{\left(m+\frac{1}{2}\right)\left(\mathbf{1}-\boldsymbol{\sigma}_x\right)}_{\text{nonBO}}\right\} \begin{pmatrix} ia_{+,n} \\ a_{-,n} \end{pmatrix} \\
&\overset{n=-m}{=} \left(m+\frac{1}{2}\right)\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} ia_{+,m} \\ a_{-,m} \end{pmatrix} \qquad \Rightarrow \qquad \begin{pmatrix} ia_{+,m} \\ a_{-,m} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} B_m \cdot A_m \qquad \text{(G.17)}
\end{aligned}
$$

where B_m is another arbitrary coefficient.

Thus the wave function for $m \geq 0$ has the form

$$
\begin{pmatrix} i\mathcal{R}_+(R) \\ \mathcal{R}_-(R) \end{pmatrix} = A_m R^m \begin{pmatrix} -1 + B_m R + O(R^2) \\ 1 + B_m R + O(R^2) \end{pmatrix} \qquad \text{(G.18)}
$$

We observe that we can choose the coefficients A_m, B_m such that the first terms of the power series expansion for m and $-m-1$ are complex conjugates of each other. We will see below in section G.1 that this is a general property of the solutions.

Another observation is that the nuclear wave function is not differentiable in all orders at the origin. Rather the $m =$ solution exhibits a kink at the origin.

G.3 Optimization of the radial wave functions

There are different ways to solve the nuclear radial Schrödinger equation numerically, which are discussed in the following two sections.

The one described here exploits Ritz' variational principle. It starts from an expression for the total energy as functional of a nuclear wave function. By minimizing this energy we obtain the ground state. Minimizing the same functional again again, but with the restriction that it is orthogonal to the ground state, will yield the first excited state. Similarly, we can successively build up all radial wave functions.

We write down the total energy of an ensemble of electron and nuclear states. The states can be expressed entirely by the nuclear radial functions. The occupations, which are denoted as $P_{m,j}$. The are Boltzmann weights.

In order to deal with real functions we introduce

$$
\begin{pmatrix} f_+ \\ f_- \end{pmatrix} \overset{\text{def}}{=} \begin{pmatrix} i\mathcal{R}_+ \\ \mathcal{R}_- \end{pmatrix} \qquad \text{(G.19)}
$$

The total energy functional to be minimized is obtained from the differential equation. Here we include all excited states, but give them probabilities $P_{m,j}$. The probabilities can be chosen fairly arbitrarily. While they are inspired by the expressions for a thermal ensemble, they are used here only as a numerical tool. Without proof let me state the following: (1) If all probabilities differ from each other, the resulting wave functions are eigenstates of the stationary Schrödinger equation. (2) The states are ordered such that the energies for states with higher probability are lower.

$$
\begin{aligned}
E[\vec{f}_{j,m}(R)] = &\sum_{m,j} P_{m,j} \int_0^\infty dR \underbrace{(2\pi R)}_{\text{ang. integr.}} \\
&\times \left\{ \vec{f}^*_{j,m}(R) \left[\frac{-\hbar^2\left(R^2\partial_R^2 + R\partial_R - m^2\right)}{2MR^2}\mathbf{1} + gR\boldsymbol{\sigma}_z + wR^2\mathbf{1} + \frac{\hbar^2(m+\frac{1}{2})}{2MR^2}\left(\mathbf{1}+\boldsymbol{\sigma}_x\right) \right] \vec{f}_{j,m}(R) \right\} \\
&+ \sum_m \sum_{j,j'} \Lambda_{m,j',j} \left\{ \left[\int_0^\infty dR \underbrace{(2\pi R)}_{\text{ang. integr.}} \vec{f}^*_{j,m}\vec{f}_{j',m} \right] - \delta_{j,j'} \right\} \qquad \text{(G.20)}
\end{aligned}
$$

This expression acts as potential energy in a fictitious Lagrangian

$$\mathcal{L}[\{\vec{f}_{j,m}(R),\dot{\vec{f}}_{j,m}(R)\}]=\sum_{m,j}P_{m,j}\int drm_f\dot{\vec{f}}_{j,m}(R)\dot{\vec{f}}_{j,m}(R)-E[\{\vec{f}_{jm}(R)]\tag{G.21}$$

and derive an equation of motion for the radial wave function in the spirit of the Car-Parrinello method[87]. The time defining the "velocity of $\vec{f}_{m,j}$", is not a real time but simply a coordinate describing dynamics of the wave functions during the optimization process.

The first variation of the total energy with respect to the radial functions is

$$dE=\sum_{m,j}P_{m,j}\int_0^\infty dR\ (2\pi R)\,d\vec{f}^*_{j,m}\vec{y}_{j,m}+\text{c.c.}\tag{G.22}$$

which yields an effective gradient of the functional

$$\begin{aligned}\vec{y}_{j,m}=\Big\{&\left[\frac{-\hbar^2\left(R^2\partial_R^2+R\partial_R-m^2\right)}{2MR^2}\mathbf{1}+gR\boldsymbol{\sigma}_z+wR^2\mathbf{1}+\frac{\hbar^2(m+\frac{1}{2})}{2MR^2}\left(\mathbf{1}+\boldsymbol{\sigma}_x\right)\right]\vec{f}_{j,m}(R)\\&+\sum_{j'}\vec{f}_{j',m}\Lambda_{m,j',j}\Big\}\end{aligned}\tag{G.23}$$

The Euler Lagrange equations for the Lagrangian Eq. G.21 produce an equation of motion

$$m_f\ddot{\vec{f}}_{m,j}(R)=-\vec{y}_{j,m}(R)\tag{G.24}$$

We did not consider here the constraint term, because we will enforce this constraint separately. This equation of motion is augmented by a friction term and discretized using the Verlet algorithm[88].

In the **Verlet algorithm**[88] a differential equation is discretized by replacing the first and second derivatives by the differential quotients

$$\begin{aligned}\dot{x}&\to\frac{x(t+\Delta)-x(t-\Delta)}{2\Delta}\\\ddot{x}&\to\frac{x(t+\Delta)-2x(t)+x(t-\Delta)}{\Delta^2}\end{aligned}\tag{G.25}$$

This allows to iteratively obtain the solution of a differential equation $m\ddot{x}=F(x)-\gamma\dot{x}$ for $x(j\Delta)$.

$$x(t+\Delta)=x(t)\frac{2}{1+a}-x(t-\Delta)\frac{1-a}{1+a}+F\frac{\Delta^2}{m(1+a)}\tag{G.26}$$

with $a=\frac{\gamma\Delta}{2m}$.

We start from an arbitrary wave function $\vec{f}_{j,m}(R)$ for each m and j. The wave function is first orthogonalized to all functions with lower j, and then it is normalized. Starting from this initial wave function we perform a Car-Parrinello like optimization in the space orthogonal to the lower-lying states.

1. calculate the gradient $\vec{y}_{j,m}(R)$

$$\vec{y}_{j,m}=\Big\{\left[\frac{-\hbar^2\left(R^2\partial_R^2+R\partial_R-m^2\right)}{2MR^2}\mathbf{1}+gR\boldsymbol{\sigma}_z+wR^2\mathbf{1}+\frac{\hbar^2(m+\frac{1}{2})}{2MR^2}\left(\mathbf{1}+\boldsymbol{\sigma}_x\right)\right]\vec{f}_{j,m}(R)\tag{G.27}$$

2. orthogonalize $\vec{y}_{j,m}(R)$ to all lower-lying functions

$$\vec{y}_{j,m}\leftarrow\vec{y}_{j,m}-\sum_{j'=1}^{j-1}\vec{f}_{j',m}\Big(\int_0^\infty dR\ R\ \vec{f}^*_{j',m}(R)\vec{y}_{j,m}(R)\tag{G.28}$$

3. propagate with the Verlet algorithm

$$\bar{\vec{f}} = \vec{f}(t)\frac{2}{1+a} - \vec{f}(t-\Delta)\frac{1-a}{1+a} - \vec{y}\frac{\Delta^2}{2m_f(1+a)} \tag{G.29}$$

4. normalize

$$\begin{aligned}
\vec{f}(t+\Delta) &= \bar{\vec{f}} + \vec{f}(t)\lambda \\
1 &\overset{!}{=} 2\pi\int_0^\infty dR\, R\, \vec{f}^*(+)\vec{f}(+) \\
&= \underbrace{\Big(2\pi\int_0^\infty dR\, R\, \bar{\vec{f}}^*\bar{\vec{f}}(+)\Big) - 1}_{C} + 2\lambda\,\underbrace{\mathrm{Re}\Big(2\pi\int_0^\infty dR\, R\, \bar{\vec{f}}^*\vec{f}(0)\Big)}_{B} + \lambda^2\,\underbrace{2\pi\int_0^\infty dR\, R\, \vec{f}^*(0)\vec{f}(0)}_{A} \\
\lambda &= -\frac{B}{A}\Big[1-\sqrt{1-\frac{A^2C}{B^2}}\Big]
\end{aligned} \tag{G.30}$$

Since the function is known for the next time step, we can begin again with the first point of the iteration sequence.

G.4 Integration of the Schrödinger equation for the radial nuclear wave functions

Editor: This is not finished!

$$0 = \Big\{-\frac{\hbar^2}{2M}\partial_R^2 - \frac{\hbar^2}{2MR}\partial_R + \frac{\hbar^2\left[m(m+1)+\frac{1}{2}\right]}{2MR^2} \pm g|\vec{R}| + w\vec{R}^2 - \mathcal{E}_{m,j}\Big\}\mathcal{R}_{\pm,m,j} \pm \frac{i\hbar^2(m+\frac{1}{2})}{2MR^2}\mathcal{R}_{\mp,m,j}$$

$$0 = \Big\{-R^2\partial_R^2 - R\partial_R + \Big[m(m+1)+\frac{1}{2}\Big] + \frac{2M}{\hbar^2}\Big[-\mathcal{E}_{m,j}R^2 \pm gR^3 + wR^4\Big]\Big\}\mathcal{R}_{\pm,m,j} \pm i\left(m+\frac{1}{2}\right)\mathcal{R}_{\mp,m,j}$$

Now we introduce an exponential Ansatz $R(x) = \mathrm{e}^x$

$$\mathcal{R}_{\pm,m,j}(\mathrm{e}^x) = f_{\pm,m,j}(x) \tag{G.31}$$

$$\begin{aligned}
R(x) &= \mathrm{e}^x \qquad \Rightarrow \qquad \frac{dR}{dx} = \mathrm{e}^x = R \qquad \Rightarrow \qquad \frac{dx}{dR} = 1/R \\
R\partial_R &= \partial_x \\
R^2\partial_R^2 &= R^2\partial_R\frac{1}{R}R\partial_R = R^2\partial_R\frac{1}{R}\partial_x = R^2\Big(-\frac{1}{R^2} + \frac{1}{R^2}R\partial_R\Big)\partial_x = -\partial_x + \partial_x^2 \\
R^2\partial_R^2 + R\partial_R &= \partial_x^2
\end{aligned} \tag{G.32}$$

This yields

$$0 = \Big\{-\partial_x^2 + \Big[m(m-1)+\frac{1}{2}\Big] + \frac{2M}{\hbar^2}\Big[-\mathcal{E}_{m,j}R^2 \pm gR^3 + wR^4\Big]\Big\}f_{\pm,m,j}(x) \mp i\left(m+\frac{1}{2}\right)f_{\mp,m,j}(x)$$

We solve this equation numerically on a grid using the Verlet algorithm[88], which replaces the second derivative with the differential quotient

$$\partial_x^2 f = \frac{f(+) - 2f(0) + f(-)}{\alpha^2} \tag{G.33}$$

where α is the grid spacing.

An important aspect is that one needs to integrate from the outside inward. The reason is that in the outer region, there is one solution that falls of towards the outside and one solution that increases exponentially. We are interested in the solution falling off.

When we integrate outward, any little numerical error will admix a fraction of the exponentially increasing solution. This small fraction will now be exponentially amplified, leading to a poor solution. On the other hand, when we integrate inward, every error will admix a partial solution that falls off in the inward direction, so that any error heals immediately.

Therefore we choose the strategy to integrate from the outside inward and match the solutions in the center.

For a given energy, we calculate the Schrödinger equation outward from the inside with two distinct boundary conditions and obtain $\Big(f^{(i)}_{1,1}(x), f^{(i)}_{2,1}(x)\Big)$ and $\Big(f^{(i)}_{1,2}(x), f^{(i)}_{2,2}(x)\Big)$. At the same time I evaluate the energy derivatives yielding $\Big(\dot{f}^{(i)}_{1,1}(x), \dot{f}^{(i)}_{2,1}(x)\Big)$ and $\Big(\dot{f}^{(i)}_{1,2}(x), \dot{f}^{(i)}_{2,2}(x)\Big)$. The boundary conditions are $\Big(f^{(i)}_{1,1}(x), f^{(i)}_{2,1}(x)\Big) = (1,0)$ and $\Big(f^{(i)}_{1,2}(x), f^{(i)}_{2,2}(x)\Big) = (0,1)$ at the first grid point. The derivatives are set to zero. For the energy derivative functions value and derivative at the first grid points are set to zero.

The condition that the two functions match with value and derivative can be approximated by the condition that the two functions match on two subsequent points x_k on the grid.

$$\sum_{\alpha=1}^{2}\Big(f^{(i)}_{i,\alpha}(x_k) + (\epsilon-\epsilon_0)\dot{f}^{(i)}_{i,\alpha}(x_k)\Big)c_\alpha = \sum_{\alpha=1}^{2}\Big(f^{(o)}_{i,\alpha}(x_k) + (\epsilon-\epsilon_0)\dot{f}^{(o)}_{i,\alpha}(x_k)\Big)c_{\alpha+2} \tag{G.34}$$

This value must be fulfilled for $k = 1,2$ and $i = 1,2$. This is an eigenvalue equation

$$\begin{pmatrix} f^{(i)}_{1,1}(x_a) & f^{(i)}_{1,2}(x_a) & -f^{(o)}_{1,1}(x_a) & -f^{(o)}_{1,2}(x_a) \\ f^{(i)}_{2,1}(x_a) & f^{(i)}_{2,2}(x_a) & -f^{(o)}_{2,1}(x_a) & -f^{(o)}_{2,2}(x_a) \\ f^{(i)}_{1,1}(x_b) & f^{(i)}_{1,2}(x_b) & -f^{(o)}_{1,1}(x_b) & -f^{(o)}_{1,2}(x_b) \\ f^{(i)}_{2,1}(x_b) & f^{(i)}_{2,2}(x_b) & -f^{(o)}_{2,1}(x_b) & -f^{(o)}_{2,2}(x_b) \end{pmatrix} + (\epsilon-\epsilon_0)\begin{pmatrix} \dot{f}^{(i)}_{1,1}(x_a) & \dot{f}^{(i)}_{1,2}(x_a) & -\dot{f}^{(o)}_{1,1}(x_a) & -\dot{f}^{(o)}_{1,2}(x_a) \\ \dot{f}^{(i)}_{2,1}(x_a) & \dot{f}^{(i)}_{2,2}(x_a) & -\dot{f}^{(o)}_{2,1}(x_a) & -\dot{f}^{(o)}_{2,2}(x_a) \\ \dot{f}^{(i)}_{1,1}(x_b) & \dot{f}^{(i)}_{1,2}(x_b) & -\dot{f}^{(o)}_{1,1}(x_b) & -\dot{f}^{(o)}_{1,2}(x_b) \\ \dot{f}^{(i)}_{2,1}(x_b) & \dot{f}^{(i)}_{2,2}(x_b) & -\dot{f}^{(o)}_{2,1}(x_b) & -\dot{f}^{(o)}_{2,2}(x_b) \end{pmatrix} = 0 \tag{G.35}$$

We define the matrix

$$\boldsymbol{A} \stackrel{\text{def}}{=} \begin{pmatrix} \dot{f}^{(i)}_{1,1}(x_a) & \dot{f}^{(i)}_{1,2}(x_a) & -\dot{f}^{(o)}_{1,1}(x_a) & -\dot{f}^{(o)}_{1,2}(x_a) \\ \dot{f}^{(i)}_{2,1}(x_a) & \dot{f}^{(i)}_{2,2}(x_a) & -\dot{f}^{(o)}_{2,1}(x_a) & -\dot{f}^{(o)}_{2,2}(x_a) \\ \dot{f}^{(i)}_{1,1}(x_b) & \dot{f}^{(i)}_{1,2}(x_b) & -\dot{f}^{(o)}_{1,1}(x_b) & -\dot{f}^{(o)}_{1,2}(x_b) \\ \dot{f}^{(i)}_{2,1}(x_b) & \dot{f}^{(i)}_{2,2}(x_b) & -\dot{f}^{(o)}_{2,1}(x_b) & -\dot{f}^{(o)}_{2,2}(x_b) \end{pmatrix}^{-1} \begin{pmatrix} f^{(i)}_{1,1}(x_a) & f^{(i)}_{1,2}(x_a) & -f^{(o)}_{1,1}(x_a) & -f^{(o)}_{1,2}(x_a) \\ f^{(i)}_{2,1}(x_a) & f^{(i)}_{2,2}(x_a) & -f^{(o)}_{2,1}(x_a) & -f^{(o)}_{2,2}(x_a) \\ f^{(i)}_{1,1}(x_b) & f^{(i)}_{1,2}(x_b) & -f^{(o)}_{1,1}(x_b) & -f^{(o)}_{1,2}(x_b) \\ f^{(i)}_{2,1}(x_b) & f^{(i)}_{2,2}(x_b) & -f^{(o)}_{2,1}(x_b) & -f^{(o)}_{2,2}(x_b) \end{pmatrix} \tag{G.36}$$

and diagonalize it using the LAPACK routine ZGEEV. The eigenvalues are an estimate for the next energy eigenvalue. The eigenvalues are $\epsilon - \epsilon_0$. Using the eigenvalues only one can zoom into the correct energy. That is, ϵ_0 is adjusted to the new estimate for the energy eigenvalue until $\epsilon - \epsilon_0$ nearly vanishes for the chosen eigenvector.

Finally, one can evaluate the right-hand eigenvector, which provides the coefficients of the inner and outer partial solutions.

The final solution is then constructed

$$\begin{aligned} \vec{f}(x) &= \Big(\vec{f}^{(i)}_1 + (\epsilon-\epsilon_0)\dot{\vec{f}}^{(i)}_1\Big)c_1 + \Big(\vec{f}^{(i)}_2 + (\epsilon-\epsilon_0)\dot{\vec{f}}^{(i)}_2\Big)c_2 \\ \vec{f}(x) &= \Big(\vec{f}^{(o)}_1 + (\epsilon-\epsilon_0)\dot{\vec{f}}^{(o)}_1\Big)c_3 + \Big(\vec{f}^{(o)}_2 + (\epsilon-\epsilon_0)\dot{\vec{f}}^{(o)}_2\Big)c_4 \end{aligned} \tag{G.37}$$

Depending on the value of x, either the equation with the inner or the outer partial solutions is used.

The solution is then normalized so that

$$
\begin{aligned}
2\pi \int dR\, R \cdot R^{2m}\Big(f^*_{+,m,j}(x(R))f_{+,m,j}(x(R)) + f^*_{-,m,j}(x(R))f_{-,m,j}(x(R))\Big) &= 1 \\
2\pi \int_{-\infty}^{\infty} dx\, \frac{dR}{dx} R \cdot R^{2m}\Big(f^*_{+,m,j}(x(R))f_{+,m,j}(x(R)) + f^*_{-,m,j}(x(R))f_{-,m,j}(x(R))\Big) &= 1 \\
2\pi \int_{-\infty}^{\infty} dx\, \mathrm{e}^{(2m+2)x}\Big(f^*_{+,m,j}(x)f_{+,m,j}(x) + f^*_{-,m,j}(x)f_{-,m,j}(x)\Big) &= 1 \\
2\pi\alpha \sum_{x_i} \mathrm{e}^{(2m+2)x_i}\Big(f^*_{+,m,j}(x_i)f_{+,m,j}(x_i) + f^*_{-,m,j}(x_i)f_{-,m,j}(x_i)\Big) &= 1 \qquad \text{(G.38)}
\end{aligned}
$$

The factor $2\pi r$ results from the angular integration. Please check!

Appendix H

Notation for spin indices

We use a notation that combines the continuous spatial indices $\vec{r}$ with the discrete spin indices into a four dimensional vector $\vec{x}$. This shortcut notation has a more rigorous basis[61], which we will show here.

We introduce an artificial spin coordinate q. This is not a spatial coordinate, but it is a coordinate in some other abstract one-dimensional space. On this space, we assume that there is a complete[1] and orthonormal basis with only two functions, namely $\alpha(q)$ and $\beta(q)$. They obey

$$\int dq\, \alpha^*(q)\alpha(q) = 1$$
$$\int dq\, \beta^*(q)\beta(q) = 1$$
$$\int dq\, \alpha^*(q)\beta(q) = 0$$

A wave function may now depend on a spatial coordinate and this fictitious spin variable q. They form together a four dimensional vector

$$\vec{x} \overset{\text{def}}{=} (\vec{r}, q)$$

An electronic wave function depends on all these variables:

$$\langle \vec{x}|\psi\rangle = \langle \vec{r}, q|\psi\rangle = \psi(\vec{r}, q)$$

We can now decompose this vector into its components

$$\psi(\vec{r}, \uparrow) \overset{\text{def}}{=} \int dq\, \alpha^*(q)\psi(\vec{r}, q)$$
$$\psi(\vec{r}, \downarrow) \overset{\text{def}}{=} \int dq\, \beta^*(q)\psi(\vec{r}, q)$$

---

[1] This is where the argument is weak: a complete set of functions on a finite interval is always infinite. Probably one needs Grassman variables

Appendix I

Some group theory for symmetries

Symmetry operators are important to break down a Hamiltonian into a block diagonal matrix, which greatly simplifies the construction of the Hamilton eigenvalues. For this purpose we need to determine distinct sets of states which are eigenstates with respect to the symmetry operators of the Hamiltonian.

The following follows closely the Chemwiki
`http://chemwiki.ucdavis.edu/Theoretical_Chemistry/Symmetry/Group_Theory%3A_Theory`

I.0.1 Definitin of a group

The symmetry operators form a **group**.

A group consists of a set of elements and an operation. In our case the elements are the symmetry operations and the operations is the operator multiplication. To be a group the following properties need to be fulfilled.

- **Closure:** For any two members $\hat{A}$ and $\hat{B}$ of the group, their product $\hat{C} = \hat{A}\hat{B}$ is a member of the group.
- **Associativity:** The order in which the multiplications are performed does not affect the result.
$$\forall_{\hat{A},\hat{B},\hat{C}\in G}\ (\hat{A}\hat{B})\hat{C} = \hat{A}(\hat{B}\hat{C}) \tag{I.1}$$
- **Identity:** There is one element $\hat{1}$ in the group, called identity, which leaves all elements in the group unchanged when operated on it from the left or from the right.
$$\exists_{\hat{1}\in G}\forall_{\hat{A}\in G}\ \hat{1}\hat{A} = \hat{A}\hat{1} \tag{I.2}$$
- **Inverse:** For each element $\hat{A}$ in the group, there is one element in the group, called the inverse $\hat{A}^{-1}$ of $\hat{A}$, which produces the identity when operated on $\hat{A}$.
$$\forall_{\hat{A}\in G}\exists_{\hat{A}^{-1}\in G}\hat{A}^{-1}\hat{A} = \hat{A}\hat{A}^{-1} = \hat{1} \tag{I.3}$$

I.0.2 Symmetry class

A **similarity transform**[1] with an invertible operator $\hat{X}$ transforms an operator $\hat{A}$ into an operator $\hat{B}$ by

$$\hat{X}^{-1}\hat{A}\hat{X} = \hat{B} \tag{I.4}$$

---

[1] Source: `https://en.wikipedia.org/wiki/Matrix_similarity`, `http://mathworld.wolfram.com/SimilarityTransformation.html`.

To construct the symmetry class C_A containing operator A, calculate all operators obtained by a symilarity transfrom with any of the operators in the group.

$$C_A = \{\hat{X}^{-1}\hat{A}\hat{X}\}_{\forall\hat{X}\in G} \tag{I.5}$$

If $\hat{B}$ is in the symmetry class C_A containing $\hat{A}$, then $\hat{A}$ is member of the symmetry class C_B.

Proof: If $\hat{B}$ is member of $\hat{C}_A$, then there is one group operation $\hat{X}$ so that $\hat{B} = \hat{X}^{-1}\hat{A}\hat{X}$. Thus also $\hat{A} = \hat{X}\hat{B}\hat{X}^{-1}$. Because also the inverse $\hat{X}^{-1}$ is member of the group, $\hat{A}$ is member of the class C_B containing $\hat{B}$.

The classes produced by two operators in the same class are identical.

$$\hat{B} \in C_A \Rightarrow C_B = C_A \tag{I.6}$$

Appendix J

Downfolding

Often, a Schroedinger equation can be divided into an important and an unimportant portion of Hilbert space. This could be a dividion into an impurity, which is important, and which is embedded into an "unimportant" bath, which is spatially separated.

It could also be that we need to focus onto certain bands in a band structure of non-interacting electrons in a periodic potential, which the "unimportant" states shall not be completely ignored.

Another application is that of a many-particle problem, where few Slater-determinants are important. They constitute what chemists call "static correlation", while other Slater determinants are of secondary importance. Their contribution is termed "dynamic correlation" in the context of quantum chemistry.[1]

Let us consider a Hilbert (or Fock-) space, which is divided into an important subspace $A=\big\{|\chi_j^A\rangle\big\}$ with states $|\chi_j^A\rangle$ and an unimportant subspace $B=\big\{|\chi_j^B\rangle\big\}$ with states $|\chi_j^A\rangle$. The matrix elements of the hamiltonian $\hat{H}$ are

$$\begin{aligned}
H_{AA,i,j}&=\langle\chi_i^A|\hat{H}|\rangle\chi_j^A\rangle\\
H_{AB,i,j}&=\langle\chi_i^A|\hat{H}|\rangle\chi_j^B\rangle\\
H_{BA,i,j}&=\langle\chi_i^B|\hat{H}|\rangle\chi_j^A\rangle\\
H_{BB,i,j}&=\langle\chi_i^B|\hat{H}|\rangle\chi_j^B\rangle
\end{aligned}\tag{J.1}$$

The stationary Schrödinger equation has the form

$$\begin{pmatrix}\boldsymbol{H}_{AA}-\epsilon\boldsymbol{1}_{AA} & \boldsymbol{H}_{AB}\\ \boldsymbol{H}_{BA} & \boldsymbol{H}_{BB}-\epsilon\boldsymbol{1}_{BB}\end{pmatrix}\begin{pmatrix}\vec{c}_A\\ \vec{c}_B\end{pmatrix}=0\tag{J.2}$$

with the norm

$$\vec{c}_A^{\,2}+\vec{c}_B^{\,2}=1\tag{J.3}$$

We use the symbols $\boldsymbol{1}_{AA}$ and $\boldsymbol{1}_{BB}$ for the unit matrices in the subspaces A and B, respectively.

From the second line of the Schrödinger equation Eq. J.2, we obtain

$$\begin{aligned}
&\boldsymbol{H}_{BA}\vec{c}_A+(\boldsymbol{H}_{BB}-\epsilon\boldsymbol{1}_{BB})\vec{c}_B=0\\
\Rightarrow\quad&\vec{c}_B=-(\boldsymbol{H}_{BB}-\epsilon\boldsymbol{1}_{BB})^{-1}\boldsymbol{H}_{BA}\vec{c}_A
\end{aligned}\tag{J.4}$$

---

[1] In the context of solid state physics and Green's functions the notion of static and dynamic correlation is entirely different: Here a static correlation is attributed to a self energy that is frequency independent, while dynamic correlations are related with a frequency dependency of the self energy.

The result Eq. J.4 is inserted into the first line of the Schrödinger equation Eq. J.2

$$\begin{aligned} & (\boldsymbol{H}_{AA}-\epsilon\boldsymbol{1}_{AA})\vec{c}_A+\boldsymbol{H}_{AB}\vec{c}_B=0 \\ \overset{\text{Eq. J.4}}{\Rightarrow} \quad & \Big[\underbrace{\boldsymbol{H}_{AA}-\boldsymbol{H}_{AB}(\boldsymbol{H}_{BB}-\epsilon\boldsymbol{1}_{BB})^{-1}\boldsymbol{H}_{BA}}_{\boldsymbol{M}(\epsilon)}-\epsilon\boldsymbol{1}_{AA}\Big]\vec{c}_A=0 \end{aligned} \tag{J.5}$$

Let us introduce a new symbol, namely

$$\boldsymbol{M}(\epsilon)\stackrel{\text{def}}{=}\boldsymbol{H}_{AA}-\boldsymbol{H}_{AB}(\boldsymbol{H}_{BB}-\epsilon\boldsymbol{1}_{BB})^{-1}\boldsymbol{H}_{BA} \tag{J.6}$$

Eq. J.5 is an equation that acts only in A, the subspace of the important degrees of freedom. This is good. The price to pay is that the equation is now non-linear in the energy. We will come to that later.

We still need to translate the norm condition:

$$\begin{aligned} & \vec{c}_A^2+\vec{c}_B^2=1 \\ \overset{\text{Eq. J.4}}{\Rightarrow} \quad & \vec{c}_A\Big[\boldsymbol{1}_{AA}+\underbrace{\boldsymbol{H}_{AB}}_{\boldsymbol{H}_{BA}^\dagger}(\boldsymbol{H}_{BB}-\epsilon\boldsymbol{1}_{BB})^{-2}\boldsymbol{H}_{BA}\Big]\vec{c}_A=1 \end{aligned} \tag{J.7}$$

Let us introduce another new symbol, namely

$$\boldsymbol{O}(\epsilon)\stackrel{\text{def}}{=}\boldsymbol{1}_{AA}+\boldsymbol{H}_{AB}(\boldsymbol{H}_{BB}-\epsilon\boldsymbol{1}_{BB})^{-2}\boldsymbol{H}_{BA} \tag{J.8}$$

The matrix $\boldsymbol{O}(\epsilon)$ is a matrix in the subspace A can be considered an "overlap operator". It accounts for the fact that the complete wave function has also a contribution in subspace B, and therefore the probability weight inside A is less than one.

It turns out that the overlap operator $\boldsymbol{O}$ can be expressed as a derivative of the matrix $\boldsymbol{M}$.

$$\boldsymbol{O}(\epsilon)=\boldsymbol{1}_{AA}-\partial_\epsilon\boldsymbol{M}(\epsilon) \tag{J.9}$$

An eigenstate $|\psi\rangle$ of the full Hamiltonian can be expressed now in the form

$$\begin{aligned} |\psi\rangle &= \sum_{j\in A}|\chi_{A,j}\rangle c_{A,j}+\sum_{k\in B}|\chi_{B,k}\rangle c_{B,k} \\ &\overset{\text{Eq. J.4}}{=} \sum_{j\in A}|\chi_{A,j}\rangle c_{A,j}-\sum_{k,l,m\in B}\sum_{j\in A}|\chi_{B,k}\rangle(\boldsymbol{H}_{BB}-\epsilon\boldsymbol{1}_{BB})^{-1}_{l,m}H_{BA,m,j}c_{A,j} \\ &= \sum_{j\in A}\Big(|\chi_{A,j}\rangle-\sum_{k,l,m\in B}|\chi_{B,k}\rangle(\boldsymbol{H}_{BB}-\epsilon\boldsymbol{1}_{BB})^{-1}_{l,m}H_{BA,m,j}\Big)c_{A,j} \end{aligned} \tag{J.10}$$

Thus, we arrived at a definition of energy-dependent basis states

$$|\tilde{\chi}_{A,j}(\epsilon)\rangle=|\chi_{A,j}\rangle-\sum_{k,l,m\in B}|\chi_{B,k}\rangle(\boldsymbol{H}_{BB}-\epsilon\boldsymbol{1}_{BB})^{-1}_{l,m}H_{BA,m,j} \tag{J.11}$$

The "unimportant" contribtion from subspace B have been **downfolded** into the states connected to subsystem A.

The nonlinear matrix equation is recovered as

$$\boldsymbol{M}(\epsilon)-\boldsymbol{1}=\langle\chi_{A,j}|\hat{H}-\epsilon\boldsymbol{1}|\tilde{\chi}_{A,k}(\epsilon)\rangle=\langle\tilde{\chi}_{A,j}(\epsilon)|\hat{H}-\epsilon\boldsymbol{1}|\tilde{\chi}_{A,k}(\epsilon)\rangle=0 \tag{J.12}$$

J.1 Downfolding with Green's functions

The Green's function is defined as

$$\Big(\epsilon\mathbf{1}-\boldsymbol{H}\Big)\boldsymbol{G}(\epsilon)=\mathbf{1} \tag{J.13}$$

Note, that there are different conventions for the Green's function, that differ in the sign.

$$\begin{pmatrix}\epsilon\mathbf{1}_{AA}-\boldsymbol{H}_{AA} & -\boldsymbol{H}_{AB}\\ -\boldsymbol{H}_{BA} & \epsilon\mathbf{1}_{BB}-\boldsymbol{H}_{BB}\end{pmatrix}\begin{pmatrix}\boldsymbol{G}_{AA}(\epsilon) & \boldsymbol{G}_{AB}(\epsilon)\\ \boldsymbol{G}_{BA}(\epsilon) & \boldsymbol{G}_{BB}(\epsilon)\end{pmatrix}=\begin{pmatrix}\mathbf{1}_{AA} & \mathbf{0}_{AB}\\ \mathbf{0}_{BA} & \mathbf{1}_{BB}\end{pmatrix} \tag{J.14}$$

From the lower left submatrix of the above equation we obtain

$$\begin{aligned}&-\boldsymbol{H}_{BA}\boldsymbol{G}_{AA}(\epsilon)+(\epsilon\mathbf{1}_{BB}-\boldsymbol{H}_{BB})\boldsymbol{G}_{BA}(\epsilon)=\mathbf{0}\\ \Rightarrow\quad&\boldsymbol{G}_{BA}(\epsilon)=(\epsilon\mathbf{1}_{BB}-\boldsymbol{H}_{BB})^{-1}\boldsymbol{H}_{BA}\boldsymbol{G}_{AA}(\epsilon)\end{aligned} \tag{J.15}$$

This result we insert into the upper left submatrix

$$\begin{aligned}&\Big(\epsilon\mathbf{1}_{AA}-\boldsymbol{H}_{AA}\Big)\boldsymbol{G}_{AA}(\epsilon)-\boldsymbol{H}_{AB}\boldsymbol{G}_{BA}(\epsilon)=\mathbf{1}_{AA}\\ \Rightarrow\quad&\Big[\epsilon\mathbf{1}_{AA}-\boldsymbol{H}_{AA}-\boldsymbol{H}_{AB}\Big(\epsilon\mathbf{1}_{BB}-\boldsymbol{H}_{BB}\Big)^{-1}\boldsymbol{H}_{BA}\Big]\boldsymbol{G}_{AA}(\epsilon)=\mathbf{1}_{AA}\end{aligned} \tag{J.16}$$

This allows one to calculate the sub-block for the Green's function for subsystem A directly

$$\Rightarrow\quad\boldsymbol{G}_{AA}(\epsilon)=\Big[\epsilon\mathbf{1}_{AA}-\boldsymbol{H}_{AA}-\boldsymbol{H}_{AB}\Big(\epsilon\mathbf{1}_{BB}-\boldsymbol{H}_{BB}\Big)^{-1}\boldsymbol{H}_{BA}\Big]^{-1} \tag{J.17}$$

This method is uses for example to describe non-interacting electrons near interfaces: It allows to remove part of the system and replace it by another one.

J.2 Taylor expansion in the energy

In order to solve the equation

$$\left[\boldsymbol{M}(\epsilon)-\mathbf{1}\epsilon\right]\vec{c}=0 \tag{J.18}$$

we need to follow the eigenvalues of $\boldsymbol{M}(\epsilon)-\epsilon\mathbf{1}$ and determine those energies where one or more eigenvalues become zero. This can be done by computing the determinant of $\boldsymbol{M}(\epsilon)-\mathbf{1}\epsilon$ as function of energy and determine the zeros. The corresponding eigenvectors with zero eigenvalue are the solutions that occur at the specific energies.

This process is much more complicated that the diagonalization of an anergy-independent Hamiltonian. Therefore one usually searches for approximate solutions.

The most simple approach is the linearization about some energy ϵ_ν

$$\begin{aligned}\left[\boldsymbol{M}(\epsilon_\nu)+(\epsilon-\epsilon_\nu)\,\partial_\epsilon|_{\epsilon_\nu}\boldsymbol{M}-\mathbf{1}\epsilon\right]\vec{c}&=0\\ \Big[\Big(\boldsymbol{M}(\epsilon_\nu)-\epsilon_\nu\,\partial_\epsilon|_{\epsilon_\nu}\boldsymbol{M}\Big)-\Big(\mathbf{1}-\partial_\epsilon|_{\epsilon_\nu}\boldsymbol{M}\Big)\epsilon\Big]\vec{c}&=0\\ \Big[\Big(\boldsymbol{M}(\epsilon_\nu)+\epsilon_\nu[\boldsymbol{O}(\epsilon_\nu)-\mathbf{1}]\Big)-\boldsymbol{O}(\epsilon_\nu)\epsilon\Big]\vec{c}&=0\end{aligned} \tag{J.19}$$

Thus, for each energy ϵ_ν we have a linear Schrödinger equation. Those eigenstates that lie close to the expansion energy will be accurate.

The Taylor expansion however has one important caveat: Every Taylor expansion has a convergence radius beyond which the Taylor expansion does no more converge tp the correct solution. Usually the convergence radius is given by the distance of the next pole or discontinuity from the expansion point. In the present case the eigenstates of $\hat{H}_{BB}$, eigenstates of the isolated subsystem B introduce a singularity. Thus, at these energies the Taylor expansion will become unreliable.

Appendix K

Sum rules related to the f-sum rule

Source: W. Thomas, "Über die Zahl der Dispersionselectronen, die einem stationären Zustande zugeordnet sind. Vorläufige Mitteilung", Naturwiss. 13, 627 (1925); W. Kuhn, "Über die Gesamtstr̆ke der von einem Zustande ausgehenden Absorptionslinien", Z. Phys. 33, 408-412 (1925); F. Reiche and W. Thomas, "Über die Zahl der Dispersionselektronen, die einem stationären Zustand zugeordnet sind", ibid. 34, 510-525 (1925)

K.1 General derivation by Wang

I follow here the very general derivation given by Wang[89], from which a numberof other sum rules can be derived in a straightforward manner.

For a general operator $\hat{A}$ we obtain

$$\begin{aligned}\sum_n (E_n - E_0)\left|\langle 0|\hat{A}|n\rangle\right|^2 &= \sum_n (E_n - E_0)\langle 0|\hat{A}|n\rangle\langle n|\hat{A}^\dagger|0\rangle \\ &= \sum_n \Big(-\langle 0|\hat{H}\hat{A}|n\rangle + \langle 0|\hat{A}\hat{H}|n\rangle\Big)\langle n|\hat{A}^\dagger|0\rangle \\ &= \sum_n \langle 0|[\hat{A},\hat{H}]_-|n\rangle\langle n|\hat{A}^\dagger|0\rangle \\ &= \langle 0|[\hat{A},\hat{H}]_-\hat{A}^\dagger|0\rangle \qquad \text{(K.1)}\end{aligned}$$

We proceed completely analogously but for the commutator with the second matrix element

$$\begin{aligned}\sum_n (E_n - E_0)\left|\langle 0|\hat{A}^\dagger|n\rangle\right|^2 &= \sum_n (E_n - E_0)\langle 0|\hat{A}^\dagger|n\rangle\langle n|\hat{A}|0\rangle \\ &= \sum_n \langle 0|\hat{A}^\dagger|n\rangle\Big(\langle n|\hat{H}\hat{A}|0\rangle - \langle n|\hat{A}\hat{H}|0\rangle\Big) \\ &= -\sum_n \langle 0|\hat{A}^\dagger|n\rangle\langle n|[\hat{A},\hat{H}]|0\rangle \\ &= -\langle 0|\hat{A}^\dagger[\hat{A},\hat{H}]|0\rangle \qquad \text{(K.2)}\end{aligned}$$

Adding the two results, Eq. K.1 and Eq. K.2, yields

WANG'S SUM RULE

$$\sum_n (E_n - E_0)\left(\left|\langle 0|\hat{A}|n\rangle\right|^2 + \left|\langle 0|\hat{A}^\dagger|n\rangle\right|^2\right) = \langle 0| \left[\hat{A}^\dagger, \left[\hat{A}, \hat{H}\right]_-\right]_- |0\rangle \tag{K.3}$$

This is a rather general result, from which the f-sum rule can be obtained by choosing $\hat{A} = \hat{x}$ and the one-dimensional hamilton operator of the form $\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x})$.

K.1.1 f-sum rule or Thomas-Reiche-Kuhn sum rule

The **f-sum rule** or **Thomas-Reiche-Kuhn sum rule** can be obtained from Eq. K.3 with $\hat{A} = \hat{\vec{r}}$ and the one-dimensional hamilton operator of the form $\hat{H} = \frac{\hat{\vec{p}}^2}{2m} + V(\hat{\vec{r}})$. Here we first calculate the one-dimensional case from which the three-dimensional can be obtained immediately.

Hence, we use $\hat{A} = \hat{x}$ and the one-dimensional hamilton operator of the form $\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x})$.

$$\begin{aligned}
\left[\hat{H}, \hat{A}\right]_- &= \left[\frac{\hat{p}^2}{2m} + V(\hat{x}), \vec{x}\right]_- \\
&= \frac{1}{2m}\left(\hat{p}^2\vec{x} \underbrace{-\hat{p}\vec{x}\hat{p} + \hat{p}\vec{x}\hat{p}}_{=0} - \vec{x}\hat{p}^2\right) \\
&= \frac{1}{2m}\left(\hat{p}\,[\hat{p}, \vec{x}] + [\hat{p}, \vec{x}]_-\,\hat{p}\right) \\
&= \frac{1}{2m}\frac{\hbar}{i}\left(\hat{p} + \hat{p}\right) \\
&= \frac{-i\hbar}{m}\hat{p}
\end{aligned} \tag{K.4}$$

$$\begin{aligned}
\left[\hat{A}^\dagger, \left[\hat{H}, \hat{A}\right]_-\right]_- &= \left[\hat{x}^\dagger, \frac{-i\hbar}{m}\hat{p}\right]_- \\
&= \frac{-i\hbar}{m}\left[\hat{x}^\dagger, \hat{p}\right]_- \\
&= \frac{\hbar^2}{m}
\end{aligned} \tag{K.5}$$

Thus we obtain

$$\sum_n (E_n - E_0)\,|\langle 0|\hat{x}|n\rangle|^2 = \frac{\hbar^2}{2m} \tag{K.6}$$

which is the one-dimensional version of the f-sum rule.

Now we simply sum over the three coordinates and obtain

F-SUM RULE OR THOMAS-REICHE-KUHN SUM RULE

$$\sum_n (E_n - E_0)\,|\langle 0|\vec{r}|n\rangle|^2 = \frac{3\hbar^2}{2m} \tag{K.7}$$

K.1.2 Bethe sum rule

The **Bethe sum rule** can be obtained from Eq. K.3 with $\hat{A} = \mathrm{e}^{i\vec{q}\vec{r}}$ and the Hamiltonian $\hat{H} = \frac{\hat{\vec{p}}^2}{2m} + V(\hat{\vec{r}})$.

$$\begin{aligned}
\left[\hat{H}, \hat{A}\right]_- &= \left[\frac{\hat{\vec{p}}^2}{2m} + V(\hat{\vec{r}}), \mathrm{e}^{i\vec{q}\vec{r}}\right]_- \\
&= \frac{1}{2m}\left(\hat{\vec{p}}^2\mathrm{e}^{i\vec{q}\vec{r}} \underbrace{-\hat{\vec{p}}\mathrm{e}^{i\vec{q}\vec{r}}\hat{\vec{p}} + \hat{\vec{p}}\mathrm{e}^{i\vec{q}\vec{r}}\hat{\vec{p}}}_{=0} - \mathrm{e}^{i\vec{q}\vec{r}}\hat{\vec{p}}^2\right) \\
&= \frac{1}{2m}\left(\hat{\vec{p}}\left[\hat{\vec{p}}, \mathrm{e}^{i\vec{q}\vec{r}}\right]_- + \left[\hat{\vec{p}}, \mathrm{e}^{i\vec{q}\vec{r}}\right]_- \hat{\vec{p}}\right) \\
&= \frac{1}{2m}\left(\hat{\vec{p}}\hbar\vec{k}\mathrm{e}^{i\vec{q}\vec{r}} + \hbar\vec{k}\mathrm{e}^{i\vec{q}\vec{r}}\hat{\vec{p}}\right) \\
&= \frac{\hbar\vec{k}}{2m}\left(\hat{\vec{p}}\mathrm{e}^{i\vec{q}\vec{r}} + \mathrm{e}^{i\vec{q}\vec{r}}\hat{\vec{p}}\right) \qquad \text{(K.8)}
\end{aligned}$$

$$\begin{aligned}
\left[\hat{A}^\dagger, \left[\hat{H}, \hat{A}\right]_-\right]_- &= \left[\mathrm{e}^{-i\vec{q}\vec{r}}, \frac{\hbar\vec{k}}{2m}\left(\hat{\vec{p}}\mathrm{e}^{i\vec{q}\vec{r}} + \mathrm{e}^{i\vec{q}\vec{r}}\hat{\vec{p}}\right)\right]_- \\
&= \frac{\hbar\vec{k}}{2m}\left(\mathrm{e}^{-i\vec{q}\vec{r}}\hat{\vec{p}}\mathrm{e}^{i\vec{q}\vec{r}} + \hat{\vec{p}} - \hat{\vec{p}} - \mathrm{e}^{i\vec{q}\vec{r}}\hat{\vec{p}}\mathrm{e}^{-i\vec{q}\vec{r}}\right) \\
&= \frac{\hbar\vec{q}}{2m}\Big((\hbar\vec{q} + \hat{\vec{p}}) - (-\hbar\vec{q}) + \hat{\vec{p}} + \hat{\vec{p}})\Big) \\
&= \frac{(\hbar\vec{q})^2}{m} \qquad \text{(K.9)}
\end{aligned}$$

BETHE SUM RULE

$$\sum_n (E_n - E_0)\left|\langle 0|\mathrm{e}^{i\vec{q}\vec{r}}|n\rangle\right|^2 = \frac{(\hbar\vec{q})^2}{2m} \qquad \text{(K.10)}$$

Appendix L

Time-inversion symmetry

Time inversion symmetry says that it is not possible from a conservative classical trajectory to find out if it time is running forward or backward in time. If we take a physical trajectory $\vec{x}(t)$ and let the time run backwards, i.e. $\vec{x'}(t) = \vec{x}(-t)$, the new trajectory $\vec{x'}(t)$ still fulfills the equations of motion.

Electrodynamics and gravitation obey time-inversion symmetry exactly. Howver, while time time-inversion symmetry is one of the most fundamental properties of natural laws, it is, taken alone, not a fundamental symmetry of nature: The weak interaction, which is for example responsible for the β decay of nuclei, violates it. Time inversion must be replaced by the weaker CPT-inversion symmetry. This is the so-called **CPT-theorem** posed by Gerhart Lüders and Wolfgang Pauli. The CPT theorem says that the fundamental laws of nature must obey a symmetry under simultaneous application of three operations:

- charge inversion (C)
- space inversion (P for Parity)
- time inversion (T)

The CPT theorem is based on the assumptions of Lorentz invariance, causality, locality and the existence of a Hamilton operator that is bounded by below. Electrodynamics and gravitation are symmetric under the three symmetry operations individually.

As we are not concerned with weak interactions we can assume exact time inversion symmetry.

L.1 Schrödinger equation

Let us now investigate what time inversion symmetry implies in quantum mechanics:

Let us consider the Schrödinger equation in a magnetic field

$$i\hbar\partial_t\Psi(\vec{r},t) = \left[\frac{(\frac{\hbar}{i}\vec{\nabla} - q\vec{A})^2}{2m} + q\Phi\right]\Psi(\vec{r},t) \tag{L.1}$$

Let us take the complex conjugate of the Eq. L.1

$$-i\hbar\partial_t\Psi^*(\vec{r},t) = \left[\frac{(-\frac{\hbar}{i}\vec{\nabla} - q\vec{A})^2}{2m} + q\Phi\right]\Psi^*(\vec{r},t) = \left[\frac{(\frac{\hbar}{i}\vec{\nabla} + q\vec{A})^2}{2m} + q\Phi\right]\Psi^*(\vec{r},t) \tag{L.2}$$

Next we look for the equation obeyed by $\psi(\vec{r},-t)$, if Eq. L.2 holds

$$i\hbar\partial_t\Psi^*(\vec{r},-t) = \left[\frac{(\frac{\hbar}{i}\vec{\nabla} + q\vec{A}(\vec{r},-t))^2}{2m} + q\Phi(\vec{r},-t)\right]\Psi^*(\vec{r},-t) \tag{L.3}$$

One can see immediately that Eq. L.3 is identical to Eq. L.1 if we revert the sign of the vector potential $\vec{A}$, when we invert the time. Thus, the Schrödinger equation is symmetric under the time inversion symmetry as stated below:

TIME-INVERSION

$$\begin{aligned}\vec{A}(\vec{r},t) &\rightarrow -\vec{A}(\vec{r},-t)\\ \Phi(\vec{r},t) &\rightarrow \Phi(\vec{r},-t)\\ \Psi(\vec{r},t) &\rightarrow \Psi^*(\vec{r},-t)\end{aligned}$$

For the time-independent Schrödinger equation we obtain

$$\Psi(\vec{r},t) = \Psi_\epsilon(\vec{r})\mathrm{e}^{-\frac{i}{\hbar}\epsilon t}$$

so that

$$\Psi^*(\vec{r},-t) = \Psi^*_\epsilon(\vec{r})\mathrm{e}^{-\frac{i}{\hbar}\epsilon t}$$

Thus, the time inversion symmetry applied to energy eigenstates has the effect that the wave function is turned into its complex conjugate.

Let us look at the problem from stationary Schrödinger equation.

$$\left[\frac{(\frac{\hbar}{i}\vec{\nabla} - q\vec{A})^2}{2m} + q\Phi - \epsilon\right]\Psi_\epsilon(\vec{r}) = 0$$

We take the complex conjugate of this equation

$$\left[\frac{(\frac{\hbar}{i}\vec{\nabla} + q\vec{A})^2}{2m} + q\Phi - \epsilon\right]\Psi^*_\epsilon(\vec{r}) = 0$$

Thus, we see that the complex conjugate of the wave function solves the same Schrödinger equation with the magnetic field reversed. If there is no magnetic field, the complex conjugate is also a solution of the original wave function. Thus, we can superimpose the two solutions to obtain two real solutions, namely the real part and the imaginary part. Thus, in the absence of a magnetic field the wave functions can be assumed to be purely real.

L.2 Pauli equation

The proper theory of electrons is the **Dirac equation**, which describes electrons by a four component spinor, that describes spin up and spin-down electrons as well as their antiparticles, the positrons. In the non-relativistic limit electrons and positrons become independent. Now electrons and positrons obey the so-called Pauli equation.

The **Pauli equation** has the form

$$i\hbar\partial_t|\psi(t)\rangle = \left[\frac{(\vec{p} - q\vec{A})^2}{2m_e} + q\Phi - \frac{q}{m_e}\hat{\vec{S}}\vec{B}\right]|\psi\rangle \qquad \text{(L.4)}$$

The wave function is now a two-component spinor with a spin-up and a spin-down component. The spin-operator is represented by the Pauli matrices

$$\vec{S} \overset{\text{Eq. 3.1}}{=} \frac{\hbar}{2}\left(\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\right)$$

The Pauli matrices are given in Eq. 3.2 on 46. The magnetic field is related to the vector potential via $\vec{B} = \vec{\nabla} \times \vec{A}$.

Expressed with explicit spinor components, the Pauli equation has the form

$$i\hbar\partial_t\psi(\vec{r}, \sigma, t) = \sum_{\sigma'} \left[\underbrace{\left(\frac{(\frac{\hbar}{i}\vec{\nabla} - q\vec{A})^2}{2m_e} + q\Phi \right) \delta_{\sigma,\sigma'}}_{\hat{H}_0} - \frac{q}{m_e}\vec{B}\vec{S}_{\sigma,\sigma'} \right] \psi(\vec{r}, \sigma', t) \tag{L.5}$$

We proceed as we did for the Schrödinger equation by taking the complex conjugate of the Pauli equation Eq. L.5

$$-i\hbar\partial_t\psi^*(\vec{r}, \sigma, t) = \sum_{\sigma} \left[\left(\frac{(-\frac{\hbar}{i}\vec{\nabla} - q\vec{A})^2}{2m_e} + q\Phi \right) \delta_{\sigma,\sigma'} - \frac{q}{m_e}\vec{B}\vec{S}^*_{\sigma,\sigma'} \right] \psi^*(\vec{r}, \sigma', t) \tag{L.6}$$

Now we revert the time argument[1] in Eq. L.6 and perform the transformation of the potentials $\vec{A}'(\vec{r}, t) = -\vec{A}(\vec{r}, -t)$ and $\Phi'(\vec{r}, t) = \Phi(\vec{r}, -t)$. The transformed magnetic field is $\vec{B}'(\vec{r}, t) = \vec{\nabla} \times \vec{A}'(\vec{r}, t) = -\vec{\nabla} \times \vec{A}(\vec{r}, -t) = -\vec{B}(\vec{r}, -t)$.

$$i\hbar\partial_t\psi^*(\vec{r}, \sigma, -t) = \sum_{\sigma'} \Bigg[\underbrace{\left(\frac{(\frac{\hbar}{i}\vec{\nabla} - q\vec{A}'(\vec{r}, t))^2}{2m_e} + q\Phi'(\vec{r}, t) \right) \delta_{\sigma,\sigma'}}_{\hat{H}'_0} + \frac{q}{m_e}\vec{B}'(\vec{r}, t)\vec{S}^*_{\sigma,\sigma'} \Bigg] \psi^*(\vec{r}, \sigma', -t) \tag{L.7}$$

We observe, that it is no more sufficient to replace the wave function by its complex conjugate of the time-reverted function as in the Schrödinger equation. This is because the complex conjugate of the spin operator $\vec{S}_{\sigma,\sigma'}$ is not identical to its complex conjugate. The y-component is purely imaginary, because the corresponding Pauli matrix is purely imaginary. Thus, the y-component changes its sign, when the complex conjugate is taken.

In order to find the transformation of the wave functions, let us rewrite the equation in components. The original equation Eq. L.5 written in components looks like

$$(i\hbar\partial_t - \hat{H}_0) \begin{pmatrix} \psi(\vec{r}, \uparrow, t) \\ \psi(\vec{r}, \downarrow, t) \end{pmatrix} \overset{\text{Eq. L.5}}{=} -\frac{\hbar q}{2m_e} \begin{pmatrix} B_z & B_x - iB_y \\ B_x + iB_y & -B_z \end{pmatrix} \begin{pmatrix} \psi(\vec{r}, \uparrow, t) \\ \psi(\vec{r}, \downarrow, t) \end{pmatrix} \tag{L.8}$$

The last equation, Eq. L.7, written in component notation has the form

$$(i\hbar\partial_t - \hat{H}'_0) \begin{pmatrix} \psi^*(\vec{r}, \uparrow, -t) \\ \psi^*(\vec{r}, \downarrow, -t) \end{pmatrix} \overset{\text{Eq. L.7}}{=} \underbrace{-\frac{\hbar q}{2m_e} \begin{pmatrix} -B'_z & -B'_x - iB'_y \\ -B'_x + iB'_y & B'_z \end{pmatrix}}_{+\frac{q}{m_e}\vec{B}'\vec{S}^*} \begin{pmatrix} \psi^*(\vec{r}, \uparrow, -t) \\ \psi^*(\vec{r}, \downarrow, -t) \end{pmatrix} \tag{L.9}$$

In order to get an idea on how to proceed, let us consider the special case $B'_x = B'_y = 0$. In this case the diagonal elements of the equation could be brought into the form of Eq. L.8 by interchanging the spin indices of the wave functions.

Guided by this idea, we interchange the spin indices of Eq. L.9. This interchanges the columns and the rows of the 2×2 matrix. We obtain

$$\left(i\hbar\partial_t - \hat{H}'\right) \begin{pmatrix} \psi^*(\vec{r}, \downarrow, -t) \\ \psi^*(\vec{r}, \uparrow, -t) \end{pmatrix} = -\frac{\hbar q}{2m_e} \begin{pmatrix} B'_z & -B'_x + iB'_y \\ -B'_x - iB'_y & -B'_z \end{pmatrix} \begin{pmatrix} \psi^*(\vec{r}, \downarrow, -t) \\ \psi^*(\vec{r}, \uparrow, -t) \end{pmatrix}$$

---

[1] Consider a replacement $t = -t'$. Thus, $\partial_{t'} = -\partial_t$ and $\psi(t) = \psi(-t')$. After the transformation we drop the prime.

The off-diagonal elements of the equation still differ by the sign from those in Eq. L.8. This problem can be remedied by replacing the upper spinor component by its negative. While rewriting the equation for the new convention, the sign of the off-diagonal elements of the matrix is changed.

$$\left(i\hbar\partial_t - \hat{H}'\right)\begin{pmatrix} -\psi^*(\vec{r},\downarrow,-t) \\ \psi^*(\vec{r},\uparrow,-t)\end{pmatrix} = -\frac{\hbar q}{2m_e}\begin{pmatrix} B'_z & B'_x - iB'_y \\ B'_x + iB'_y & -B'_z\end{pmatrix}\begin{pmatrix} -\psi^*(\vec{r},\downarrow,-t) \\ \psi^*(\vec{r},\uparrow,-t)\end{pmatrix} \tag{L.10}$$

Thus, we arrived at the desired form Eq. L.8.

By comparing Eq. L.10 with the original Pauli equation Eq. L.8, we see that the for any solution of the Pauli equation also the result of the time inversion operation Eq. L.11 for two-component spinors and electromagnetic fields is a solution of the same Pauli equation.

TIME-INVERSION FOR TWO-COMPONENT SPINORS AND ELECTROMAGNETIC FIELDS

$$\begin{aligned}\psi'(\vec{r},\uparrow,t) &= -\psi^*(\vec{r},\downarrow,-t) \\ \psi'(\vec{r},\downarrow,t) &= +\psi^*(\vec{r},\uparrow,-t) \\ \vec{A}'(\vec{r},t) &= -\vec{A}(\vec{r},-t) \\ \Phi'(\vec{r},t) &= \Phi(\vec{r},-t)\end{aligned} \tag{L.11}$$

If we investigate the resulting transformation of the spin-expectation values, we see that the spin is inverted. This is expected under time-inversion symmetry, if we interpret the spin as a angular momentum. If we invert the time, the particle is spinning in the opposite direction.

L.3 Time inversion for Bloch states

Often we require the implications of this symmetry for wave in a representation of Bloch waves. A Bloch state is given by as product of a periodic function $u_{\vec{k},n}$ and a phase factor $e^{i\vec{k}\vec{r}}$, so that

$$\psi_{\vec{k},n}(\vec{r},\sigma,t) = u_{\vec{k},n}(\vec{r},\sigma)e^{i(\vec{k}\vec{r}-\vec{\omega}_n t)} \tag{L.12}$$

If time inversion symmetry is obeyed, for example in the absence of magnetic fields with static potentials, we can use $\psi(\vec{r},\uparrow,t) = -\psi^*(\vec{r},\downarrow,-t)$ and $\psi(\vec{r},\downarrow,t) = -\psi^*(\vec{r},\uparrow,-t)$, so that

$$\begin{aligned}u_{\vec{k},n}(\vec{r},\uparrow)e^{i(\vec{k}\vec{r}-\vec{\omega}t)} = \psi(\vec{r},\uparrow,t) = -\psi^*(\vec{r},\downarrow,-t) &= -u^*_{-\vec{k},n}(\vec{r},\downarrow)e^{i(\vec{k}\vec{r}-\vec{\omega}t)} \\ u_{\vec{k},n}(\vec{r},\uparrow) &= -u^*_{-\vec{k},n}(\vec{r},\downarrow)\end{aligned} \tag{L.13}$$

and

$$\begin{aligned}u_{\vec{k},n}(\vec{r},\downarrow)e^{i(\vec{k}\vec{r}-\vec{\omega}t)} = \psi(\vec{r},\downarrow,t) = \psi^*(\vec{r},\uparrow,-t) &= u^*_{-\vec{k},n}(\vec{r},\uparrow)e^{i(\vec{k}\vec{r}-\vec{\omega}t)} \\ u_{\vec{k},n}(\vec{r},\downarrow) &= u^*_{-\vec{k},n}(\vec{r},\uparrow)\end{aligned} \tag{L.14}$$

Thus, we obtain for the periodic parts of the wave functions

$$\begin{pmatrix} u_{\vec{k},n}(\vec{r},\uparrow) \\ u_{\vec{k},n}(\vec{r},\downarrow)\end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0\end{pmatrix}\begin{pmatrix} u_{-\vec{k},n}(\vec{r},\uparrow) \\ u_{-\vec{k},n}(\vec{r},\downarrow)\end{pmatrix}^* \tag{L.15}$$

Appendix M

Slater determinants for parallel and antiparallel spins

Why can spin-up and spin-down electrons be treated as non-identical particle, even through they are only spinor-components of identical particles.

Here, we try to give an answer by showing that a one-particle orbitals can be occupied independently of each other with one spin-up and one spin-down electron. That is for each placement of the two electrons into one-particle orbitals there is a Slater determinant. For two spin-up electrons or for two spin-down electrons, the same one-particle orbital can only be occupied once.

This implies that the statistics of spin up and spin-down electrons is such that spin up electrons are treated as a class of identical particles and spin-down electrons are treated as a different class of independent particles.

Let us consider two spatial one-particle orbitals $\chi_a(\vec{r})$ and $\chi_b(\vec{r})$. The two orbitals shall be orthonormal. Out of these two orbitals, we construct four spin orbitals, namely

$$\begin{aligned}
\phi_{a\uparrow}(\vec{r},\sigma) &= \chi_a(\vec{r})\delta_{\sigma,\uparrow} \\
\phi_{a\downarrow}(\vec{r},\sigma) &= \chi_a(\vec{r})\delta_{\sigma,\downarrow} \\
\phi_{b\uparrow}(\vec{r},\sigma) &= \chi_b(\vec{r})\delta_{\sigma,\uparrow} \\
\phi_{b\downarrow}(\vec{r},\sigma) &= \chi_b(\vec{r})\delta_{\sigma,\downarrow}
\end{aligned}$$

or using $\alpha \in \{a, b\}$ and $s \in \{\uparrow, \downarrow\}$.

$$\phi_{\alpha,s}(\vec{r},\sigma) = \chi_\alpha(\vec{r})\delta_{s,\sigma}$$

Out of these one-particle spin orbitals, we construct the Slater determinants

$$\begin{aligned}
\Phi_{\alpha,\beta,s,s'}(\vec{r},\sigma,\vec{r'},\sigma') &= \frac{1}{\sqrt{2}}\Big(\phi_{\alpha,s}(\vec{r},\sigma)\phi_{\beta,s'}(\vec{r'},\sigma') - \phi_{\beta,s'}(\vec{r},\sigma)\phi_{\alpha,s}(\vec{r'},\sigma')\Big) \\
&= \frac{1}{\sqrt{2}}\Big(\chi_\alpha(\vec{r})\chi_\beta(\vec{r'})\delta_{s,\sigma}\delta_{s',\sigma'} - \chi_\beta(\vec{r})\chi_\alpha(\vec{r'})\delta_{s,\sigma'}\delta_{s',\sigma}\Big)
\end{aligned}$$

Some of these Slater determinants are zero states, but that does not matter at the moment.

These Slater determinants are products of a spatial two-particle wave function and a two-particle wave function in the spin space. It will be convenient to introduce such two-particle states that are symmetric and antisymmetric under particle exchange. (Exchange of the arguments.) These are

$$
\begin{aligned}
\Psi^{+}_{\alpha,\beta}(\vec{r},\vec{r'}) &\stackrel{\text{def}}{=} \frac{1}{\sqrt{2}}\Big(\chi_\alpha(\vec{r})\chi_\beta(\vec{r'})+\chi_\beta(\vec{r})\chi_\alpha(\vec{r'})\Big)\\
\Psi^{-}_{\alpha,\beta}(\vec{r},\vec{r'}) &\stackrel{\text{def}}{=} \frac{1}{\sqrt{2}}\Big(\chi_\alpha(\vec{r})\chi_\beta(\vec{r'})-\chi_\beta(\vec{r})\chi_\alpha(\vec{r'})\Big)\\
G^{+}_{s,s'}(\sigma,\sigma') &\stackrel{\text{def}}{=} \frac{1}{\sqrt{2}}\Big(\delta_{s,\sigma}\delta_{s',\sigma'}+\delta_{s',\sigma'}\delta_{s,\sigma}\Big)\\
G^{-}_{s,s'}(\sigma,\sigma') &\stackrel{\text{def}}{=} \frac{1}{\sqrt{2}}\Big(\delta_{s,\sigma}\delta_{s',\sigma'}-\delta_{s',\sigma'}\delta_{s,\sigma}\Big)
\end{aligned}
$$

The Slater determinants have the form

$$
\begin{aligned}
\Phi_{\alpha,\beta,s,s'}(\vec{r},\sigma,\vec{r'},\sigma') &= \frac{1}{\sqrt{2}}\left[\frac{1}{2}\Big(\Psi^{+}_{\alpha,\beta}+\Psi^{-}_{\alpha,\beta}\Big)\Big(G^{+}_{s,s'}+G^{-}_{s,s'}\Big)-\frac{1}{2}\Big(\Psi^{+}_{\alpha,\beta}-\Psi^{-}_{\alpha,\beta}\Big)\Big(G^{+}_{s,s'}-G^{-}_{s,s'}\Big)\right]\\
&= \frac{1}{\sqrt{2}}\Big[\Psi^{+}_{\alpha,\beta}(\vec{r},\vec{r'})G^{-}_{s,s'}(\sigma,\sigma')+\Psi^{-}_{\alpha,\beta}(\vec{r},\vec{r'})G^{+}_{s,s'}(\sigma,\sigma')\Big] \qquad \text{(M.1)}
\end{aligned}
$$

Now we can analyze the result using

$$
\begin{aligned}
G^{-}_{s,s'}(\sigma,\sigma') &\equiv 0 \qquad \text{for} \qquad s=s'\\
\Psi^{-}_{\alpha,\beta}(\vec{r},\vec{r'}) &\equiv 0 \qquad \text{for} \qquad \alpha=\beta
\end{aligned}
$$

Let us now consider two spin up electrons. Because of Eq. M.1 and $G^{-}_{\uparrow,\uparrow}=0$, there is only a single non-zero Slater determinant, namely

$$
|\Phi_{a,b,\uparrow,\uparrow}\rangle = \Psi^{-}_{a,b}\left(\frac{1}{\sqrt{2}}G^{+}_{s,s'}\right)
$$

The Slater determinant $|\Phi_{b,a,\uparrow,\uparrow}\rangle$ differs from $|\Phi_{a,b,\uparrow,\uparrow}\rangle$ only by a sign change. Thus, we see that the spatial wave function is antisymmetric. Hence, the Pauli principle exists in the spatial coordinates.

To show the difference let us investigate the states with different spin.

$$
\begin{aligned}
|\Phi_{a,a,\uparrow,\downarrow}\rangle &= \Big(\frac{1}{\sqrt{2}}\Psi^{+}_{a,a}\Big)G^{-}_{\uparrow,\downarrow}\\
|\Phi_{a,b,\uparrow,\downarrow}\rangle &= \frac{1}{\sqrt{2}}\Big(\Psi^{+}_{a,b}G^{-}_{\uparrow,\downarrow}+\Psi^{-}_{a,b}G^{+}_{\uparrow,\downarrow}\Big)\\
|\Phi_{b,a,\uparrow,\downarrow}\rangle &= \frac{1}{\sqrt{2}}\Big(\Psi^{+}_{a,b}G^{-}_{\uparrow,\downarrow}-\Psi^{-}_{a,b}G^{+}_{\uparrow,\downarrow}\Big)\\
|\Phi_{b,b,\uparrow,\downarrow}\rangle &= \Big(\frac{1}{\sqrt{2}}\Psi^{+}_{b,b}\Big)G^{-}_{\uparrow,\downarrow}
\end{aligned}
$$

We see that the two electrons with different spin can be placed without restriction into the two spatial one-particle orbitals.

This indicates that electrons that the statistics of electrons with different spin is identical to that of two non-identical particles. The statistics of electrons with like spin is like identical particles.

M.1 Spatial symmetry for parallel and antiparallel spins

Here, we show that

- the spatial wave functions for two electrons with parallel spin is antisymmetric
- the spatial wave functions for two electrons with anti-parallel spin is symmetric

under exchange of the coordinates.

A common misconception is that the Slater determinant of two one-particle spin-orbitals with opposite spin describes two electrons with anti-parallel spin. For such a state the z-component of the spin vanishes, but it is a superposition of a two states with $S_z = 0$, namely one with antiparallel spin and the other with parallel spin with a total spin lying in the xy-plane.

In order to determine states with parallel and antiparallel spin, we need to determine the spin eigenstates and eigenvalues. If the eigenvalue for $\hat{S}^2_{tot} = (\hat{\vec{S}}_1 + \hat{\vec{S}}_2)^2$ vanishes, the wave function has antiparallel spin. If the eigenvalue is $2\hbar^2$ the spins are parallel.

Using the ladder operators $\hat{S}_\pm$ we write the total spin as

$$
\begin{aligned}
\hat{S}^2_{tot} &= (\hat{\vec{S}}_1 + \hat{\vec{S}}_2)^2 \\
&= \hat{\vec{S}}^2_1 + \hat{\vec{S}}^2_2 + 2\hat{\vec{S}}_1\hat{\vec{S}}_2 \\
&= \hat{\vec{S}}^2_1 + \hat{\vec{S}}^2_2 + 2\Big[\frac{1}{2}\hat{S}_{1,+}\hat{S}_{2,-} + \frac{1}{2}\hat{S}_{1,-}\hat{S}_{2,+} + \hat{S}_{1,z}\hat{S}_{2,z}\Big] \\
&= \hat{\vec{S}}^2_1 + \hat{\vec{S}}^2_2 + \hat{S}_{1,+}\hat{S}_{2,-} + \hat{S}_{1,-}\hat{S}_{2,+} + 2\hat{S}_{1,z}\hat{S}_{2,z}\Big]
\end{aligned}
$$

The ladder operators are defined as

$$
\begin{aligned}
\hat{S}_+ &= \hat{S}_x + i\hat{S}_y \\
\hat{S}_- &= \hat{S}_x - i\hat{S}_y
\end{aligned}
$$

and obey

$$
\begin{aligned}
&\hat{S}_-|\uparrow\rangle = |\downarrow\rangle\hbar \qquad \text{and} \qquad \hat{S}_-|\downarrow\rangle = |\varnothing\rangle \\
&\hat{S}_+|\uparrow\rangle = |\varnothing\rangle \qquad \text{and} \qquad \hat{S}_+|\downarrow\rangle = |\uparrow\rangle\hbar
\end{aligned}
$$

Thus

$$
\begin{aligned}
\Big(\hat{S}_{1+}\hat{S}_{2-} + \hat{S}_{1-}\hat{S}_{2+}\Big)|G^\pm_{\uparrow,\downarrow}\rangle &= \Big(\hat{S}_{1+}\hat{S}_{2-} + \hat{S}_{1-}\hat{S}_{2+}\Big)\frac{1}{\sqrt{2}}\Big(|\uparrow,\downarrow\rangle \pm |\downarrow,\uparrow\rangle\Big) \\
&= \frac{1}{\sqrt{2}}\Big(|\downarrow,\uparrow\rangle \pm |\uparrow,\downarrow\rangle\Big)\hbar^2 \\
&= |G^\pm_{\uparrow,\downarrow}\rangle\Big(\pm\hbar^2\Big) \\
\Big(\hat{S}_{1+}\hat{S}_{2-} + \hat{S}_{1-}\hat{S}_{2+}\Big)|G^\pm_{\uparrow,\uparrow}\rangle &= |\varnothing\rangle \\
\Big(\hat{S}_{1+}\hat{S}_{2-} + \hat{S}_{1-}\hat{S}_{2+}\Big)|G^\pm_{\downarrow,\downarrow}\rangle &= |\varnothing\rangle \\
\hat{S}_{1,z}\hat{S}_{2,z}|G^\pm_{\uparrow,\downarrow}\rangle &= |G^\pm_{\uparrow,\downarrow}\rangle\Big(-\frac{\hbar^2}{4}) \\
\hat{S}_{1,z}\hat{S}_{2,z}|G^\pm_{\uparrow,\uparrow}\rangle &= |G^\pm_{\uparrow,\uparrow}\rangle\Big(+\frac{\hbar^2}{4}) \\
\hat{S}_{1,z}\hat{S}_{2,z}|G^\pm_{\downarrow,\downarrow}\rangle &= |G^\pm_{\downarrow,\downarrow}\rangle\Big(+\frac{\hbar^2}{4}) \\
\hat{\vec{S}}^2_i|G^\pm_{s,s'}\rangle &= |G^\pm_{s,s'}\rangle\frac{3\hbar^2}{4}
\end{aligned}
$$

$$
\begin{aligned}
\hat{S}^2_{tot}|G^\pm_{\uparrow,\downarrow}\rangle &= |G^\pm_{\uparrow,\downarrow}\rangle\left[\frac{3\hbar^2}{4} + \frac{3\hbar^2}{4} \pm \hbar^2 - \frac{\hbar^2}{2}\right] = |G^\pm_{\uparrow,\downarrow}\rangle(1 \pm 1)\hbar^2 \\
\hat{S}^2_{tot}|G^\pm_{\uparrow,\uparrow}\rangle &= |G^\pm_{\uparrow,\uparrow}\rangle\left[\frac{3\hbar^2}{4} + \frac{3\hbar^2}{4} + \frac{\hbar^2}{2}\right] = |G^\pm_{\uparrow,\uparrow}\rangle 2\hbar^2 \\
\hat{S}^2_{tot}|G^\pm_{\downarrow,\downarrow}\rangle &= |G^\pm_{\downarrow,\downarrow}\rangle\left[\frac{3\hbar^2}{4} + \frac{3\hbar^2}{4} + \frac{\hbar^2}{2}\right] = |G^\pm_{\downarrow,\downarrow}\rangle 2\hbar^2
\end{aligned}
$$

Thus, we see that $|G^+_{\uparrow,\downarrow}\rangle$, $|G^+_{\uparrow,\uparrow}\rangle$ and $|G^+_{\downarrow,\downarrow}\rangle$ describe two electrons with parallel spin. Note that $|G^-_{s,s'}\rangle$ vanishes for parallel spin.

The only solution with antiparallel spin is $|G^-_{\uparrow,\downarrow}\rangle$.

From Eq. M.1 we know that the antisymmetric spin wave function is always combined with the symmetric spatial orbital and vice versa. Hence, the wave functions describing two electrons with anti-parallel spin are symmetric in their spatial coordinates, while the ones with parallel spin have an antisymmetric spatial wave function.

M.2 An intuitive analogy for particle with spin

Consider balls that are painted on the one side green and on the other side red.

If we place two such balls on a table with the green side up, turn around and look at them again, we cannot tell if the two balls habe been interchanged.

No we place them with opposite colors up. If we only allow that the positions of the two balls are interchanged, but exclude that they are turned around, we can tell the two spheres apart from their orientation. Hence, we can treat them as non-identical spheres.

However, if we consider exchanges of position *and* orientation, we are again unable to tell, whether the spheres have been interchanged or not.

Thus, if we can exclude that the spins of the particles –or the orientation of our spheres– change, we can divide electrons into two classes of particles, namely spin-up and spin-down electrons. Particles within each class are indistinguishable, but spin-up and spin-down electrons can be distinguished. When the spin is preserved, the Hamiltonian is invariant with respect to spin-rotation, and spin is a good quantum number.

However, if there is a magnetic field, which can cause a rotation of the spin direction, the division into spin-up and spin-down particles is no more a useful concept. In our model, if the orientation of the spheres can change with time, we cannot use it as distinguishing feature.

Appendix N

Hartree-Fock of the free-electron gas

Here, we derive the changes in the dispersion relation $\epsilon(\vec{k})$ of the free-electron gas due to the exchange potential discussed in section 6.6 on p. 110.

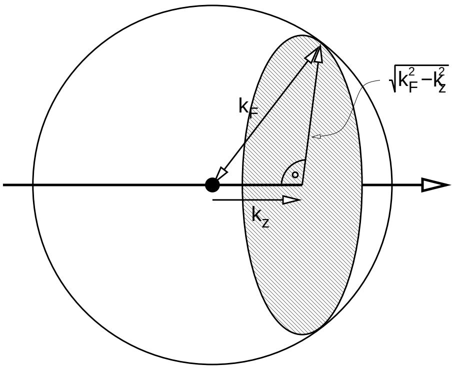
Because of the translational symmetry, we can assume that the charge density is spatially constant.[1] Since the problem is translationally invariant, we can furthermore deduce that the eigenstates are plane waves.

N.1 Exchange potential as non-local potential

Let us evaluate the non-local potential as defined in Eq. 6.24 using plane waves $\phi_{\vec{k},\sigma}(\vec{r},\sigma') = \langle\vec{r},\sigma'|\phi_{\vec{k},\sigma}\rangle = \frac{1}{\Omega}e^{i\vec{k}\vec{r}}\delta_{\sigma,\sigma'}$ as defined in Eq. 4.91 as basis functions.

$$
\begin{aligned}
V_x(\vec{x},\vec{x'}) &\overset{Eq.\,6.24}{=} -\left[\sum_j \frac{e^2\phi_j^*(\vec{x})\phi_j(\vec{x'})}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\right] \\
&\overset{Eq.\,4.91}{=} \frac{-e^2}{4\pi\epsilon_0\Omega}\left[\sum_{\vec{k}}\theta(k_F-|\vec{k}|)\overbrace{\underbrace{\frac{(2\pi)^3}{\Omega}}_{\to d^3k}\frac{\Omega}{(2\pi)^3}}^{1}\frac{e^{i\vec{k}(\vec{r'}-\vec{r})}\delta_{\sigma,\sigma'}}{|\vec{r}-\vec{r'}|}\right] \\
&= \frac{-e^2}{4\pi\epsilon_0|\vec{r}-\vec{r'}|}\frac{\delta_{\sigma,\sigma'}}{(2\pi)^3}\int d^3k\;\theta(k_F-|\vec{k}|)e^{i\vec{k}(\vec{r'}-\vec{r})}
\end{aligned}
\tag{N.1}
$$

We see that the integral is isotropic in $\vec{r'}-\vec{r}$. Thus, we can assume, without loss of generality, that the distance vector points in z-direction, that is $\vec{r'}-\vec{r}=\vec{e}_z s$.



---

[1] We assume here that there is no symmetry breaking, which is not guaranteed.

$$
\begin{aligned}
\int_{\vec{k}\le k_F} d^3k\ \mathrm{e}^{ik_zs} &= \int_{-k_F}^{k_F} dk_z \mathrm{e}^{ik_zs} \underbrace{\int_{k_x^2+k_y^2<k_F^2-k_z^2} dk_x dk_y}_{\pi(k_F^2-k_z^2)} \\
&= \int_{-k_F}^{k_F} dk_z \left(\pi(k_F^2-k_z^2)\right) \mathrm{e}^{ik_zs} \\
&= \pi k_F^2 \int_{-k_F}^{k_F} dk_z\ \mathrm{e}^{ik_zs} - \pi \int_{-k_F}^{k_F} dk_z\ k_z^2 \mathrm{e}^{ik_zs} \\
&= \pi k_F^2 \int_{-k_F}^{k_F} dk_z\ \mathrm{e}^{ik_zs} + \pi \frac{d^2}{ds^2} \int_{-k_F}^{k_F} dk_z\ \mathrm{e}^{ik_zs} \\
&= \pi \left[k_F^2 + \frac{d^2}{ds^2}\right] \left[\frac{1}{is}\mathrm{e}^{ik_zs}\right]_{-k_F}^{k_F} \\
&= \pi \left[k_F^2 + \frac{d^2}{ds^2}\right] \frac{\mathrm{e}^{ik_Fs} - \mathrm{e}^{-ik_Fs}}{is} \\
&= 2\pi k_F^3 \left[1 + \frac{d^2}{d(k_Fs)^2}\right] \frac{\sin(k_Fs)}{k_Fs} \\
&= 2\pi k_F^3 \left[\left(1 + \frac{d^2}{dx^2}\right) \frac{\sin(x)}{x}\right]_{x=k_Fs} \\
&= 2\pi k_F^3 \left[\frac{\sin(x)}{x} - \frac{\sin(x)}{x} - \frac{2\cos(x)}{x^2} + \frac{2\sin(x)}{x^3}\right]_{x=k_Fs} \\
&= \frac{4\pi}{3} k_F^3 \left[-3\frac{\cos(k_Fs)}{(k_Fs)^2} + 3\frac{\sin(k_Fs)}{(k_Fs)^3}\right]
\end{aligned}
\tag{N.2}
$$

Now we insert the result of the integral into the expression for the nonlocal potential.

$$
\begin{aligned}
V_x(s,\sigma,\sigma') &\overset{Eqs.\ N.1,N.2}{=} \frac{-e^2\delta_{\sigma,\sigma'}}{4\pi\epsilon_0 s} \frac{1}{(2\pi)^3} \underbrace{4\pi k_F^3 \left[-\frac{\cos(k_Fs)}{(k_Fs)^2} + \frac{\sin(k_Fs)}{(k_Fs)^3}\right]}_{\int d^3k\ \theta(k_F-|\vec{k}|)\mathrm{e}^{i\vec{k}(\vec{r'}-\vec{r})}} \\
&= \frac{e^2\delta_{\sigma,\sigma'}}{4\pi\epsilon_0 s} \frac{1}{(2\pi)^3} \frac{4\pi}{3} k_F^3 \left[3\frac{(k_Fs)\cos(k_Fs) - \sin(k_Fs)}{(k_Fs)^3}\right]
\end{aligned}
\tag{N.3}
$$

The non-local exchange potential acts only between electrons of the same spin. It is important to realize that this non-local potential is not an interaction potential between two electrons, but it is a one-particle potential acting on each electron individually. Rather is is a potential that produces a large energy correction for a particle if its wave function is concentrated in one region, as compared to a very delocalized electron density. Nevertheless this potential expresses an interaction: The shape of this non-local potential depends on the state of all the other electrons.

N.2 Energy-level shifts by the exchange potential

Now we need to evaluate the expectation values of this potential in order to obtain the energy shifts:

$$
\begin{aligned}
d\epsilon_{\vec{k},\sigma} &= \frac{1}{\Omega}\int d^3r \int d^3r'\ V_x(|\vec{r}-\vec{r'}|,\sigma,\sigma')\mathrm{e}^{i\vec{k}(\vec{r}-\vec{r'})} \\
&\overset{\vec{s}\overset{\mathrm{def}}{=}\vec{r}-\vec{r'}}{=} \underbrace{\frac{1}{\Omega}\int d^3r}_{1} \int d^3s\ V_x(|\vec{s}|,\sigma,\sigma')\mathrm{e}^{i\vec{k}\vec{s}}
\end{aligned}
$$

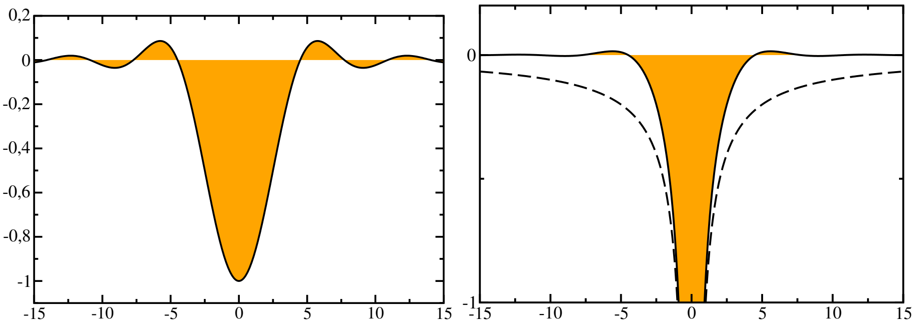


Fig. N.1: The shape of the non-local exchange potential (right) for a free electron gas as calculated in the Hartree-Fock method. The dashed line corresponds to a Coulomb interaction. The function $3\frac{x\cos(x)-sin(x)}{x^3}$ is shown on the left-hand side.

Now we decompose the plane wave into spherical harmonics(see appendix "Distributionen, d-Funktionen und Fourier transformationen in the text book of Messiah[90])

$$e^{i\vec{k}\vec{r}}=4\pi\sum_{\ell=1}^{\infty}\sum_{m=-\ell}^{\ell}i^{\ell}j_{\ell}(|\vec{k}||\vec{r}|)Y^{*}_{\ell,m}(\vec{k})Y^{*}_{\ell,m}(\vec{r})$$

Thus, we obtain

$$\begin{aligned}d\epsilon_{\vec{k},\sigma}&=\int d^3s\;V_x(|\vec{s}|,\sigma,\sigma')4\pi\sum_{\ell=1}^{\infty}\sum_{m=-\ell}^{\ell}i^{\ell}j_{\ell}(|\vec{k}||\vec{s}|)Y^{*}_{\ell,m}(\vec{k})Y^{*}_{\ell,m}(\vec{s})\\&=4\pi\sum_{\ell=1}^{\infty}\sum_{m=-\ell}^{\ell}i^{\ell}Y^{*}_{\ell,m}(\vec{k})\int d^3s\;V_x(|\vec{s}|,\sigma,\sigma')j_{\ell}(|\vec{k}||\vec{s}|)Y^{*}_{\ell,m}(\vec{s})\end{aligned}$$

Because the non-local potential is isotropic, only the term with $\ell=0$ contributes.

$$\begin{aligned}d\epsilon_{\vec{k},\sigma}&=4\pi\underbrace{Y^{*}_{0,0}(\vec{k})}_{\frac{1}{\sqrt{4\pi}}}\int d^3s\;V_x(|\vec{s}|,\sigma,\sigma)\underbrace{j_0(|\vec{k}||\vec{s}|)}_{j_0(x)=\frac{\sin(x)}{x}}\underbrace{Y^{*}_{0,0}(\vec{s})}_{\frac{1}{\sqrt{4\pi}}}\\&=\int d^3s\;V_x(|\vec{s}|,\sigma,\sigma)\frac{\sin(|\vec{k}||\vec{s}|)}{|\vec{k}||\vec{s}|}\\&\overset{s\overset{\text{def}}{=}|\vec{s}|}{=}4\pi\int ds\;s^2V_x(s,\sigma,\sigma)\frac{\sin(|\vec{k}|s)}{|\vec{k}|s}\\&=\frac{4\pi}{|\vec{k}|}\int ds\;V_x(s,\sigma,\sigma)s\sin(|\vec{k}|s)\\&\overset{Eq.\;6.27}{=}\frac{4\pi}{|\vec{k}|}\int ds\;\frac{e^2}{4\pi\epsilon_0 s}\frac{1}{(2\pi)^2}\frac{4\pi}{3}k_F^3\left[3\frac{k_Fs\cos(k_Fs)-\sin(k_Fs)}{(k_Fs)^3}\right]s\;\sin(|\vec{k}|s)\\&=\frac{4\pi}{|\vec{k}|}\frac{e^2}{4\pi\epsilon_0}\frac{1}{(2\pi)^2}\frac{4\pi}{3}k_F^3\int ds\;\left[3\frac{k_Fs\cos(k_Fs)-\sin(k_Fs)}{(k_Fs)^3}\right]\sin(|\vec{k}|s)\\&\overset{r\overset{\text{def}}{=}k_Fs}{=}\frac{4\pi}{k_F|\vec{k}|}\frac{e^2}{4\pi\epsilon_0}\frac{1}{(2\pi)^2}\frac{4\pi}{3}k_F^3\int_0^{\infty}dr\;\left[3\frac{r\cos(r)-\sin(r)}{r^3}\right]\sin(\frac{|\vec{k}|}{k_F}r)\end{aligned}$$

Now we need to solve the integral

$$I=\int_0^{\infty}dx\;\left[3\frac{x\cos(x)-\sin(x)}{x^3}\right]\sin(ax)$$

where $a = |\vec{k}|/k_F$. The difficulty with this integral is that we cannot take the two terms apart, because the individual parts of the integrand diverge at the origin. Thus, during the derivation, we have to deal with divergent expressions.

$$\begin{align}
\partial_x \frac{\sin(x)}{x^{n-1}} &= -(n-1)\frac{\sin(x)}{x^n} + \frac{\cos(x)}{x^{n-1}} \notag \\
\Rightarrow \frac{\sin(x)}{x^n} &= \frac{1}{(n-1)}\left[\frac{\cos(x)}{x^{n-1}} - \partial_x \frac{\sin(x)}{x^{n-1}}\right] \tag{N.4} \\
\overset{Eq.\ N.4}{\Rightarrow} \int_0^\infty dx \frac{\sin(x)}{x^3}\sin(ax) &= \frac{1}{2}\int_0^\infty dx\ \frac{\cos(x)}{x^2}\sin(ax) - \frac{1}{2}\int_0^\infty dx\ \sin(ax)\partial_x \frac{\sin(x)}{x^2} \notag \\
&= \frac{1}{2}\int_0^\infty dx \frac{\cos(x)}{x^2}\sin(ax) \notag \\
&\quad -\frac{1}{2}\left[\sin(ax)\frac{\sin(x)}{x^2}\right]_0^\infty + \frac{a}{2}\int_0^\infty dx\ \frac{\sin(x)}{x^2}\cos(ax) \notag \\
&= \frac{a}{2} + \frac{1}{2}\int_0^\infty dx \frac{\cos(x)\sin(ax) + a\sin(x)\cos(ax)}{x^2} \tag{N.5} \\
\overset{Eq.\ N.4}{\Rightarrow} \int_0^\infty dx \frac{\sin(x)}{x^2}\cos(ax) &= \int_0^\infty dx\ \frac{\cos(x)}{x}\cos(ax) - \int_0^\infty dx\ \cos(ax)\partial_x \frac{\sin(x)}{x} \notag \\
&= \int_0^\infty dx \frac{\cos(x)}{x}\cos(ax) \notag \\
&\quad - \left[\cos(ax)\frac{\sin(x)}{x}\right]_0^\infty - a\int_0^\infty dx\ \frac{\sin(x)}{x}\sin(ax) \notag \\
&= 1 + \int_0^\infty dx \frac{\cos(x)\cos(ax) - a\sin(x)\sin(ax)}{x} \tag{N.6}
\end{align}$$

Thus, we obtain

$$\begin{aligned}
I &= 3\int_0^\infty dx\ \frac{x\cos(x)-\sin(x)}{x^3}\sin(ax) \\
&= 3\int_0^\infty dx\ \frac{\cos(x)\sin(ax)}{x^2} - 3\int_0^\infty dx\ \frac{\sin(x)\sin(ax)}{x^3} \\
&\overset{Eq.\ N.5}{=} 3\int_0^\infty dx\ \frac{\cos(x)\sin(ax)}{x^2} - \frac{3a}{2} - \frac{3}{2}\int_0^\infty dx\ \frac{\cos(x)\sin(ax)+a\sin(x)\cos(ax)}{x^2} \\
&= -\frac{3a}{2} + \frac{3}{2}\int_0^\infty dx\ \frac{\cos(x)\sin(ax)}{x^2} - \frac{3a}{2}\int_0^\infty dx\ \frac{\sin(x)\cos(ax)}{x^2} \\
&= -\frac{3a}{2} + \frac{3a}{2}\int_0^\infty dy\ \frac{\cos(\frac{1}{a}y)\sin(y)}{y^2} - \frac{3a}{2}\int_0^\infty dx\ \frac{\sin(x)\cos(ax)}{x^2} \\
&\overset{Eq.\ N.6}{=} -\frac{3a}{2} + \frac{3a}{2}\left[1+\int_0^\infty dy\ \frac{\cos(\frac{1}{a}y)\cos(y)-\frac{1}{a}\sin(\frac{1}{a}y)\sin(y)}{y}\right] \\
&\quad -\frac{3a}{2}\left[1+\int_0^\infty dx\ \frac{\cos(ax)\cos(x)-a\sin(ax)\sin(x)}{x}\right] \\
&= -\frac{3a}{2} + \frac{3a}{2}\left[\int_0^\infty dx\ \frac{\cos(x)\cos(ax)-\frac{1}{a}\sin(x)\sin(ax)}{x}\right] \\
&\quad -\frac{3a}{2}\left[\int_0^\infty dx\ \frac{\cos(ax)\cos(x)-a\sin(ax)\sin(x)}{x}\right] \\
&= -\frac{3a}{2} + \frac{3a}{2}(-\frac{1}{a}+a)\left[\int_0^\infty dx\ \frac{\sin(x)\sin(ax)}{x}\right] \\
&= -\frac{3a}{2} + \frac{3}{2}(a^2-1)\left[\int_0^\infty dx\ \frac{\sin(x)\sin(ax)}{x}\right]
\end{aligned}$$

We recognize that the integral is the fourier transform of $\frac{\sin(x)}{x}$. From the Fourier transform tables of Bronstein[91] we take

$$\begin{aligned}
\sqrt{\frac{2}{\pi}}\int_0^\infty dx\ \sin(xy)\frac{\sin(ax)}{x} &\overset{Bronstein}{=} \frac{1}{\sqrt{2\pi}}\ln\left|\frac{y+a}{y-a}\right| \\
\overset{a\to1;y\to a}{\Rightarrow} \int_0^\infty dx\ \sin(ax)\frac{\sin(x)}{x} &= \frac{1}{2}\ln\left|\frac{a+1}{a-1}\right| = \frac{1}{2}\ln\left|\frac{1+a}{1-a}\right|
\end{aligned}$$

Thus, we obtain

$$I = -\frac{3a}{2} + \frac{3}{4}(a^2-1)\ln\left|\frac{1+a}{1-a}\right|$$

Now we are done with the Integral. We need to insert the result in the correction for the energy eigenvalues

$$\begin{aligned}
d\epsilon_{\vec{k},\sigma} &= \frac{4\pi}{k_F|\vec{k}|}\frac{e^2}{4\pi\epsilon_0}\frac{1}{(2\pi)^2}\frac{4\pi}{3}k_F^3\left[-\frac{3a}{2}+\frac{3(a^2-1)}{4}\ln\left|\frac{a+1}{a-1}\right|\right] \\
&= \frac{4\pi}{k_F^2}\frac{e^2}{4\pi\epsilon_0}\frac{1}{(2\pi)^2}\frac{4\pi}{3}k_F^3\left[-\frac{3}{2}+\frac{3(a^2-1)}{4a}\ln\left|\frac{a+1}{a-1}\right|\right] \\
&= -\frac{e^2}{4\pi\epsilon_0}\frac{2k_F}{\pi}\left[\frac{1}{2}+\frac{1-a^2}{4a}\ln\left|\frac{1+a}{1-a}\right|\right]
\end{aligned}$$

The function in parenthesis is shown in Fig. N.2 and the resulting dispersion relation is discussed in section 6.6 on p. 110.

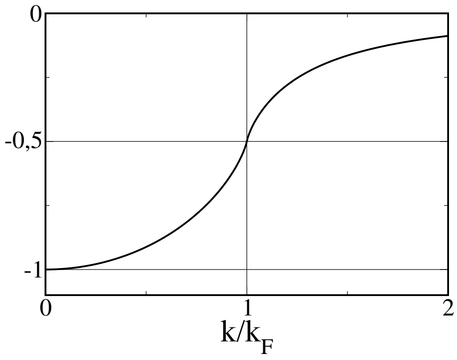


Fig. N.2: The function $f(a) = -\left(\frac{1}{2} + \frac{1-a^2}{4a}\ln\left|\frac{1+a}{1-a}\right|\right)$ as function of $a = k/k_F$. Note that the slope for $k = k_F$ is infinite. For $k >> k_F$ the function approaches zero.

Appendix O

Derivation of Slater-Condon rules

In this section we derive the Slater-Condon rules spelled out in section 8.5. on p. 160.

We assume that the Slater determinants are in maximum coincidence spelled out on p 160 The principle of maximum coincidence can be used to place the one-particle orbitals that differ between the two Slater determinants to the front. The other one-particle orbitals, which are identical in both Slater determinants are placed in the identical position of both determinants.

O.1 Identical Slater determinants

The result has already been obtained when we worked out the equations for the Hartree-Fock method, where we obtained Eq. 6.6 and Eq. 6.13 on p. 101ff. They are identical to the first Slater-Condon rule Eq. 8.4 given above.

O.2 One-particle operator with one different orbital

First, we consider Slater determinants

$$\langle \vec{x}_1,\ldots,\vec{x}_N|\Psi\rangle = \frac{1}{\sqrt{N!}}\sum_{i_1,\ldots,i_N=1}^{N}\epsilon_{i_1,\ldots,i_N}\langle\vec{x}_1|\psi_{i_1}\rangle\cdots\langle\vec{x}_N|\psi_{i_N}\rangle \tag{O.1}$$

$$\langle \vec{x}_1,\ldots,\vec{x}_N|\Phi\rangle = \frac{1}{\sqrt{N!}}\sum_{j_1,\ldots,j_N=1}^{N}\epsilon_{j_1,\ldots,j_N}\langle\vec{x}_1|\phi_{j_1}\rangle\cdots\langle\vec{x}_N|\phi_{j_N}\rangle \tag{O.2}$$

that differ by exactly one one-particle orbital. Due to the form of **maximum coincidence** defined in section 8.5, we can place these two orbitals in the first position. Hence, $|\psi_1\rangle \neq |\phi_1\rangle$, but $|\psi_i\rangle = |\phi_i\rangle$ for $i \neq 1$. Thus, $\langle\psi_i|\phi_j\rangle = \delta_{i,j}(1-\delta_{i,1})$.

$$\begin{aligned}
\langle\Psi|\hat{A}_1|\Phi\rangle &= \frac{1}{N!}\sum_{i_1,\ldots,i_N=1}^{N}\sum_{j_1,\ldots,j_N=1}^{N}\epsilon_{i_1,i_2,\ldots,i_N}\epsilon_{j_1,j_2,\ldots,j_N}\\
&\quad\cdot\langle\psi_{i_1}|\hat{A}|\phi_{j_1}\rangle\underbrace{\langle\psi_{i_2}|\phi_{j_2}\rangle}_{\delta_{i_2,j_2}(1-\delta_{i_2,1})}\cdots\underbrace{\langle\psi_{i_N}|\phi_{j_N}\rangle}_{\delta_{i_N,j_N}(1-\delta_{i_N,1})}\\
&= \frac{1}{N!}\sum_{i_1,\ldots,i_N=1}^{N}\sum_{j_1=1}^{N}\epsilon_{i_1,i_2,\ldots,i_N}\epsilon_{j_1,i_2,\ldots,i_N}\\
&\quad\cdot\langle\psi_{i_1}|\hat{A}|\phi_{j_1}\rangle(1-\delta_{i_2,1})\ldots(1-\delta_{i_N,1})
\end{aligned} \tag{O.3}$$

Because the Levi-Civita symbol is only non-zero if all indices differ, the product of the two Levi-Civita symbols can only be non-zero if their first index is identical.

$$\langle\Psi|\hat{A}_1|\Phi\rangle = \frac{1}{N!}\sum_{i_1,\dots,i_N=1}^{N}\left(\epsilon_{i_1,i_2,\dots,i_N}\right)^2\langle\psi_{i_1}|\hat{A}|\phi_{i_1}\rangle(1-\delta_{i_2,1})\dots(1-\delta_{i_N,1})$$

Furthermore the there is a contribution only if the first index i_1 has the value 1. If that is not the case there is a term $(1-\delta_{i,1})$ that vanishes.

$$\langle\Psi|\hat{A}_1|\Phi\rangle = \frac{1}{N!}\sum_{i_2\dots,i_N=1}^{N}\left(\epsilon_{1,i_2,\dots,i_N}\right)^2\langle\psi_1|\hat{A}|\phi_1\rangle \tag{O.4}$$

The Levi-Civita symbols contribute, if all indices are different, This happens $(N-1)!$ times, which corresponds to the number of permutations of the indices $2,\dots,N$. Thus, we obtain

$$\langle\Psi|\hat{A}_1|\Phi\rangle = \frac{1}{N}\langle\psi_1|\hat{A}|\phi_1\rangle \tag{O.5}$$

Since the operator has the form $\hat{A}=\sum_{i=1}^{N}\hat{A}_i$, and since all operators $\hat{A}_i$ contribute the same result, we obtain

$$\langle\Psi|\hat{A}|\Phi\rangle = \langle\psi_1|\hat{A}|\phi_1\rangle \tag{O.6}$$

which corresponds to the first line in the second Slater-Condon rule Eq. 8.5.

O.3 Two-particle operator with one different orbital

We consider the matrix element between two Slater determinants

$$\langle\vec{x}_1,\dots,\vec{x}_N|\Psi\rangle = \frac{1}{\sqrt{N!}}\sum_{i_1,\dots,i_N=1}^{N}\epsilon_{i_1,\dots,i_N}\langle\vec{x}_1|\psi_{i_1}\rangle\cdots\langle\vec{x}_N|\psi_{i_N}\rangle \tag{O.7}$$

$$\langle\vec{x}_1,\dots,\vec{x}_N|\Phi\rangle = \frac{1}{\sqrt{N!}}\sum_{j_1,\dots,j_N=1}^{N}\epsilon_{j_1,\dots,j_N}\langle\vec{x}_1|\phi_{j_1}\rangle\cdots\langle\vec{x}_N|\phi_{j_N}\rangle \tag{O.8}$$

with all one-particle orbitals pairwise equal, except for the two first orbitals, that is

$$\langle\psi_i|\phi_j\rangle = \left\{\begin{matrix}\delta_{i,j} \text{ for } & i,j\in\{2,\dots,N\} & \\ 0 \text{ for } & i=1 \quad \text{and/or} & j=1\end{matrix}\right\} = \delta_{i,j}(1-\delta_{i,1}) \tag{O.9}$$

Note, that the Slater determinants are chosen in **maximum coincidence**, defined in section 8.5: the two orbitals that are present in only one of the two Slater determinants stand at the same position, namely the first. The orbitals which are identical for both Slater determinants are in the same position.

Now, we can work out the matrix element of the interaction operator $\hat{W}_{1,2}$ that acts exclusively on the first two particle coordinates. If the interaction would act onto any other orbital, at least one of the orbitals in the first position would be connected with another one by an overlap matrix element. This overlap matrix element vanishes because the orbitals in the first position are orthogonal to all other orbitals involved.

Later, we will see that the result is the same for each pair of coordinates, so that the sum over pairs is done easily at the end of the calculation.

$$
\begin{aligned}
\langle\Psi|\hat{W}_{1,2}|\Phi\rangle &= \frac{1}{N!}\sum_{i_1,\ldots,i_N=1}^{N}\sum_{j_1,\ldots,j_N=1}^{N}\epsilon_{i_1,i_2,\ldots,i_N}\epsilon_{j_1,j_2,\ldots,j_N} \\
&\quad\cdot\langle\psi_{i_1}\psi_{i_2}|\hat{W}_{1,2}|\phi_{j_1}\phi_{j_2}\rangle\underbrace{\langle\psi_{i_3}|\phi_{j_3}\rangle}_{\delta_{i_3,j_3}(1-\delta_{i_3,1})}\ldots\underbrace{\langle\psi_{i_N}|\phi_{j_N}\rangle}_{\delta_{i_N,j_N}(1-\delta_{i_N,1})} \\
&= \frac{1}{N!}\sum_{i_1,\ldots,i_N=1}^{N}\sum_{j_1,j_2=1}^{N}\epsilon_{i_1,i_2,i_3,\ldots,i_N}\epsilon_{j_1,j_2,i_3,\ldots,i_N} \\
&\quad\cdot\langle\psi_{i_1}\psi_{i_2}|\hat{W}|\phi_{j_1}\phi_{j_2}\rangle(1-\delta_{i_3,1})\ldots(1-\delta_{i_N,1})
\end{aligned}
$$

Now we can work out the sum over the indices j_1, j_2. Because the antisymmetric tensor vanishes, whenever two indices are equal, the indices j_1, j_2 must be both in the set $\{i_1, i_2\}$. Furthermore the two indices j_1, j_2 must differ, again to avoid that the antisymmetric tensor vanishes. Hence, only two terms from the second sum survive, namely the one with $(j_1, j_2) = (i_1, i_2)$ and the one where the two indices are interchanged $(j_1, j_2) = (i_2, i_1)$.

$$
\begin{aligned}
\langle\Psi|\hat{W}_{1,2}|\Phi\rangle &= \frac{1}{N!}\sum_{i_1,\ldots,i_N=1}^{N}\Big[(\epsilon_{i_1,i_2,\ldots,i_N})^2\,\langle\psi_{i_1}\psi_{i_2}|\hat{W}_{1,2}|\phi_{i_1}\phi_{i_2}\rangle \\
&\qquad\qquad +\epsilon_{i_1,i_2,i_3,\ldots,i_N}\epsilon_{i_2,i_1,i_3,\ldots,i_N}\langle\psi_{i_1}\psi_{i_2}|\hat{W}_{1,2}|\phi_{i_2}\phi_{i_1}\rangle\Big] \\
&\quad\cdot(1-\delta_{i_3,1})\ldots(1-\delta_{i_N,1}) \\
&= \frac{1}{N!}\sum_{i_1,\ldots,i_N=1}^{N}(\epsilon_{i_1,i_2,\ldots,i_N})^2\Big(\langle\psi_{i_1}\psi_{i_2}|\hat{W}_{1,2}|\phi_{i_1}\phi_{i_2}\rangle-\langle\psi_{i_1}\psi_{i_2}|\hat{W}_{1,2}|\phi_{i_2}\phi_{i_1}\rangle\Big) \\
&\quad\cdot(1-\delta_{i_3,1})\ldots(1-\delta_{i_N,1})
\end{aligned}
$$

Furthermore, there is a contribution only if one of the first two indices i_1, i_2 has the value 1. If that is not the case, there is a term $(1-\delta_{i,1})$ that vanishes.

$$
\begin{aligned}
\langle\Psi|\hat{W}_{1,2}|\Phi\rangle &= \frac{1}{N!}\sum_{i_2,i_3,\ldots,i_N=1}^{N}(\epsilon_{1,i_2,\ldots,i_N})^2\Big(\langle\psi_1\psi_{i_2}|\hat{W}_{1,2}|\phi_1\phi_{i_2}\rangle-\langle\psi_1\psi_{i_2}|\hat{W}_{1,2}|\phi_{i_2}\phi_1\rangle\Big) \\
&\quad+\frac{1}{N!}\sum_{i_1,i_3,\ldots,i_N=1}^{N}(\epsilon_{i_1,1,,i_3\ldots,i_N})^2\Big(\langle\psi_{i_1}\psi_1|\hat{W}_{1,2}|\phi_{i_1}\phi_1\rangle-\langle\psi_{i_1}\psi_1|\hat{W}_{1,2}|\phi_1\phi_{i_1}\rangle\Big) \\
&= \frac{2}{N!}\sum_{n=2}^{N}\Big(\langle\psi_1\psi_n|\hat{W}_{1,2}|\phi_1\phi_n\rangle-\langle\psi_1\psi_n|\hat{W}_{1,2}|\phi_n\phi_1\rangle\Big)\underbrace{\sum_{i_3,\ldots,i_N=1}^{N}(\epsilon_{1,n,\ldots,i_N})^2}_{(N-2)!} \\
&= \underbrace{\frac{2(N-2)!}{N!}}_{\frac{2}{N(N-1)}}\sum_{n=2}^{N}\Big(\langle\psi_1\psi_n|\hat{W}_{1,2}|\phi_1\phi_n\rangle-\langle\psi_1\psi_n|\hat{W}_{1,2}|\phi_n\phi_1\rangle\Big)
\end{aligned}
$$

Since the operator has the form $\hat{W}=\frac{1}{2}\sum_{i\neq j}^{N}\hat{W}_{i,j}$. Each pair contributes the same result. As there are $N(N-1)$ distinct pairs in the double-sum, we obtain

$$
\langle\Psi|\hat{W}|\Phi\rangle=\sum_{n=1}^{N}\left(\langle\psi_1\psi_n|\hat{W}_{1,2}|\phi_1\phi_n\rangle-\langle\psi_1\psi_n|\hat{W}_{1,2}|\phi_n\phi_1\rangle\right)
$$

This result corresponds to the second line in the second Slater-Condon rule Eq. 8.5. The sum runs now over all terms because the element with $n = 1$ cancels anyway.

O.4 One-particle operator with two different orbitals

$$
\begin{aligned}
\langle\Psi|\hat{A}_1|\Phi\rangle = \frac{1}{N!}\sum_{i_1,\ldots,i_N=1}^{N}\sum_{j_1,\ldots,j_N=1}^{N}\epsilon_{i_1,i_2,\ldots,i_N}\epsilon_{j_1,j_2,\ldots,j_N}\\
\cdot\langle\psi_{i_1}|\hat{A}|\phi_{j_1}\rangle \underbrace{\langle\psi_{i_2}|\phi_{j_2}\rangle}_{\delta_{i_2,j_2}(1-\delta_{i_2,1})(1-\delta_{i_2,2})} \ldots \underbrace{\langle\psi_{i_N}|\phi_{j_N}\rangle}_{\delta_{i_N,j_N}(1-\delta_{i_N,1})(1-\delta_{i_N,2})}
\end{aligned}
$$

It is evident that this matrix element vanishes, because in each term there is at least one scalar product between to orbitals that differ in the two slater determinants.

This result corresponds to the first line in the third Slater-Condon rule Eq. 8.6.

O.5 Two-particle operator with two different orbitals

$$
\begin{aligned}
\langle\Psi|\hat{W}_{x_1,x_2}|\Phi\rangle &= \frac{1}{N!}\sum_{i_1,\ldots,i_N}\sum_{j_1,\ldots,j_N}\epsilon_{i_1,i_2,\ldots,i_N}\epsilon_{j_1,j_2,\ldots,j_N}\\
&\quad\cdot\langle\psi_{i_1}\psi_{i_2}|\hat{W}|\phi_{j_1}\phi_{j_2}\rangle \underbrace{\langle\psi_{i_3}|\phi_{j_3}\rangle}_{\delta_{i_3,j_3}(1-\delta_{i_3,1})(1-\delta_{i_3,2})} \ldots \underbrace{\langle\psi_{i_N}|\phi_{j_N}\rangle}_{\delta_{i_N,j_N}(1-\delta_{i_N,1})(1-\delta_{i_N,2})}\\
&= \frac{1}{N!}\sum_{i_1,i_2=1}^{2}\sum_{j_1,j_2=1}^{2}\langle\psi_{i_1}\psi_{i_2}|\hat{W}|\phi_{j_1}\phi_{j_2}\rangle\sum_{i_3,\ldots,i_N}\epsilon_{i_1,i_2,i_3\ldots,i_N}\epsilon_{j_1,j_2,i_3\ldots,i_N}\\
&= \frac{1}{N!}\sum_{i_1,i_2=1}^{2}\sum_{j_1,j_2=1}^{2}\langle\psi_{i_1}\psi_{i_2}|\hat{W}|\phi_{j_1}\phi_{j_2}\rangle(N-2)!\left(\delta_{i_1,j_1}\delta_{i_2,j_2}-\delta_{i_1,j_2}\delta_{i_2,j_1}\right)\\
&= \frac{2}{N(N-1)}\left(\langle\psi_1\psi_2|\hat{W}|\phi_1\phi_2\rangle-\langle\psi_1\psi_2|\hat{W}|\phi_2\phi_1\rangle\right)
\end{aligned}
$$

With the interaction $\hat{W}=\frac{1}{2}\sum_{i\neq j}^{N}\hat{W}_{i,j}$ we obtain with $\langle\Psi|W_{i,j}|\Phi\rangle=\langle\Psi|W_{1,2}|\Phi\rangle$

$$
\langle\Psi|\hat{W}|\Phi\rangle=\langle\psi_1\psi_2|\hat{W}|\phi_1\phi_2\rangle-\langle\psi_1\psi_2|\hat{W}|\phi_2\phi_1\rangle
$$

This result corresponds to the second line in the third Slater-Condon rule Eq. 8.6.

O.6 More than two different one-particle orbitals

The result vanishes. The argument is analogous to that about in Section O.4

Appendix P

One- and two-particle operators expressed by field operators

Here, we will explicitly derive the form of one- and two-particle operators by field operators as given by Eqs. 8.8 and 8.9 on p. 162.

OPERATORS EXPRESSED BY ANNIHILATION AND CREATION OPERATORS

A one-particle operator has the form

$$\hat{A}=\sum_{i,j}A_{i,j}\hat{c}_i^\dagger\hat{c}_j \tag{P.1}$$

and a two-particle operator has the form

$$\hat{W}=\frac{1}{2}\sum_{i,j,k,l}W_{i,j,k,l}\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_l\hat{c}_k \tag{P.2}$$

Note the reversed order of the annihilators relative to the indices!

Our approach will be to work out the matrix elements of a product of creation and annihilation operators between Slater determinants Then the result will be compared with the Slater-Condon rules Eqs. 8.4 -8.6 on p. 161.

Our argument is restricted to Slater determinants constructed from the zero state by a product of creation operators in a unique basis of one-particle orbitals. basis. Thus, these states are eigenstates of the occupation-number operator

$$\hat{n}_k=\hat{c}_k^\dagger\hat{c}_k$$

P.1 Matrix elements between identical Slater determinants

One-particle operator

Let us determine the matrix element of a one-particle operator between two identical Slater determinants $|\Phi\rangle$.

In the first step we used that $\langle\Phi|\hat{c}_i^\dagger\hat{c}_j|\Phi\rangle=0$ for $i\neq j$: Applied to a Slater determinant $\hat{c}_i^\dagger\hat{c}_i$ produces a different Slater determinant that is orthogonal to the original one. Thus, only the diagonal terms contribute.

$$\langle\Phi|\hat{c}_i^\dagger\hat{c}_j|\Phi\rangle = \delta_{i,j}\langle\Phi|\underbrace{\hat{n}_i}_{\hat{c}_i^\dagger\hat{c}_i}|\Phi\rangle = \delta_{i,j}n_i$$

$$\Rightarrow \langle\Phi|\sum_{i,j}A_{i,j}\hat{c}_i^\dagger\hat{c}_j|\Phi\rangle = \sum_i n_iA_{i,i} \qquad \text{(P.3)}$$

Eq. P.3 is identical to the first Slater-Condon rule Eq. 8.4 for one-particle operators on p.161. In Eq. P.3 we sum over all one-particle basis functions and the occupation numbers n_i select those that are present in the Slater determinant. In Eq. 8.4 the sum goes directly over the one-orbitals in the determinant.

Two-particle operator

To work out the expressions for a two-particle operator, we use a similar argument. Unless the operator $\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l$ creates all orbitals that it has destroyed, the resulting Slater determinant will differ from and be orthogonal to the original Slater determinant. There are just two possibilities, namely

- case 1: $i=k$ and $j=l$.

$$\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_i\hat{c}_j \overset{[\hat{c}_i^\dagger,\hat{c}_j]_+=\delta_{i,j}}{=} \hat{c}_i^\dagger\delta_{i,j}\hat{c}_j - \hat{c}_i^\dagger\hat{c}_i\hat{c}_j^\dagger\hat{c}_j = \delta_{i,j}\hat{n}_i - \hat{n}_i\hat{n}_j$$

- case 2: $i=l$ and $j=k$

$$\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_j\hat{c}_i \overset{[\hat{c}_j,\hat{c}_i]_+=0}{=} -\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_i\hat{c}_j \overset{[\hat{c}_i^\dagger,\hat{c}_j]_+=\delta_{i,j}}{=} -\hat{c}_i^\dagger\delta_{i,j}\hat{c}_j + \hat{c}_i^\dagger\hat{c}_i\hat{c}_j^\dagger\hat{c}_j = -\delta_{i,j}\hat{n}_i + \hat{n}_i\hat{n}_j$$

- All other cases vanish because the creators and annihilators do not match.

Thus, we obtain

$$\begin{aligned}
\langle\Phi|\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l|\Phi\rangle &= \delta_{i,k}\delta_{j,l}\Big(\delta_{i,j}n_i - n_in_j\Big) + \delta_{i,l}\delta_{j,k}\Big(-\delta_{i,j}n_i + n_in_j\Big)\\
&= \underbrace{\delta_{i,k}\delta_{j,l}\delta_{i,j}n_i}_{a} - \underbrace{\delta_{i,k}\delta_{j,l}n_in_j}_{b} - \underbrace{\delta_{i,l}\delta_{j,k}\delta_{i,j}n_i}_{c} + \underbrace{\delta_{i,l}\delta_{j,k}n_in_j}_{d}\\
&\overset{a=c}{=} \underbrace{\delta_{i,l}\delta_{j,k}n_in_j}_{d} - \underbrace{\delta_{i,k}\delta_{j,l}n_in_j}_{b}
\end{aligned}$$

which yields

$$\Rightarrow \langle\Phi|\sum_{i,j,k,l}W_{i,j,k,l}\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_l\hat{c}_k|\Phi\rangle = \sum_{i,j}n_in_j\left[W_{i,j,i,j} - W_{i,j,j,i}\right]$$

Note, that we interchanged the two annihilators which changes the sign of the expression.

Let us compare this result again with that of the first Slater-Condon rule Eq. 8.4 on p. 161.

P.2 Matrix elements between Slater determinants differing by one orbital

We can construct such pairs of N-particle Slater determinants from a common $(N-1)$-particle Slater determinant $|\Phi\rangle$ by creating two different one-particle orbitals (one for each). We obtain $\hat{c}_a^\dagger|\Phi\rangle$ and $\hat{c}_b^\dagger|\Phi\rangle$. They differ by exactly one one-particle orbital since $a \neq b$ and they are automatically in maximum coincidence. Remember, that the bra related to $\hat{c}_a^\dagger|\Phi\rangle$ is $\langle\hat{c}_a^\dagger\Phi| = \langle\Phi|\hat{c}_a$.

Now we construct the matrix element of a one-particle operator

$$\begin{aligned}\langle\Phi|\hat{c}_a\hat{c}_i^\dagger\hat{c}_j\hat{c}_b^\dagger|\Phi\rangle &= \delta_{a,i}(1-n_i)\delta_{j,b}(1-n_b)\\ \Rightarrow \langle\Phi|\hat{c}_a\left(\sum_{i,j}A_{i,j}\hat{c}_i^\dagger\hat{c}_j\right)\hat{c}_b^\dagger|\Phi\rangle &= A_{a,b}(1-n_a)(1-n_b)\end{aligned}$$

This result can be compared to the Slater-Condon rule Eq. 8.5 on p. 161. The role of the terms $(1-n_a)$ and $(1-n_b)$ is to test if the corresponding one-particle orbital is already present in $|\Phi\rangle$. If it is present that creator, $c_a^\dagger$ or $c_b^\dagger$, would turn this state into a zero state, for which the matrix element vanishes.

For the two-particle operator we have to connect all annihilators with creators of the same orbital.

$$\begin{aligned}\langle\Phi|\hat{c}_a\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l\hat{c}_b^\dagger|\Phi\rangle &= \delta_{a,i}\delta_{k,j}\delta_{l,b}\langle\Phi|\hat{c}_a\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l\hat{c}_b^\dagger|\Phi\rangle\\ &+\delta_{a,i}\delta_{k,b}\delta_{l,j}\langle\Phi|\hat{c}_a\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l\hat{c}_b^\dagger|\Phi\rangle\\ &+\delta_{a,j}\delta_{k,i}\delta_{l,b}\langle\Phi|\hat{c}_a\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l\hat{c}_b^\dagger|\Phi\rangle\\ &+\delta_{a,j}\delta_{k,b}\delta_{l,i}\langle\Phi|\hat{c}_a\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l\hat{c}_b^\dagger|\Phi\rangle\\ &+\underbrace{\delta_{a,b}}_{=0}\delta_{k,i}\delta_{l,j}\langle\Phi|\hat{c}_a\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l\hat{c}_b^\dagger|\Phi\rangle\\ &+\underbrace{\delta_{a,b}}_{=0}\delta_{k,j}\delta_{l,i}\langle\Phi|\hat{c}_a\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l\hat{c}_b^\dagger|\Phi\rangle\end{aligned}$$

The last tow terms vanish because we required the two orbitals $|\phi_a\rangle$ and $|\phi_b\rangle$ to differ. Hence, $a\neq b$.

With the help of the delta functions, we can reduce the number if differing indices in the products of creation and annihilation operators. This yields

$$\begin{aligned}\langle\Phi|\hat{c}_a\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l\hat{c}_b^\dagger|\Phi\rangle &\overset{a\neq b}{=} \delta_{a,i}\delta_{k,j}\delta_{l,b}\langle\Phi|\hat{c}_a\hat{c}_a^\dagger\hat{c}_j^\dagger\hat{c}_j\hat{c}_b\hat{c}_b^\dagger|\Phi\rangle\\ &+\ \delta_{a,i}\delta_{k,b}\delta_{l,j}\langle\Phi|\hat{c}_a\hat{c}_a^\dagger\hat{c}_j^\dagger\hat{c}_b\hat{c}_j\hat{c}_b^\dagger|\Phi\rangle\\ &+\ \delta_{a,j}\delta_{k,i}\delta_{l,b}\langle\Phi|\hat{c}_a\hat{c}_i^\dagger\hat{c}_a^\dagger\hat{c}_i\hat{c}_b\hat{c}_b^\dagger|\Phi\rangle\\ &+\ \delta_{a,j}\delta_{k,b}\delta_{l,i}\langle\Phi|\hat{c}_a\hat{c}_i^\dagger\hat{c}_a^\dagger\hat{c}_b\hat{c}_i\hat{c}_b^\dagger|\Phi\rangle\end{aligned}$$

Now we interchange the operators so that every pair having the same index stands next to each other. We only need to interchange annihilation with annihilation operators and creation with creation operators. Due to the anticommutator relations, each interchange causes a sign change.

$$\begin{aligned}\langle\Phi|\hat{c}_a\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l\hat{c}_b^\dagger|\Phi\rangle &= \delta_{a,i}\delta_{k,j}\delta_{l,b}\langle\Phi|\hat{c}_a\hat{c}_a^\dagger\hat{c}_j^\dagger\hat{c}_j\hat{c}_b\hat{c}_b^\dagger|\Phi\rangle\\ &-\delta_{a,i}\delta_{k,b}\delta_{l,j}\langle\Phi|\hat{c}_a\hat{c}_a^\dagger\hat{c}_j^\dagger\hat{c}_j\hat{c}_b\hat{c}_b^\dagger|\Phi\rangle\\ &-\delta_{a,j}\delta_{k,i}\delta_{l,b}\langle\Phi|\hat{c}_a\hat{c}_a^\dagger\hat{c}_i^\dagger\hat{c}_i\hat{c}_b\hat{c}_b^\dagger|\Phi\rangle\\ &+\delta_{a,j}\delta_{k,b}\delta_{l,i}\langle\Phi|\hat{c}_a\hat{c}_a^\dagger\hat{c}_i^\dagger\hat{c}_i\hat{c}_b\hat{c}_b^\dagger|\Phi\rangle\end{aligned}$$

Finally exploit that our Slater determinant $|\Phi\rangle$ is an eigenstate of the occupation-number operators

$$\begin{aligned}\hat{c}_i^\dagger\hat{c}_i|\Phi\rangle &= |\Phi\rangle n_i\\ \hat{c}_i\hat{c}_i^\dagger|\Phi\rangle &= |\Phi\rangle(1-n_i)\end{aligned}$$

where n_i are the eigenvalues, which are the occupations of the orbital in $|\Phi\rangle$.

$$\begin{aligned}\langle\Phi|\hat{c}_a\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l\hat{c}_b^\dagger|\Phi\rangle &= \delta_{a,i}\delta_{k,j}\delta_{l,b}(1-n_b)n_j(1-n_a)\\ &-\delta_{a,i}\delta_{k,b}\delta_{l,j}(1-n_b)n_j(1-n_a)\\ &-\delta_{a,j}\delta_{k,i}\delta_{l,b}(1-n_b)n_i(1-n_a)\\ &+\delta_{a,j}\delta_{k,b}\delta_{l,i}(1-n_b)n_i(1-n_a)\end{aligned}$$

Thus, we obtain

$$
\begin{aligned}
\langle\Phi|\hat{c}_a\hat{W}\hat{c}_b^\dagger|\Phi\rangle &= \langle\Phi|\hat{c}_a\left(\frac{1}{2}\sum_{i,j,k,l}W_{i,j,k,l}\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_l\hat{c}_k\right)\hat{c}_b^\dagger|\Phi\rangle \\
&= -\frac{1}{2}\sum_{i,j,k,l}W_{i,j,k,l}\langle\Phi|\hat{c}_a\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l\hat{c}_b^\dagger|\Phi\rangle \\
&= -\frac{1}{2}\sum_j\Big(W_{a,j,j,b}(1-n_a)n_j(1-n_b) - W_{a,j,b,j}(1-n_a)n_j(1-n_b) \\
&\quad - W_{j,a,j,b}(1-n_a)n_j(1-n_b) + W_{j,a,b,j}(1-n_a)n_j(1-n_b)\Big) \\
&= -(1-n_a)(1-n_b)\frac{1}{2}\sum_j n_j\Big(W_{a,j,j,b} - W_{a,j,b,j} - W_{j,a,j,b} + W_{j,a,b,j}\Big)
\end{aligned}
$$

Let us have a look at the definition Eq. 8.3 of the matrix elements from p. 161.

$$
W_{i,j,k,l} \overset{\text{Eq. 8.3}}{\overset{\text{def}}{=}} \int d^4x \int d^4x'\ \phi_i^*(\vec{x})\phi_j^*(\vec{x'})W(\vec{x},\vec{x'})\phi_k(\vec{x})\phi_l(\vec{x'})
$$

We see that we can interchange simultaneously the first two and the last two indices, if $W(\vec{x},\vec{x'}) = W(\vec{x'},\vec{x})$. For the interaction this is indeed the case due to "actio=reactio". Thus, $W_{a,j,j,b} = W_{j,a,b,j}$ and $W_{a,j,b,j} = W_{j,a,j,b}$ and

$$
\begin{aligned}
\langle\Phi|\hat{c}_a\hat{W}\hat{c}_b^\dagger|\Phi\rangle &= -(1-n_a)(1-n_b)\frac{1}{2}\sum_j n_j\Big(W_{a,j,j,b} - W_{a,j,b,j}\Big) \\
&= (1-n_a)(1-n_b)\frac{1}{2}\sum_{j=1}^{\infty} n_j\Big(W_{a,j,b,j} - W_{a,j,j,b}\Big)
\end{aligned}
$$

The first term ensures, that the Result vanishes, if $|\Phi\rangle$ already contains the orbital $|\phi_a\rangle$ or the orbital $|\phi_b\rangle$.

Eq. P.4 is the result of the second Slater-Condon rule Eq. 8.5 on p. 161.

P.3 Matrix elements between Slater determinants differing by two orbitals

The procedure is analogous to the above case. We construct the N-electron Slater determinants $\hat{c}_a^\dagger\hat{c}_b^\dagger|\Phi\rangle$ and $\hat{c}_c^\dagger\hat{c}_d^\dagger|\Phi\rangle$ from the same $N-2$ electron Slater determinant $|\Phi\rangle$. Then we select those terms where all creators are paired with an annihilator of the same orbital. All other terms vanish. Note that the bra related to $\hat{c}_a^\dagger\hat{c}_b^\dagger|\Phi\rangle$ is $\langle\Phi|\hat{c}_b\hat{c}_a$, because $\left(\hat{c}_a^\dagger\hat{c}_b^\dagger\right)^\dagger = \hat{c}_b\hat{c}_a$.

For the one-particle operator, the matrix element vanishes consistent with the third Slater-Condon rule Eq. 8.6, because the one-particle operator can only match one annihilator.

For the two-particle operator, we obtain

$$
\begin{array}{cl}
 & \langle\Phi|\hat{c}_b\hat{c}_a\left(\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l\right)\hat{c}_c^\dagger\hat{c}_d^\dagger|\Phi\rangle \\
= & \delta_{b,i}\delta_{a,j}\delta_{k,c}\delta_{l,d}\langle\Phi|\hat{c}_b\hat{c}_a\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l\hat{c}_c^\dagger\hat{c}_d^\dagger|\Phi\rangle \\
+ & \delta_{b,i}\delta_{a,j}\delta_{k,d}\delta_{l,c}\langle\Phi|\hat{c}_b\hat{c}_a\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l\hat{c}_c^\dagger\hat{c}_d^\dagger|\Phi\rangle \\
+ & \delta_{b,j}\delta_{a,i}\delta_{k,c}\delta_{l,d}\langle\Phi|\hat{c}_b\hat{c}_a\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l\hat{c}_c^\dagger\hat{c}_d^\dagger|\Phi\rangle \\
+ & \delta_{b,j}\delta_{a,i}\delta_{k,d}\delta_{l,c}\langle\Phi|\hat{c}_b\hat{c}_a\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_k\hat{c}_l\hat{c}_c^\dagger\hat{c}_d^\dagger|\Phi\rangle \\
\hline
= & \delta_{b,i}\delta_{a,j}\delta_{k,c}\delta_{l,d}\langle\Phi|\hat{c}_b\hat{c}_a\hat{c}_b^\dagger\hat{c}_a^\dagger\hat{c}_c\hat{c}_d\hat{c}_c^\dagger\hat{c}_d^\dagger|\Phi\rangle \\
+ & \delta_{b,i}\delta_{a,j}\delta_{k,d}\delta_{l,c}\langle\Phi|\hat{c}_b\hat{c}_a\hat{c}_b^\dagger\hat{c}_a^\dagger\hat{c}_d\hat{c}_c\hat{c}_c^\dagger\hat{c}_d^\dagger|\Phi\rangle \\
+ & \delta_{b,j}\delta_{a,i}\delta_{k,c}\delta_{l,d}\langle\Phi|\hat{c}_b\hat{c}_a\hat{c}_a^\dagger\hat{c}_b^\dagger\hat{c}_c\hat{c}_d\hat{c}_c^\dagger\hat{c}_d^\dagger|\Phi\rangle \\
+ & \delta_{b,j}\delta_{a,i}\delta_{k,d}\delta_{l,c}\langle\Phi|\hat{c}_b\hat{c}_a\hat{c}_a^\dagger\hat{c}_b^\dagger\hat{c}_d\hat{c}_c\hat{c}_c^\dagger\hat{c}_d^\dagger|\Phi\rangle \\
\hline
\overset{c\neq d;a\neq b}{=} & \delta_{b,i}\delta_{a,j}\delta_{k,c}\delta_{l,d}\langle\Phi|\hat{c}_b{\color{red}\hat{c}_b^\dagger\hat{c}_a}\hat{c}_a^\dagger\hat{c}_c{\color{blue}\hat{c}_c^\dagger\hat{c}_d}\hat{c}_d^\dagger|\Phi\rangle \\
- & \delta_{b,i}\delta_{a,j}\delta_{k,d}\delta_{l,c}\langle\Phi|{\color{blue}\hat{c}_a\hat{c}_a^\dagger\hat{c}_b\hat{c}_b^\dagger}\hat{c}_d{\color{red}\hat{c}_d^\dagger\hat{c}_c\hat{c}_c^\dagger}|\Phi\rangle \\
- & \delta_{b,j}\delta_{a,i}\delta_{k,c}\delta_{l,d}\langle\Phi|{\color{red}\hat{c}_a\hat{c}_a^\dagger\hat{c}_b}\hat{c}_b^\dagger\hat{c}_c{\color{blue}\hat{c}_c^\dagger\hat{c}_d}\hat{c}_d^\dagger|\Phi\rangle \\
+ & \delta_{b,j}\delta_{a,i}\delta_{k,d}\delta_{l,c}\langle\Phi|{\color{red}\hat{c}_a\hat{c}_a^\dagger\hat{c}_b}\hat{c}_b^\dagger\hat{c}_d{\color{blue}\hat{c}_d^\dagger\hat{c}_c\hat{c}_c^\dagger}|\Phi\rangle \\
\hline
= & \delta_{b,i}\delta_{a,j}\delta_{k,c}\delta_{l,d}(1-n_d)(1-n_c)(1-n_a)(1-n_b) \\
- & \delta_{b,i}\delta_{a,j}\delta_{k,d}\delta_{l,c}(1-n_d)(1-n_c)(1-n_a)(1-n_b) \\
- & \delta_{b,j}\delta_{a,i}\delta_{k,c}\delta_{l,d}(1-n_d)(1-n_c)(1-n_a)(1-n_b) \\
+ & \delta_{b,j}\delta_{a,i}\delta_{k,d}\delta_{l,c}(1-n_d)(1-n_c)(1-n_a)(1-n_b)
\end{array}
$$

Thus, we obtain

$$
\begin{aligned}
&\langle\Phi|\hat{c}_b\hat{c}_a\left(\frac{1}{2}\sum_{i,j,k,l}W_{i,j,k,l}\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_l\hat{c}_k\right)\hat{c}_c^\dagger\hat{c}_d^\dagger|\Phi\rangle \\
&={\color{red}-}\frac{1}{2}\sum_{i,j,k,l}W_{i,j,k,l}\langle\Phi|\hat{c}_b\hat{c}_a\left(\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_{\color{red}k}\hat{c}_{\color{red}l}\right)\hat{c}_c^\dagger\hat{c}_d^\dagger|\Phi\rangle \\
&=-\frac{1}{2}\Big(W_{b,a,c,d}-W_{b,a,d,c}-W_{a,b,c,d}+W_{a,b,d,c}\Big) \\
&\quad\cdot(1-n_d)(1-n_c)(1-n_a)(1-n_b)
\end{aligned}
$$

Now we exploit that a joint interchange of the first two indices and the last two indices does not change the matrix element, that is $W_{i,j,k,l}=W_{j,i,l.k}$. This is allowed because the interaction is inversion symmetric, that is $V(\vec{x},\vec{x}')=V(\vec{x},\vec{x}')$, which is a special property of the Coulomb interaction.

$$
\begin{aligned}
&\langle\Phi|\hat{c}_b\hat{c}_a\left(\frac{1}{2}\sum_{i,j,k,l}W_{i,j,k,l}\hat{c}_i^\dagger\hat{c}_j^\dagger\hat{c}_l\hat{c}_k\right)\hat{c}_c^\dagger\hat{c}_d^\dagger|\Phi\rangle \\
&=-\Big(W_{a,b,d,c}-W_{a,b,c,d}\Big)(1-n_d)(1-n_c)(1-n_a)(1-n_b) \\
&=+\Big(W_{a,b,c,d}-W_{a,b,d,c}\Big)(1-n_d)(1-n_c)(1-n_a)(1-n_b)
\end{aligned}
$$

This result corresponds directly to the third Slater-Condon rule Eq. 8.6 on p. 161.

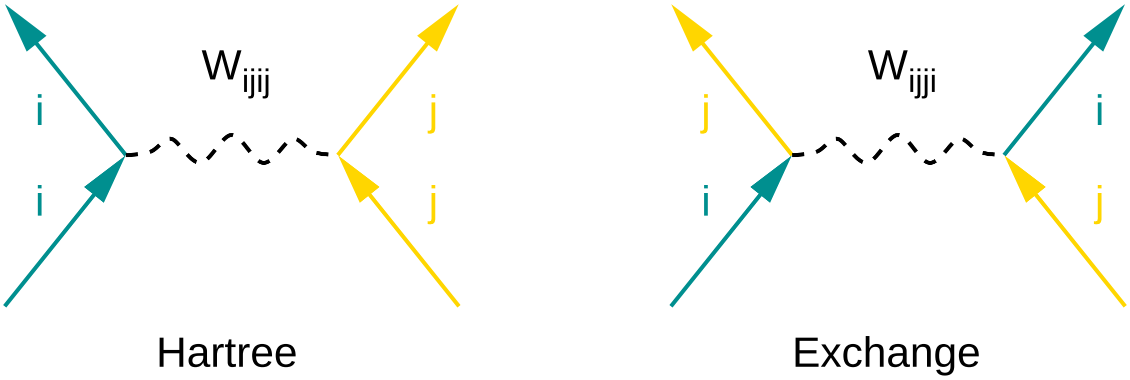


Fig. P.1: The left diagram describes that two particles are scattered by the Coulomb interaction. The right diagram describes the same process, but the two electrons are exchanged. The second process is possible, because the two electrons are indistinguishable so that we cannot detect if the two electrons are still the same or not.

Appendix Q

Green's function and division of systems

Before we return to quantum field theory I would like to show one important application of Green's functions in the one-particle theory. Green's function are important whenever we couple to quantum mechanical systems.

Let us partition a system in two subsystems. Consider for example a defect atom in a host crystal. The defect atom would be described by a Hamiltonian $\boldsymbol{H}_{11}$ and the host crystal by a Hamiltonian $\boldsymbol{H}_{22}$. More precisely we divide Hilbert space up into two subspaces. The subspace of the impurity consists of the wave functions $|\psi_i\rangle$ and the Hilbert space of the host consists of the wave functions $|\phi_i\rangle$. Then the matrix $\boldsymbol{H}_{11}$ is defined by

$$(\boldsymbol{H}_{11})_{i,j} = \langle\psi_i|\hat{H}|\psi_j\rangle$$

First we show how the Green's function is determined

$$\begin{pmatrix} G_{11}(\epsilon) & G_{12}(\epsilon) \\ G_{21}(\epsilon) & G_{22}(\epsilon) \end{pmatrix} \begin{pmatrix} H_{11}-\epsilon & H_{12} \\ H_{21} & H_{22}-\epsilon \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

We obtain four equations

$$G_{11}(\epsilon)\left(H_{11}-\epsilon\right)+G_{12}(\epsilon)H_{21}=1 \tag{Q.1}$$

$$G_{21}(\epsilon)\left(H_{11}-\epsilon\right)+G_{22}(\epsilon)H_{21}=0 \tag{Q.2}$$

$$G_{11}(\epsilon)H_{12}+G_{12}(\epsilon)\left(H_{22}-\epsilon\right)=0 \tag{Q.3}$$

$$G_{21}(\epsilon)H_{12}+G_{22}(\epsilon)\left(H_{22}-\epsilon\right)=1 \tag{Q.4}$$

We first determine the matrices $G_{12}(\epsilon)$ and $G_{21}(\epsilon)$ as function of the diagonal elements

$$G_{21}(\epsilon)=-G_{22}(\epsilon)H_{21}\left(H_{11}-\epsilon\right)^{-1} \tag{Q.5}$$

$$G_{12}(\epsilon)=-G_{11}(\epsilon)H_{12}\left(H_{22}-\epsilon\right)^{-1} \tag{Q.6}$$

and insert the result into the other two equations.

$$G_{11}(\epsilon)\left(H_{11}-\epsilon\right)-G_{11}(\epsilon)H_{12}\left(H_{22}-\epsilon\right)^{-1}H_{21}=1 \tag{Q.7}$$

$$+G_{22}(\epsilon)\left(H_{22}-\epsilon\right)-G_{22}(\epsilon)H_{21}\left(H_{11}-\epsilon\right)^{-1}H_{12}=1 \tag{Q.8}$$

which can be written as

$$G_{11}(\epsilon)=\Big[H_{11}-\epsilon-H_{12}\left(H_{22}-\epsilon\right)^{-1}H_{21}\Big]^{-1} \tag{Q.9}$$

$$G_{22}(\epsilon)=\Big[H_{22}-\epsilon-H_{21}\left(H_{11}-\epsilon\right)^{-1}H_{12}\Big]^{-1} \tag{Q.10}$$

The uncoupled Green's functions G^0_{11} and G^0_{22}, that is the Green's functions of the separated systems, are

$$G^0_{11}(\epsilon) = [H_{11} - \epsilon]^{-1} \tag{Q.11}$$
$$G^0_{22}(\epsilon) = [H_{22} - \epsilon]^{-1} \tag{Q.12}$$

which allows us to rewrite the result for the full Green's function as

$$G_{11}(\epsilon) = \Big[\big(G^0_{11}(\epsilon)\big)^{-1} - H_{12}G^0_{22}(\epsilon)H_{21}\Big]^{-1} \tag{Q.13}$$
$$G_{22}(\epsilon) = \Big[\big(G^0_{22}(\epsilon)\big)^{-1} - H_{21}G^0_{11}(\epsilon)H_{12}\Big]^{-1} \tag{Q.14}$$

These equations provide us in a straightforward fashion to couple for example different impurities into a host crystal. We need to know the Green's functions for the separate systems. Then we need to set up an energy grid and perform a number of matrix inversions for each energy.

Q.1 Relation to a Schrödinger equation

The Green's function itself allows to extract a wealth of useful information. However to get a better insight it may be instructive to derive an effective Schrödinger equation for the defect atom alone, which however also includes the effect of the host crystal accurately. In the following we introduce the Green's function G_{22} of the host crystal which is not to be confused with the 22 element of the total Green's function.

$$\begin{pmatrix} H_{11} - \epsilon & \lambda H_{12} \\ \lambda H_{21} & H_{22} - \epsilon \end{pmatrix} \begin{pmatrix} \vec{c}_1 \\ \vec{c}_2 \end{pmatrix} = 0 \tag{Q.15}$$

We use the second line an resolve it for $\vec{c}_2$

$$\lambda H_{21}\vec{c}_1 + (H_{22} - \epsilon)\vec{c}_2 = 0 \qquad \Rightarrow \qquad \vec{c}_2 = -\underbrace{(H_{22} - \epsilon)^{-1}}_{G_{22}} H_{21}\vec{c}_1$$

and insert the result in the first line of Eq. Q.15.

$$\begin{aligned} (H_{11} - \epsilon)\vec{c}_1 + H_{12}\vec{c}_2 &= 0 \\ \big[(H_{11} - \epsilon) - H_{12}(H - \epsilon)^{-1}H_{21}\big]\,\vec{c}_1 &= 0 \\ [H_{11} - H_{12}G_{22}(\epsilon)H_{21} - \epsilon]\,\vec{c}_1 &= 0 \end{aligned} \tag{Q.16}$$

This equation already looks like a Schrödinger equation in which the environment acts on the first system with an energy-dependent potential $H_{12}G_{22}(\epsilon)H_{21}$.

Before we continue, note however that Eq. Q.16 can also be written as

$$G^{-1}_{11}(\epsilon)\vec{c}_1 = 0$$

Thus, we can find the eigenstates by looking for the poles of the full Green's function. The Green's function has an eigenvalue of zero at the energies where it has poles. The corresponding eigenvalues are the vectors $\vec{c}_1$, which only need to be normalized.

However we still need to consider the normalization condition for $\vec{c}_1$.

$$\begin{aligned} \vec{c}_1^*\vec{c}_1 + \vec{c}_2^*\vec{c}_2 &= 1 \\ \vec{c}_1^*\vec{c}_1 + \underbrace{\vec{c}_1^* H^\dagger_{21}G^\dagger_{22}(\epsilon)}_{\vec{c}_2^*}\underbrace{G_{22}(\epsilon)H_{21}\vec{c}_1}_{\vec{c}_2} &= 1 \\ \vec{c}_1^*\left(1 + H_{12}G^2_{22}(\epsilon)H_{21}\right)\vec{c}_1 &= 1 \end{aligned}$$

The operator in the middle plays a role of an overlap operator.

We can combine the equation Eq. Q.16 with the normalization condition to an energy-dependent Schrödinger-like equation.

$$
\begin{aligned}
\left[H_{11}-H_{12}G_{22}(\epsilon)H_{21}+H_{12}\epsilon G^2_{22}(\epsilon)H_{21}-\epsilon\left(1+H_{12}G^2_{22}(\epsilon)H_{21}\right)\right]\vec{c}_1&=0\\
\left[H_{11}\underbrace{-H_{12}\left(G_{22}(\epsilon)-\epsilon G^2_{22}(\epsilon)\right)H_{21}}_{V^{emb}_{11}}-\epsilon\underbrace{\left(1+H_{12}G^2_{22}(\epsilon)H_{21}\right)}_{O^{eff}_{11}}\right]\vec{c}_1&=0
\end{aligned}
\tag{Q.17}
$$

We may now interpret this equation as a Schrödinger equation

$$\left(H_{11}+V^{emb}_{11}-\epsilon O^{eff}_{11}\right)\vec{c}_1=0$$

of system 1 with an energy dependent overlap matrix

$$O^{eff}_{11}=1+H_{12}G^2_{22}(\epsilon)H_{21}$$

and an energy dependent embedding potential V^{emb}

$$V^{emb}_{11}=-H_{12}\left(G_{22}(\epsilon)-\epsilon G^2_{22}(\epsilon)\right)H_{21}$$

Q.1.1 Another derivation of the effective Schrödinger equation

The same equation can also be obtained by linearizing the equation Eq. Q.16

$$
\begin{aligned}
\left[H_{11}-H_{12}G^0_{22}(\epsilon)H_{21}-\epsilon\right]\vec{c}_1&=0\\
\left[H_{11}-H_{12}G^0_{22}(\epsilon_0)H_{21}-H_{12}\left.\frac{dG^0_{22}}{d\epsilon}\right|_{\epsilon_0}H_{21}(\epsilon-\epsilon_0)-\epsilon\right]\vec{c}_1&=0\\
\left[\underbrace{H_{11}-H_{12}\left[G^0_{22}(\epsilon_0)-\epsilon_0\left.\frac{dG^0_{22}}{d\epsilon}\right|_{\epsilon_0}\right]H_{21}}_{H^{eff}_{11}}-\epsilon\underbrace{\left[1+H_{12}\left.\frac{dG^0_{22}}{d\epsilon}\right|_{\epsilon_0}H_{21}\right]}_{O^{eff}_{11}}\right]\vec{c}_1&=0
\end{aligned}
\tag{Q.18}
$$

It can be shown that the overlap matrix can directly be obtained from the energy derivative of $G^{-1}_{11}(\epsilon)$ or equivalently the differential operator in Eq. Q.16. Precisely,

$$O^{eff}_{11}(\epsilon)=-\frac{dG^{-1}_{11}(\epsilon)}{d\epsilon}$$

Let us first determine the energy derivative of the decoupled Green's function $G^{(0)}_{22}(\epsilon)$. We begin with the defining equation

$$
\begin{aligned}
\left(H_{22}-(\epsilon+d\epsilon)\right)G^{(0)}_{22}(\epsilon+d\epsilon)&=1\\
\left(H_{22}-\epsilon-d\epsilon\right)\left(G^{(0)}_{22}(\epsilon)+\frac{G^{(0)}_{22}}{d\epsilon}\right)&=1+O(d\epsilon^2)\\
\left(H_{22}-\epsilon)\right)G^{(0)}_{22}(\epsilon)+d\epsilon\left[G^{(0)}_{22}(\epsilon)+\left(H_{22}-\epsilon\right)\frac{G^{(0)}_{22}}{d\epsilon}\right]&=1+O(d\epsilon^2)\\
\left[G^{(0)}_{22}(\epsilon)+\left(H_{22}-\epsilon\right)\frac{G^{(0)}_{22}}{d\epsilon}\right]&=O(d\epsilon)\\
\frac{G^{(0)}_{22}}{d\epsilon}=-\underbrace{\left(H_{22}-\epsilon\right)^{-1}}_{G^{(0)}_{22}(\epsilon)}G^{(0)}_{22}(\epsilon)&=-\left(G^{(0)}_{22}(\epsilon)\right)^2
\end{aligned}
$$

Now we can work out the energy derivative of the full Green's function

$$\begin{aligned}\partial_\epsilon G_{11}^{-1}(\epsilon) &= \partial_\epsilon \left[H_{11} - H_{12}G_{22}H_{21} - \epsilon\right]\\ &= -H_{12}\frac{dG_{22}^{(0)}}{d\epsilon}H_{21} - 1\\ &= -\left[1 + H_{12}\left(G_{22}^{(0)}\right)^2 H_{21}\right] = -O_{11}^{eff}\end{aligned}$$

which is what we wanted to show. With this result it is easy to show the equivalence of Eq. Q.18 and Eq. Q.17.

Further information about one-particle Green's functions can be found in the Book by Economou[92].

Appendix R

Time-ordered exponential, Propagators etc.

R.1 Time-ordered exponential

Often we encounter equations of the form of the time-dependent Schrödinger equation with a time-dependent Hamiltonian.

TIME-ORDERED EXPONENTIAL

The solution of a Schrödinger equation with a time-dependent Hamiltonian

$$i\hbar\partial_t|\psi(t)\rangle = \hat{H}(t)|\psi(t)\rangle \tag{R.1}$$

can be solved in the form (Eq. 31 of[93])

$$|\psi(t)\rangle = \left(\mathcal{T}_D \mathrm{e}^{-\frac{i}{\hbar}\int_0^t dt'\,\hat{H}(t')}\right)|\psi(0)\rangle \tag{R.2}$$

Here, $\mathcal{T}_D$ is **Dyson's time ordering operator**[a] or **P-product** (Eq. 29 of [93]) which reorders the time dependent operators in ascending order from right to left, that is

$$\mathcal{T}_D\hat{A}(t)\hat{B}(t') = \begin{cases} \hat{A}(t)\hat{B}(t') & \text{for } t \geq t' \\ \hat{B}(t')\hat{A}(t) & \text{for } t' > t \end{cases} \tag{R.3}$$

---

[a] Dyson's time ordering operator must not be confused with **Wick's time ordering operator**[73], which changes the sign for each permutation of two fermionic operators. If the Hamiltonian is particle number conserving, i.e. if creation and annihilation operators occur in even-numbered products, the action of Dyson's time ordering operator on the Hamiltonian is identical to Wick's time ordering operator acting on individual creation and annihilation operators.

In a first step we convert the differential equation Eq. R.1 into an integral equation

$$|\psi(t)\rangle = |\psi(0)\rangle - \frac{i}{\hbar}\int_0^t dt'\,\hat{H}(t')|\psi(t)\rangle \tag{R.4}$$

When we repeatedly insert the left hand side of Eq. R.4 on the right-hand side, we obtain

$$\begin{aligned}
|\psi(t)\rangle &= |\psi(0)\rangle - \frac{i}{\hbar}\int_0^t dt'\; \hat{H}(t')|\psi(t)\rangle \\
&= |\psi(0)\rangle + \frac{1}{i\hbar}\int_0^t dt'\; \hat{H}(t')|\psi(0)\rangle + \left(\frac{1}{i\hbar}\right)^2 \int_0^t dt'\; \hat{H}(t') \int_0^{t'} dt''\; \hat{H}(t'')|\psi(t)\rangle \\
&= \left[\sum_{n=0}^{\infty} \left(\frac{1}{i\hbar}\right)^n \int_0^t dt_1 \ldots \int_0^{t_{n-1}} dt_n\; \hat{H}(t_1)\cdots\hat{H}(t_n)\right]|\psi(0)\rangle \qquad \text{(R.5)}
\end{aligned}$$

This result is only valid if the series converges. Convergence of the series is a non-trivial issue.

Using Dyson's time ordering operator we can rewrite Eq. R.5 with integrals that all have the same bounds, namely 0 and t.

$$\begin{aligned}
|\psi(t)\rangle &= \left[\sum_{n=0}^{\infty} \mathcal{T}_D \frac{1}{n!}\left(\frac{1}{i\hbar}\right)^n \int_0^t dt_1 \ldots \int_0^t dt_n\; \hat{H}(t_1)\cdots\hat{H}(t_n)\right]|\psi(0)\rangle \\
&= \mathcal{T}_D\left[\sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{1}{i\hbar}\int_0^t dt'\; \hat{H}(t')\right)^n\right]|\psi(0)\rangle \qquad \text{(R.6)}
\end{aligned}$$

In the form used previously, in Eq. R.5, the Hamilton operators are in a specific order, namely ascending from right to left. Here, in Eq. R.6 on on the other hand, all orderings are allowed, but each term is then ordered by the time-ordering operator. However, now we have,instead of one, $n!$ terms according to the $n!$ possible arrangement of the time arguments. This overcounting is corrected by the factor $1/n!$.

We recognize the similarity of this expression Eq. R.6 with the Taylor expansion of the exponential function $\mathrm{e}^x = \sum_{n=0}^{\infty} \frac{1}{n!}x^n$, which allows us to write

$$|\psi(t)\rangle = \left[\mathcal{T}_D \mathrm{e}^{-\frac{i}{\hbar}\int_0^t dt'\; \hat{H}(t')}\right]|\psi(0)\rangle \qquad \text{(R.7)}$$

which proves Eq. R.2.

R.2 Wick's time ordering operator

In his seminal paper[73] Wick laying the foundation for the diagrammatic expansion of the Green's function, he introduced his time ordering operator, which we denote by $\mathcal{T}_W$ or, without subscript as $\mathcal{T}$. Like for Dyson's time ordering operator, we do not add a hat above the symbol, because this operator differs from the operators that act on states. The time ordering operators are fundamentally different: They are merely a prescription on how to rearrange regular operators.

R.3 Propagator in real time

R.3.1 Propagator in the Schrödinger picture

As shown above in section R.1, the solution of a Schrödinger equation

$$i\hbar\partial_t|\psi(t)\rangle = \hat{H}(t)|\psi(t)\rangle$$

can be expressed by the propagator $\hat{U}(t,0)$ to the initial state

$$|\psi(t)\rangle = \hat{U}(t,t')|\Psi(t')\rangle \qquad \text{(R.8)}$$

where

$$\hat{U}(t_2,t_1) \overset{\text{Eq. R.2}}{=} \mathcal{T}_D\; \mathrm{e}^{-\frac{i}{\hbar}\int_{t_1}^{t_2} dt'\; \hat{H}(t')} \qquad \text{(R.9)}$$

Propagator in the interaction picture

If the Hamilton operator can be split into a time-independent term $\hat{H}_0$ and a time-dependent term $\hat{W}$, we usually use the interaction picture: A state $|\psi_I(t)\rangle$ in the interaction picture is related to a state $|\psi_S(t)\rangle$ in the Schrödinger picture as

$$|\psi_I(t)\rangle = \mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t}|\psi_S(t)\rangle$$

This transforms the Schrödinger equation into

$$\begin{aligned}
i\hbar\partial_t|\psi_I(t)\rangle &= -\hat{H}_0|\psi_I(t)\rangle + \mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t} i\hbar\partial_t|\psi_S(t)\rangle \\
&= -\hat{H}_0|\psi_I(t)\rangle + \mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t}(\hat{H}_0 + \hat{W}(t))|\psi_S(t)\rangle \\
&= \left[-\mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t}\hat{H}_0\mathrm{e}^{-\frac{i}{\hbar}\hat{H}_0 t} + \mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t}\Big(\hat{H}_0 + \hat{W}(t)\Big)\mathrm{e}^{-\frac{i}{\hbar}\hat{H}_0 t}\right]|\psi_I(t)\rangle \\
&= \mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t}\hat{W}(t)\mathrm{e}^{-\frac{i}{\hbar}\hat{H}_0 t}|\psi_I(t)\rangle \\
&= \hat{W}_I(t)|\psi_I(t)\rangle
\end{aligned}$$

where $\hat{W}_I(t)$ is the interaction operator in the interaction picture

$$\hat{W}_I(t) = \mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t}\hat{W}_S(t)\mathrm{e}^{-\frac{i}{\hbar}\hat{H}_0 t}$$

Now we can use the expression for the time-ordered exponential to obtain a solution for the propagator in the interaction picture, that is

$$|\Psi_I(t_2)\rangle = \mathcal{T}_D\left[\mathrm{e}^{-\frac{i}{\hbar}\int_{t_1}^{t_2} dt'\,\hat{W}_I(t')}\right]|\Psi_I(t_1)\rangle$$

$$|\Psi_S(t_2)\rangle = \mathrm{e}^{-\frac{i}{\hbar}\hat{H}_0 t_2}\mathcal{T}_D\left[\mathrm{e}^{-\frac{i}{\hbar}\int_{t_1}^{t_2} dt'\,\hat{W}_I(t')}\right]\mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t_1}|\Psi_S(t_1)\rangle$$

Thus, the propagator for Schrödinger states can be written in two forms, namely as

PROPAGATOR IN REAL TIME

$$\hat{U}(t_2,t_1) = \mathcal{T}_D\left[\mathrm{e}^{-\frac{i}{\hbar}\int_{t_1}^{t_2} dt'\,\left(\hat{H}_0+\hat{W}_S(t')\right)}\right] \tag{R.10}$$

$$= \mathrm{e}^{-\frac{i}{\hbar}\hat{H}_0 t_2}\mathcal{T}_D\left[\mathrm{e}^{-\frac{i}{\hbar}\int_{t_1}^{t_2} dt'\,\hat{W}_I(t')}\right]\mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t_1} \tag{R.11}$$

R.4 Propagator in imaginary time

For a brief historical overview see Abeś article.[94]. The use of Feynman diagrams in statistical physics goes back to Matsubara[95]. The original proof of Wick's theorem in the imaginary time formalism is due to Bloch and de Dominicis[96]. Abe says that this proof is complicated and a simpler one has been given later by Gaudin[97] in 1960.

R.4.1 Imaginary-time propagator in the Schrödinger picture

In the finite temperature formulation of quantum mechanics Matsubara generalized time to a complex quantity.

$$\tau = -it$$

In imaginary time, we can write down an equation analogous to the Schrödinger equation, which then obtains the form of a diffusion equation.

$$-\hbar\partial_\tau|\psi(\tau)\rangle = (\hat{H} - \mu\hat{N})|\psi(\tau)\rangle$$

This equation has the propagator $\hat{U}(\tau,\tau')$ which links the wave function between two different times.

$$|\psi(\tau)\rangle = \hat{U}(\tau,\tau')|\psi(\tau')\rangle \tag{R.12}$$

Because the Hamiltonian is independent of time, which is assumed to be the case, the propagator has the form

$$\hat{U}(\tau,\tau') = \mathrm{e}^{-(\hat{H}-\mu\hat{N})(\tau-\tau')/\hbar} \tag{R.13}$$

$\tau = 0$ corresponds to the high temperature limit.

The expectation of an operator $\hat{A}$ in a finite temperature ensemble has the simple form

$$\Big\langle A\Big\rangle_{T,\mu} = \frac{1}{\sum_n \mathrm{e}^{-\beta E_n}}\sum_n \langle\psi_n|\hat{A}|\psi_n\rangle\mathrm{e}^{-\beta E_n} = \frac{\sum_n\langle\Psi_n|\hat{A}\mathrm{e}^{-\beta(\hat{H}-\mu\hat{N})}|\Psi_n\rangle}{\sum_n\langle\Psi_n|\mathrm{e}^{-\beta(\hat{H}-\mu\hat{N})}|\Psi_n\rangle} = \frac{Tr[\hat{A}\hat{U}(\beta\hbar,0)]}{Tr[\hat{U}(\beta\hbar,0)]}$$

The partition sum Z is the denominator, which is linked to the grand potential Ω via $Z = \mathrm{e}^{-\beta\Omega}$. Thus, the grand potential is directly linked to the propagator in imaginary time

$$\Omega(T,\mu) = -k_BT\ln[Z(T,\mu)] = -k_BT\ln\Big(Tr[\hat{U}_\mu(\beta\hbar,0)]\Big)$$

Propagator in the interaction picture

For the sake of obtaining a diagrammatic expansion analogous to that in real time, one introduces an interaction picture in the imaginary-time formalism.

Again, the Hamilton operator is split into a non-interacting $\hat{H}_0$ and an interaction $\hat{W}$.

A state $|\psi_I(\tau)\rangle$ in the interaction picture is related to a state $|\psi_S(\tau)\rangle$ in the imaginary-time Schrödinger picture is

$$|\psi_I(\tau)\rangle = \mathrm{e}^{(\hat{H}_0-\mu\hat{N})\tau/\hbar}|\psi_S(\tau)\rangle$$

This transforms the Schrödinger equation into

$$\begin{aligned}
-\hbar\partial_\tau|\psi_I(\tau)\rangle &= -(\hat{H}_0-\mu\hat{N})|\psi_I(\tau)\rangle + \mathrm{e}^{(\hat{H}_0-\mu\hat{N})\tau/\hbar}(-\hbar\partial_\tau)|\psi_S(\tau)\rangle \\
&= -(\hat{H}_0-\mu\hat{N})|\psi_I(\tau)\rangle + \mathrm{e}^{(\hat{H}_0-\mu\hat{N})\tau/\hbar}(\hat{H}_0+\hat{W}-\mu\hat{N})|\psi_S(\tau)\rangle \\
&= \Big[-\mathrm{e}^{(\hat{H}_0-\mu\hat{N})\tau/\hbar}(\hat{H}_0-\mu\hat{N})\mathrm{e}^{-(\hat{H}_0-\mu\hat{N})\tau/\hbar} + \mathrm{e}^{(\hat{H}_0-\mu\hat{N})\tau/\hbar}\Big(\hat{H}_0+\hat{W}-\mu\hat{N}\Big)\mathrm{e}^{-(\hat{H}_0-\mu\hat{N})\tau/\hbar}\Big]|\psi_I(\tau)\rangle \\
&= \mathrm{e}^{(\hat{H}_0-\mu\hat{N})\tau/\hbar}\hat{W}\mathrm{e}^{-(\hat{H}_0-\mu\hat{N})\tau/\hbar}|\psi_I(\tau)\rangle \\
&= \hat{W}_I(\tau)|\psi_I(\tau)\rangle
\end{aligned}$$

where $\hat{W}_I(\tau)$ is the interaction operator in the interaction picture

$$\hat{W}_I(\tau) = \mathrm{e}^{(\hat{H}_0-\mu\hat{N})\tau/\hbar}\hat{W}_S\mathrm{e}^{-(\hat{H}_0-\mu\hat{N})\tau/\hbar}$$

Now we can use the expression for the time ordered exponential to obtain a solution for the propagator in the interaction picture, that is

$$|\Psi_I(\tau_2)\rangle = \mathcal{T}_D\Big[\mathrm{e}^{-\int_{\tau_1}^{\tau_2} d\tau'\,\hat{W}_I(\tau')}\Big]|\Psi_I(\tau_1)\rangle$$

$$|\Psi_S(\tau_2)\rangle = \mathrm{e}^{-(\hat{H}_0-\mu\hat{N})\tau_2/\hbar}\mathcal{T}_D\Big[\mathrm{e}^{-\frac{1}{\hbar}\int_{\tau_1}^{\tau_2} d\tau'\,\hat{W}_I(\tau')}\Big]\mathrm{e}^{(\hat{H}_0-\mu\hat{N})\tau_1/\hbar}|\Psi_S(\tau_1)\rangle$$

Thus, the propagator for Schrödinger states can be written in two forms, namely as

PROPAGATOR IN IMAGINARY TIME

$$\hat{U}(\beta\hbar, 0) = \mathrm{e}^{-\beta\left(\hat{H}_0+\hat{W}_S-\mu\hat{N}\right)} = \mathrm{e}^{-\beta(\hat{H}_0-\mu\hat{N})}\mathcal{T}_D\left[\mathrm{e}^{-\frac{1}{\hbar}\int_0^{\beta\hbar} d\tau'\,\hat{W}_I(\tau')}\right] \tag{R.14}$$

$$= \mathrm{e}^{-\beta(\hat{H}_0-\mu\hat{N})}\mathcal{T}_D\left[\exp\left(-\frac{1}{\hbar}\int_0^{\beta\hbar} d\tau\;\mathrm{e}^{(\hat{H}_0-\mu\hat{N})\tau/\hbar}\hat{W}_S\mathrm{e}^{-(\hat{H}_0-\mu\hat{N})\tau/\hbar}\right)\right] \tag{R.15}$$

Appendix S

Field operators in the interaction picture

FIELD OPERATORS IN THE INTERACTION PICTURE

If the non-interacting Hamiltonian is is $\hat{H}_0 = \sum_{\alpha,\beta} h_{\alpha,\beta}\hat{c}^\dagger_\alpha \hat{c}_\beta$, the field operators in the interaction picture can be written in the form

$$\begin{aligned}\hat{c}_{\alpha,I}(t) &\overset{\text{Eq. S.5}}{=} \sum_\beta \left(\mathrm{e}^{-\frac{i}{\hbar}\boldsymbol{h}t}\right)_{\alpha,\beta}\hat{c}_\beta && \text{(S.1)}\\ \hat{c}^\dagger_{n,I}(t) &= \sum_\beta \left(\mathrm{e}^{+\frac{i}{\hbar}\boldsymbol{h}t}\right)_{\alpha,\beta}\hat{c}^\dagger_\beta && \text{(S.2)}\end{aligned}$$

If the non-interacting Hamiltonian $\hat{H}_0$ is expressed in its diagonal form, that is $\hat{H}_0 = \sum_n \epsilon_n \hat{c}^\dagger_n \hat{c}_n$, the field operators in the interaction picture can be written in the form

$$\begin{aligned}\hat{c}_{n,I}(t) &= \mathrm{e}^{-\frac{i}{\hbar}\epsilon_n t}\hat{c}_n && \text{(S.3)}\\ \hat{c}^\dagger_{n,I}(t) &= \mathrm{e}^{+\frac{i}{\hbar}\epsilon_n t}\hat{c}^\dagger_n && \text{(S.4)}\end{aligned}$$

Proof for the annihilator using the Baker-Hausdorff Theorem

We start with the operator

$$\hat{c}_{\zeta,I} = \mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t}\hat{c}_{\zeta,S}\mathrm{e}^{-\frac{i}{\hbar}\hat{H}_0 t}$$

with $\hat{H}_0 = \sum_{\alpha,\beta} h_{\alpha,\beta}\hat{c}^\dagger_\alpha \hat{c}_\beta$. where $\hat{c}_{\zeta,I}$ is the annihilator in the interaction picture, and $\hat{c}_{\zeta,S} = \hat{c}_\zeta$ is the same operator in the Schrödinger picture.

With the help of the Baker-Hausdorff theorem Eq. T.3 we can rewrite this expression as

$$\mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t}\hat{c}_\zeta \mathrm{e}^{-\frac{i}{\hbar}\hat{H}_0 t} = \sum_{n=0}^{\infty}\left(\frac{i}{\hbar}t\right)^n \hat{X}_n \qquad \text{with} \qquad \hat{X}_n = \frac{1}{n+1}[\hat{H}_0, \hat{X}_n]_- \qquad \text{and} \qquad \hat{X}_0 = \hat{c}_\zeta$$

Let us now work out the terms $\hat{X}_n$

$$
\begin{aligned}
\hat{X}_0 &= \hat{c}_\zeta \\
\hat{X}_1 &\overset{Eq.\ T.3}{=} \left[\sum_{\alpha,\beta} h_{\alpha,\beta}\hat{c}^\dagger_\alpha\hat{c}_\beta,\, \hat{c}_\zeta\right]_- = \sum_{\alpha,\beta} h_{\alpha,\beta}\left(\hat{c}^\dagger_\alpha\hat{c}_\beta\hat{c}_\zeta - \hat{c}_\zeta\hat{c}^\dagger_\alpha\hat{c}_\beta\right) \\
&= \sum_{\alpha,\beta} h_{\alpha,\beta}\left(\hat{c}^\dagger_\alpha \underbrace{[\hat{c}_\beta, \hat{c}_\zeta]_+}_{=0} - \hat{c}^\dagger_\alpha\hat{c}_\zeta\hat{c}_\beta - \underbrace{[\hat{c}_\zeta, \hat{c}^\dagger_\alpha]_+}_{\delta_{\alpha,\zeta}}\hat{c}_\beta + \hat{c}^\dagger_\alpha\hat{c}_\zeta\hat{c}_\beta\right) \\
&= -\sum_\beta h_{\zeta,\beta}\hat{c}_\beta \\
\hat{X}_2 &\overset{Eq.\ T.3}{=} \frac{1}{2}[\underbrace{\sum_{\alpha,\beta} h_{\alpha,\beta}\hat{c}^\dagger_\alpha\hat{c}_\beta}_{\hat{H}_0},\ \underbrace{-\sum_\gamma h_{\zeta,\gamma}\hat{c}_\gamma}_{\hat{X}_1}]_- = -\frac{1}{2}\sum_\gamma h_{\zeta,\gamma}[\sum_{\alpha,\beta} h_{\alpha,\beta}\hat{c}^\dagger_\alpha\hat{c}_\beta,\ \hat{c}_\gamma]_- \\
&= -\frac{1}{2}\sum_\gamma h_{\zeta,\gamma}\Big(-\sum_\beta h_{\gamma,\beta}\hat{c}_\gamma\Big) = \frac{1}{2!}\Big(-\boldsymbol{h}\Big)^2_{\zeta,\beta}\hat{c}_\gamma \\
\hat{X}_n &= \frac{1}{n!}\Big(-\boldsymbol{h}\Big)^n_{\zeta,\gamma}\hat{c}_\gamma
\end{aligned}
$$

Thus, we obtain

$$
\hat{c}_{\alpha,I} = \mathrm{e}^{\frac{i}{\hbar}\hat{H}_0 t}\hat{c}_\alpha\mathrm{e}^{-\frac{i}{\hbar}\hat{H}_0 t} = \sum_{n=0}^{\infty}\frac{1}{n!}\sum_\beta(-\frac{i}{\hbar}\boldsymbol{h}t)^n_{\alpha,\beta}\hat{c}_\beta = \sum_\beta\Big(\mathrm{e}^{-\frac{i}{\hbar}\boldsymbol{h}t}\Big)_{\alpha,\beta}\hat{c}_\beta \tag{S.5}
$$

which proves Eq. S.1.

In the special case that the Hamilton matrix $\boldsymbol{h}$ is diagonal, so that $h_{\alpha,\beta} = \delta_{\alpha,\beta}\epsilon_\alpha$, we obtain

$$
\hat{c}_{\alpha,I} = \mathrm{e}^{-\frac{i}{\hbar}\epsilon_\alpha t}\hat{c}_\alpha
$$

which proves Eq. S.3.

Proof for the creator by hermitian conjugation

For the creator we exploit that creators and annihilators at equal times are hermitian conjugate to each other also in the interaction picture.

$$
\begin{aligned}
\hat{c}^\dagger_{\alpha,I}(t) &= \hat{U}^{(0)}(0,t)\hat{c}^\dagger_\alpha\hat{U}^{(0)}(t,0) = \Big(\hat{U}^{(0),\dagger}(t,0)\hat{c}_\alpha\hat{U}^{(0)\dagger}(0,t)\Big)^\dagger = \Big(\hat{U}^{(0)}(0,t)\hat{c}_\alpha\hat{U}^{(0)}(t,0)\Big)^\dagger \\
&= (\hat{c}_{\alpha,I})^\dagger
\end{aligned} \tag{S.6}
$$

This allows us to write directly

$$
\hat{c}^\dagger_{\alpha,I}(t) = \sum_\beta\Big(\mathrm{e}^{+\frac{i}{\hbar}\boldsymbol{h}t}\Big)_{\alpha,\beta}\hat{c}^\dagger_\beta \tag{S.7}
$$

which proves Eq. S.2, and for a one-particle basis, that diagonalizes the one-particle Hamiltonian, we obtain

$$
\hat{c}^\dagger_{\alpha,I}(t) = \mathrm{e}^{\frac{i}{\hbar}\epsilon_\alpha t}\hat{c}^\dagger_\alpha \tag{S.8}
$$

which proves Eq. S.4.

Appendix T

Baker-Hausdorff Theorem

The Baker-Hausdorff theorem [1][66, 67, 68, 69, 70, 71] provides a method to evaluate the product of two exponentials of non-commutating operators[98]. Specifically we look for the resulting operator $\hat{C}$ of

$$\mathrm{e}^{\hat{A}}\mathrm{e}^{\hat{B}} = \mathrm{e}^{\hat{C}}$$

for two given operators A, B.

Here, we will only investigate a few special cases.

T.1 Baker-Hausdorff Theorem

This is a specialization of the Baker Haussdorf theorem, which applies if certain commutator relations hold.

BAKER-HAUSDORFF THEOREM

$$\mathrm{e}^{\hat{A}+\hat{B}} = \mathrm{e}^{\hat{A}}\mathrm{e}^{\hat{B}}\mathrm{e}^{-\frac{1}{2}[\hat{A},\hat{B}]_-} \tag{T.1}$$

if both operators commutate with their commutator, i.e.

$$[\hat{A},[\hat{A},\hat{B}]] = [\hat{B},[\hat{A},\hat{B}]] = 0 \tag{T.2}$$

Proof:[2]

$$\begin{aligned}
F(\alpha) &\overset{\text{def}}{=} \mathrm{e}^{\alpha\hat{A}+\beta\hat{B}} \\
\frac{dF}{d\alpha} &= \frac{d}{d\alpha}\sum_{n=0}^{\infty}\frac{1}{n!}(\alpha\hat{A}+\beta\hat{B})^n \\
&= \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{j=0}^{n-1}(\alpha\hat{A}+\beta\hat{B})^j\hat{A}(\alpha\hat{A}+\beta\hat{B})^{n-j-1}
\end{aligned}$$

---

[1] Felix Hausdorff, 1868-1942 German mathematician. Professor in Leipzig, Bonn, Greifswald, Bonn. Is considered as co-founder of the modern topology. He invented the Hausdorff dimension, which is used to characterize fractals. The corresponding work is the most cited original paper in mathematics between 1910 and 1920. The Jewish Hausdorff, his wife and her sister committed suicide as they about to be deported into a concentration camp under the NS Dictatorship.

[2] from *Cavity Quantum Electrodynamics: The Strange Theory of Light in a Box*, S.M. Dutra, (John Wiley, 2005)

Let us now analyze the individual terms:

$$
\begin{aligned}
(\alpha\hat{A}+\beta\hat{B})\hat{A} &= \hat{A}(\alpha\hat{A}+\beta\hat{B})+\beta[\hat{B},\hat{A}]_- \\
(\alpha\hat{A}+\beta\hat{B})^k\hat{A} &= (\alpha\hat{A}+\beta\hat{B})^{k-1}\hat{A}(\alpha\hat{A}+\beta\hat{B})+(\alpha\hat{A}+\beta\hat{B})^{k-1}\beta[\hat{B},\hat{A}]_- \\
&\overset{\text{Eq. T.2}}{=} (\alpha\hat{A}+\beta\hat{B})^{k-1}\hat{A}(\alpha\hat{A}+\beta\hat{B})+\beta[\hat{B},\hat{A}]_-(\alpha\hat{A}+\beta\hat{B})^{k-1} \\
&= \hat{A}(\alpha\hat{A}+\beta\hat{B})^k+k\beta[\hat{B},\hat{A}]_-(\alpha\hat{A}+\beta\hat{B})^{k-1}
\end{aligned}
$$

$$
\begin{aligned}
\frac{dF}{d\alpha} &= \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{j=0}^{n-1}(\alpha\hat{A}+\beta\hat{B})^j\hat{A}(\alpha\hat{A}+\beta\hat{B})^{n-j-1} \\
&= \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{j=0}^{n-1}\Big(\hat{A}(\alpha\hat{A}+\beta\hat{B})^{n-1}+j\beta[\hat{B},\hat{A}]_-(\alpha\hat{A}+\beta\hat{B})^{n-2}\Big) \\
&= \sum_{n=0}^{\infty}\frac{1}{n!}\left(n\hat{A}(\alpha\hat{A}+\beta\hat{B})^{n-1}+\frac{n(n-1)}{2}\beta[\hat{B},\hat{A}]_-(\alpha\hat{A}+\beta\hat{B})^{n-2}\right) \\
&= \hat{A}\sum_{n=1}^{\infty}\frac{1}{(n-1)!}(\alpha\hat{A}+\beta\hat{B})^{n-1}+\frac{1}{2}\beta[\hat{B},\hat{A}]_-\sum_{n=2}^{\infty}\frac{1}{(n-2)!}(\alpha\hat{A}+\beta\hat{B})^{n-2} \\
&= \left(\hat{A}+\frac{1}{2}\beta[\hat{B},\hat{A}]_-\right)F(\alpha)
\end{aligned}
$$

This is a differential equation for $F(\alpha)$, which can be solve as if the operator were a number because only one operator, namely $\hat{A}+\frac{1}{2}\beta[\hat{B},\hat{A}]_-$ is involved.

We obtain

$$
\begin{aligned}
&F(\alpha)=F(0)\mathrm{e}^{\alpha\left(\hat{A}+\frac{1}{2}\beta[\hat{B},\hat{A}]_-\right)}=\mathrm{e}^{\beta\hat{B}}\left(\mathrm{e}^{\alpha\hat{A}}\mathrm{e}^{\frac{1}{2}\alpha\beta[\hat{B},\hat{A}]_-}\right) \\
\Rightarrow\quad &\mathrm{e}^{\alpha\hat{A}+\beta\hat{B}}=\mathrm{e}^{\beta\hat{B}}\mathrm{e}^{\alpha\hat{A}}\mathrm{e}^{\frac{1}{2}\alpha\beta[\hat{B},\hat{A}]_-}=\mathrm{e}^{\beta\hat{B}}\mathrm{e}^{\alpha\hat{A}}\mathrm{e}^{-\frac{1}{2}\alpha\beta[\hat{A},\hat{B}]_-}\qquad\text{q.e.d.}
\end{aligned}
$$

In the last step we exploited that $\hat{A}$ commutates with the commutator $[\hat{A},\hat{B}]_-$, so that we can disentangle the exponential as if the operators were numbers.

T.2 Hadamard Lemma

The Hadamard Lemma is a specialization of Zassenhaus formula[98], which again is a variant of the Baker-Campbell-Hausdorff theorem is useful to work out simple operators in the interaction picture.

BAKER-CAMPBELL-HAUSDORFF THEOREM

$$
\begin{aligned}
\mathrm{e}^{\lambda\hat{A}}\hat{B}\mathrm{e}^{-\lambda\hat{A}} &\overset{Eq.\ T.7}{=} \sum_{n=0}^{\infty}\hat{X}_n\lambda^n \\
\text{with}\quad \hat{X}_{n+1} &= \frac{1}{n+1}[\hat{A},\hat{X}_n]_- \quad\text{and}\quad \hat{X}_0=\hat{B}
\end{aligned}
\tag{T.3}
$$

$$
\begin{aligned}
\mathrm{e}^{\lambda\hat{A}}\hat{B}\mathrm{e}^{-\lambda\hat{A}} &= \left(\sum_{i=0}^{\infty}\frac{1}{i!}(\lambda\hat{A})^i\right)\hat{B}\left(\sum_{j=0}^{\infty}\frac{1}{j!}(-\lambda\hat{A})^j\right) &\text{(T.4)}\\
&= \sum_{n=0}^{\infty}\lambda^n\underbrace{\left(\sum_{j=0}^{n}\frac{(-1)^j}{(n-j)!j!}\hat{A}^{n-j}\hat{B}\hat{A}^j\right)}_{X_n} &\text{(T.5)}
\end{aligned}
$$

Let us now work out the expression denoted $\hat{X}_n$ in Eq. T.5. $\hat{X}_0 = \hat{B}$ can be determined directly. The higher $\hat{X}_n$ are obtained by recursion.

$$
\begin{aligned}
[\hat{A},\hat{X}_n]_- &= \sum_{j=0}^{n}\frac{(-1)^j}{(n-j)!j!}\hat{A}^{n+1-j}\hat{B}\hat{A}^j - \sum_{j=0}^{n}\frac{(-1)^j}{(n-j)!j!}\hat{A}^{n-j}\hat{B}\hat{A}^{j+1}\\
&= \sum_{j=0}^{n}\frac{(n+1-j)(-1)^j}{(n+1-j)!j!}\hat{A}^{n+1-j}\hat{B}\hat{A}^j + \sum_{j=0}^{n}\frac{(j+1)(-1)^{j+1}}{(n+1-(j+1))!(j+1)!}\hat{A}^{n+1-(j+1))}\hat{B}\hat{A}^{j+1}\\
&\overset{j+1\to j}{=} \sum_{j=0}^{n}\frac{(n+1-j)(-1)^j}{(n+1-j)!j!}\hat{A}^{n+1-j}\hat{B}\hat{A}^j + \sum_{k=1}^{n+1}\frac{j(-1)^j}{(n+1-j)!j!}\hat{A}^{n+1-j))}\hat{B}\hat{A}^j\\
&= (n+1)\sum_{j=0}^{n+1}(\frac{(-1)^j}{(n+1-j)!j!}\hat{A}^{n+1-j}\hat{B}\hat{A}^j\\
&= (n+1)\hat{X}_{n+1} &\text{(T.6)}
\end{aligned}
$$

Thus, we obtain the Baker-Campbell-Hausdorff Theorem saying

$$
\begin{aligned}
&\mathrm{e}^{\lambda\hat{A}}\hat{B}\mathrm{e}^{-\lambda\hat{A}} \overset{Eq.\ T.5}{=} \sum_{n=0}^{\infty}\hat{X}_n\lambda^n\\
\text{with}\quad &\hat{X}_{n+1} \overset{Eq.\ T.6}{=} \frac{1}{n+1}[\hat{A},\hat{X}_n]_- \quad\text{and}\quad \hat{X}_0 = \hat{B} &\text{(T.7)}
\end{aligned}
$$

Appendix U

Stuff

This is not for readers. I collect here uncompleted stuff that may be used later in the script.

U.1 Collection of links

- great web site on point group symmetries `http://www.phys.ncl.ac.uk/staff/njpg/symmetry/index.html`

U.2 Electrons and holes

The one-particle Green's function describes the propagation of a single particle. Now we want to generalize this concept to many-particle physics. We are not interested here in the dynamics of all particles but of a single particle in a many-particle system. The first step to such a description is to generalize the concept to a system of non-interacting particles. We can actually reformulate the problem of many electrons by excitations of electrons and holes in a Fermi sea. The **Fermi sea** contains all one-particle states that are occupied in the ground state.

Let us consider the many-particle Hamiltonian for non-interacting electrons. The Fermi sea is filled up to the Fermi level μ. We will show that we can describe either all the electrons, or we can consider only electrons and holes.

$$
\begin{aligned}
\hat{H} &= \sum_n \epsilon_n \hat{c}_n^\dagger \hat{c}_n \\
&\overset{[\hat{c}_n,\hat{c}_n^\dagger]_+=1}{=} \sum_{n;\epsilon_n<\epsilon_f} \epsilon_n \left[1-\hat{c}_n\hat{c}_n^\dagger\right] + \sum_{n;\epsilon_n>\epsilon_f} \epsilon_n \hat{c}_n^\dagger \hat{c}_n \\
&= \underbrace{\sum_{n;\epsilon_n<\epsilon_f} \epsilon_n}_{\text{energy of Fermi sea}} + \underbrace{\sum_{n;\epsilon_n>\epsilon_f} \epsilon_n \hat{c}_n^\dagger \hat{c}_n}_{\text{energy of electrons}} + \underbrace{\sum_{n;\epsilon_n<\epsilon_f} \Big(-\epsilon_n\Big) \hat{c}_n \hat{c}_n^\dagger}_{\text{energy of holes}}
\end{aligned}
$$

Thus, the Hamilton operator can be represented by terms for electrons and holes plus an additional constant that represents the energy $E_{\mathcal{G}}$ of the Fermi sea. When we talk in the present context of electrons and holes, electrons always mean electrons in states above the Fermi level and holes are electrons removed from the Fermi sea below the Fermi level.

It is convenient to express the energies of electron and hole excitations relative to the Fermi level

μ. Therefore we rewrite the Hamiltonian in the form

$$\hat{H} = E_{\mathcal{G}} + \mu \underbrace{\left(\overbrace{\sum_{n;\epsilon_n>\mu} \hat{c}_n^\dagger \hat{c}_n}^{\text{electrons}} - \overbrace{\sum_{n;\epsilon_n<\mu} \hat{c}_n \hat{c}_n^\dagger}^{\text{holes}} \right)}_{\text{net number of electrons added to Fermi sea}} + \underbrace{\sum_{n;\epsilon_n<\mu} (\mu - \epsilon_n) \hat{c}_n \hat{c}_n^\dagger}_{\text{energy of holes}} + \underbrace{\sum_{n;\epsilon_n>\mu} (\epsilon_n - \mu) \hat{c}_n^\dagger \hat{c}_n}_{\text{energies of electrons}} \quad \text{(U.1)}$$

In this picture we determine first the energy of the Fermi sea and the that of any surplus electrons or holes. The surplus electrons are first placed into the reservoir. The reservoir is characterized by a chemical potential μ. An electron in the reservoir has the energy μ and every hole in the reservoir has the energy $-\mu$. The remaining two terms is the energy required to move electrons and holes out of the reservoir into the system. Therefore, it makes sense to speak of electron excitations and hole excitations.

The ground-state energy of an "empty"[1] reservoir is set to zero. Thus, the ground-state energy of the Fermi sea combined with the reservoir is $E_{\mathcal{G}} - \mu N_{\mathcal{G}}$. $N_{\mathcal{G}}$ is the number of electrons in the Fermi sea. $E_{\mathcal{G}} - \mu N_{\mathcal{G}}$ is the energy difference between a state, where all electrons are ion the reservoir and the state where $N_{\mathcal{G}}$ electrons have been moved from the reservoir into the system where they form the Fermi sea.

When an electron is moved from the reservoir into the system, the energy required is the difference of the electron energy in the system and in the reservoir, namely $\epsilon_n - \mu$. Moving an electron into the reservoir, that is the formation of a hole requires the energy $\mu - \epsilon$. The Hamiltonian describing the energy difference to the ground state of reservoir and Fermi sea is

$$\hat{H} - \mu \hat{N} - \Big(E_{\mathcal{G}} - \mu N_{\mathcal{G}} \Big) = \underbrace{\sum_{n;\epsilon_n<\mu} (\mu - \epsilon_n) \hat{c}_n \hat{c}_n^\dagger}_{\text{energy of holes}} + \underbrace{\sum_{n;\epsilon_n>\mu} (\epsilon_n - \mu) \hat{c}_n^\dagger \hat{c}_n}_{\text{energies of electrons}} \quad \text{(U.2)}$$

In nearly all applications of many-particle quantum mechanics, the ground-state energy is ignored, because it does not affect the dynamics or the response of the system towards perturbations. One has to be aware that this is a particular choice for the energy reference.

This point of view and its effect on electron and hole excitations is demonstrated in Fig. U.1.

---

[1] Even an empty particle reservoir can provide electrons to the system. A reservoir is like a bank that gives credit.

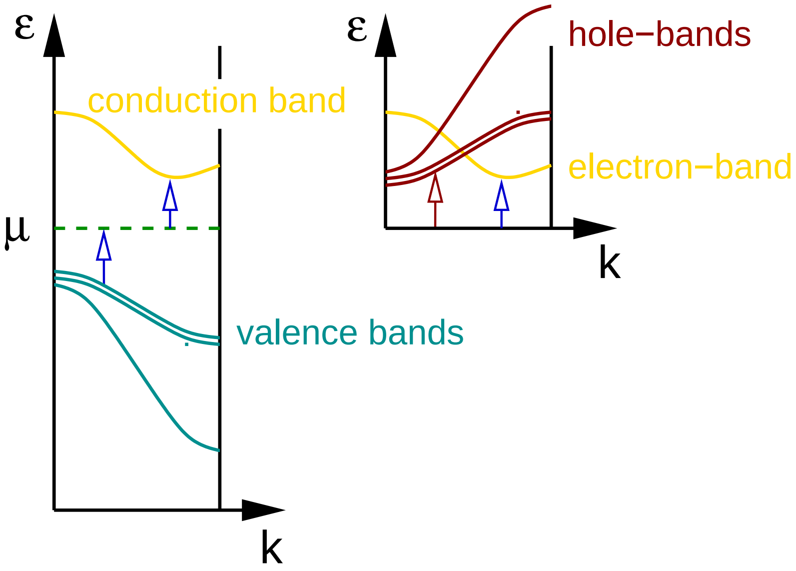


Fig. U.1: Instead of considering the energy of all electrons (left), we may focus only on the changes (right) from the ground state. The ground state is the Fermi sea combined with a particle reservoir. Instead of electrons, we consider electron- and hole-excitations. As they are excitations of the ground state they have positive energies. The creation of a hole corresponds now to lifting a hole from the Fermi level, into one of the hole bands. The hole bands are the mirror images (at μ) of the valence electron bands.

Appendix V

Dictionary

caveat	Vorbehalt
coincidence	Übereinstimmung
convolution	Faltung
cross section	Wirkungsquerschnitt
esthetics	Ästetik
factorial	Fakultät (mathematische Operation)
faculty	Fakultät (akademische Verwaltungseinheit)
for the sake of ...	um ... willen (gen)
formidable	gewaltig, eindrucksvoll
propagator	Zeitentwicklungsoperator
commutate	kommutieren
rank	Rang
scattering amplitude	Streuamplitude
without loss of generality	ohne Beschränkung der Allgemeinheit
wlog	oBdA (ohne Beschränkung der Allgemeinheit)

V.1 Explanations

- caveat: Latin, a modifying or cautionary detail to be considered

V.2 Symbols

- $\forall$ "for all".
- $[A, B]_- = [A, B] = AB - BA$ **commutator** of A and B. A and B may be matrices or operators.
- $[A, B]_+ = AB + BA$ **anti-commutator** of A and B. A and B may be matrices or operators.
- $\vec{a} \times \vec{b}$ **vector product** between two three dimensional vectors $\vec{a}$ and $\vec{b}$.

$$\vec{a} \times \vec{b} = \begin{pmatrix} a_y b_z - a_z b_y \\ a_z b_x - a_x b_z \\ a_x b_y - a_y b_x \end{pmatrix} \tag{V.1}$$

- $\vec{a} \otimes \vec{b}$ **outer product** or **dydadic product** of two vectors $\vec{a}$ and $\vec{b}$.

$$\left(\vec{a} \otimes \vec{b}\right)_{i,j} = a_i b_j \tag{V.2}$$

- $\det|\boldsymbol{A}|$ **determinant** of matrix $\boldsymbol{A}$
- $\mathrm{perm}|\boldsymbol{A}|$ **permanent** of matrix $\boldsymbol{A}$. The permanent is computed like the determinant with the exceptions that all terms are summed up without the sign changes for permutations.
- $\delta_{i,j}$ **Kronecker delta**

$$\delta_{i,j} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{else} \end{cases} \tag{V.3}$$

- $\delta(\vec{x})$ **Dirac's delta function** defined by

$$\forall_{f(x)} \int dx\, f(x)\delta(x) = f(0) \tag{V.4}$$

  which holds for all differentiable functions $f(x)$. The delta function of a vector is the product of the delta functions of its components.

- $\epsilon_{i_1,\ldots,i_N}$ **Levy-Civita Symbol** or **fully antisymmetric tensor**

Appendix W

Greek Alphabet

A	α	alpha	N	ν	nu
B	β	beta	Ξ	ξ	ksi
Γ	γ	gamma	O	$o,$	omicron
Δ	δ	delta	Π	π, ϖ	pi
E	ϵ, ε	epsilon	P	ρ, ϱ	rho
Z	ζ	zeta	Σ	σ, ς	sigma
H	η	eta	T	τ	tau
Θ	θ, ϑ	theta	Υ	υ	upsilon
I	ι	iota	Φ	ϕ, φ	phi
K	κ	kappa	X	χ	chi
Λ	λ	lambda	Ψ	ϕ	phi
M	μ	mu	Ω	ω	omega

Appendix X

Philosophy of the ΦSX Series

In the ΦSX series, I tried to implement what I learned from the feedback given by the students which attended the courses and that relied on these books as background material.

The course should be self-contained. There should not be any statements "as shown easily..." if, this is not true. The reader should not need to rely on the author, but he should be able to convince himself, if what is said is true. I am trying to be as complete as possible in covering all material that is required. The basis is the mathematical knowledge. With few exceptions, the material is also developed in a sequence so that the material covered can be understood entirely from the knowledge covered earlier.

The derivations shall be explicit. The novice should be able to step through every single step of the derivation with reasonable effort. An advanced reader should be able to follow every step of the derivations even without paper and pencil.

All units are explicit. That is, formulas contain all fundamental variables, which can be inserted in any desirable unit system. Expressions are consistent with the SI system, even though I am quoting some final results in units, that are common in the field.

The equations that enter a specific step of a derivation are noted ontop of the equation sign. The experience is that the novice does not immediately memorize all the material covered and that he is struggling with the math, so that he spends a lot of time finding the rationale behind a certain step. This time is saved by being explicit about it. The danger that the student gets dependent on these indications, is probably minor, as it requires some effort for the advanced reader to look up the assumptions, an effort he can save by memorizing the relevant material.

Important results and equations are highlighted by including them in boxes. This should facilitate the preparations for examinations.

Portraits of the key researchers and short biographical notes provide independent associations to the material. A student may not memorize a certain formula directly, but a portrait. From the portrait, he may associate the correct formula. The historical context provides furthermore an independent structure to organize the material.

The two first books are in german (That is the intended native language) in order to not add complications to the novice. After these first books, all material is in English. It is mandatory that the student masters this language. Most of the scientific literature is available only in English. English is currently the language of science, and science is absolutely dependent on international contacts.

I tried to include many graphs and figures. The student shall become used to use all his senses in particular the visual sense.

I have slightly modified the selection of the material commonly tought in most courses. Some topics, which I consider of mostly historical relevance I have removed. Others such as the Noether theorem, I have added. Some, like Chaos, Stochastic processes, etc. I have not added yet.

Appendix Y

About the Author

Prof. Dr. rer. nat Peter E. Blöchl studied physics at Karlsruhe University of Technology in Germany. Subsequently he joined the Max Planck Institutes for Materials Research and for Solid state Research in Stuttgart, where he worked on development of electronic-structure methods related to the LMTO method and on first-principles investigations of interfaces. He received his doctoral degree in 1989 from the University of Stuttgart.

Following his graduation, he joined the renowned T.J. Watson Research Center in Yorktown Heights, NY in the US on a World Trade Fellowship. In 1990 he accepted an offer from the IBM Zurich Research Laboratory in Ruschlikon, Switzerland, which had just received two Nobel prices in Physics (For the Scanning Tunneling Microscope in 1986 and for the High-Temperature Superconductivity in 1987). He spent the summer term 1995 as visiting professor at the Vienna University of Technology in Austria, from where was awarded the habilitation in 1997. In 2000 he left the IBM Research Laboratory after a 10-year period and accepted an offer to be professor for theoretical physics at Clausthal University of Technology in Germany. Since 2003, Prof. Blöchl is member of the Braunschweigische Wissenschaftliche Gesellschaft (Academy of Sciences).

The main thrust of Prof. Blöchl's research is related to ab-initio simulations, that is, parameter-free simulation of materials processes and molecular reactions based on quantum mechanics. He developed the Projector Augmented Wave (PAW) method, one of the most widely used electronic structure methods to date. This work has been cited over 15,000 times. It is among the 100 most cited scientific papers of all times and disciplines, and is among the 10 most-cited papers out of more than 500,000 published in the 120-year history of Physical Review. Next to the research related to simulation methodology, his research covers a wide area from biochemistry, solid state chemistry to solid state physics and materials science. Prof. Bloechl contributed to 8 Patents and published about 100 research publications, among others in well-known Journals such as "Nature". The work of Prof. Blöchl has been cited over 23,000 times, and he has an H-index of 41.

Bibliography

[1] W.-D. Schöne, C. Timm, and W.-D. Schotte. Theoretische festkörperphysik, teil 1. http://kolxo3.tiera.ru/_Papers/TU%20Wien%20Scripta/Physik/Festkoerperphysik/Theoretische_Fest 2001.

[2] A.L. Fetter and J.D. Walecka. Quantum Theory of Many-Particle Systems. McGraw-Hill, New York, 1971. ISBN 978-0-07-020653-3. URL http://store.doverpublications.com/0486428273.html.

[3] W. Nolting. Grundkurs Theoretische Physik 7: Viel-Teilchen Theorie. Springer, 2002.

[4] R.D. Mattuck. A Guide to Feynman Diagrams in the Many-Body Problem. McGraw-Hill, 1967.

[5] H. Eschrig. The Fundamentals of Density Functional Theory. Edition am Gutenbergplatz, Leipzig, 2003.

[6] J.W. Negele and H. Orland. Quantum Many-Particle Physics. Westview Press, 1988.

[7] M. Born and R. Oppenheimer. Zur quantentheorie der molekeln. Ann. Phys., 84:457, 1927.

[8] M. Born and K. Huang. Dynamical Theory of Crystal Lattices. Oxford University Press, Oxford, 1954.

[9] G.A. Worth and M.A. Robb. Applying direct molecular dynamics to non-adiabatic systems. Adv. Chem. Phys., 124:355, 2002.

[10] Felix T. Smith. Diabatic and adiabatic representations for atomic collision problems. Phys. Rev., 179:111–123, Mar 1969. doi: 10.1103/PhysRev.179.111. URL http://link.aps.org/doi/10.1103/PhysRev.179.111.

[11] F. Dufey. Störungsreihen für die nichtadiabatische Kopplung mit durchschneidenden Potentialflächen. PhD thesis, Technische Universität München, 2002.

[12] M. Baer, Á. Viboḱ, G. J. Halaśz, and D. J. Kouri. The electronic non-adiabatic coupling terms: On the connection between molecular physics and field theory. Adv. Quantum Chem., 44:103, 2003.

[13] Y. Aharonov and D. Bohm. Significance of electromagnetic potentials in the quantum theory. Phys. Rev., 115:485–491, Aug 1959. doi: 10.1103/PhysRev.115.485. URL http://link.aps.org/doi/10.1103/PhysRev.115.485.

[14] C. N. Yang and R. L. Mills. Conservation of isotopic spin and isotopic gauge invariance. Phys. Rev., 96:191–195, Oct 1954. doi: 10.1103/PhysRev.96.191. URL http://link.aps.org/doi/10.1103/PhysRev.96.191.

[15] J. v. Neumann and E.P. Wigner. über das verhalten von eigenwerten bei adiabatischen prozessen. Phys. Z, 30:467, 1929.

[16] D.R. Yarkony. Diabolical conical intersections. Rev. Mod. Phys., 68:985, 1996.

[17] J. H. Van Vleck. The jahnteller effect and crystalline stark splitting for clusters of the form xy6. J. Chem. Phys., 7:72, 1939.

[18] L. Butler. Chemical reaction dynamics beyond the born-oppenheimer approximation. Annu. Rev. Phys. Chem, 49:124, 1998.

[19] M. Baer and G.D. Billing, editors. The role of degenerate states in chemistry, volume 124 of Adv. Chem. Phys. John Wiley and Sons, 2002.

[20] M.S. Child. Early perspectives on geometric phase. Adv. Chem. Phys., 124:1, 2002.

[21] Michael Baer. Beyond Born-Oppenheimer: Conical Intersections and Electronic Nonadiabatic Coupling Terms. John Wiley, 2006.

[22] G.A. Worth and L.S. Cederbaum. Beyond born-oppenheimer: molecular dynamics through a conical intersection. Annu. Rev. Phys. Chem, 55:127, 2004. doi: doi:10.1146/annurev.physchem.55.091602.094335.

[23] G. Herzberg and H. C. Longuet-Higgins. Intersection of potential energy surfaces in polyatomic molecules. Discuss. Faraday Soc., 35:77–82, 1963. doi: 10.1039/DF9633500077. URL `http://dx.doi.org/10.1039/DF9633500077`.

[24] M.V. Berry. Quantal phase factors accompanying adiabatic changes. Proc. Roy. Soc. A, 392:45, 1984.

[25] David C. Clary. Geometric phase in chemical reactions. Science, 309:1195, 2005.

[26] Carl D. Anderson. The positive electron. Phys. Rev., 43:491, 1933.

[27] P. E. Blöchl. ΦSX: Theoretical Physics III, Quantum Theory. Blöchl, 2015. URL `http://www2.pt.tu-clausthal.de/atp/phisx.html`.

[28] J.C. Slater. The theory of complex spectra. Phys. Rev., 34:1293, 1929.

[29] J.C. Slater. Cohesion in monovalent metals. Phys. Rev., 35:509, 1930.

[30] P. E. Blöchl. ΦSX: Quantum Mechanics of the Chemical Bond. Blöchl, 2015. URL `http://www2.pt.tu-clausthal.de/atp/phisx.html`.

[31] P. E. Blöchl. ΦSX: Theoretical Physics IV, Statistical Physics. Blöchl, 2015. URL `http://www2.pt.tu-clausthal.de/atp/phisx.html`.

[32] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. Phys. Rev. Lett, 77:3865, 1996.

[33] V. Fock. Näherungsmethode zur lösung des quantenmechanischen mehrkörperproblems. Zeitschr. Phys., 61:126, 1930.

[34] V. Fock. "selfconsistent field" mit austausch für natrium. Zeitschr. Phys., 62:795, 1930.

[35] D.R. Hartree. The wave mechanics of an atom with a non-coulomb central field. part i-theory and methods. Proc. Cam. Phil. Soc., 24:89, 1928. Hartree approximation.

[36] D.R. Hartree. The wave mechanics of an atom with a non-coulomb central field. part ii–some results and discussion. Proc. Cam. Phil. Soc., 24:111, 1928. Hartree approximation part two.

[37] J.P. Perdew and A. Zunger. Self interaction correction to density-functional approximations for many-electron systems. Phys. Rev. B, 23:5048, 1981.

[38] W. Kolos and C.C. Roothaan. Correlated orbitals for the ground state of the hydrogen molecule. Rev. Mod. Phys., 32:205, 1960.

[39] Pierre Hohenberg and Walter Kohn. Inhomogeneous electron gas. Phys. Rev., 136:B864, 1964. doi: 10.1103/PhysRev.136.B864. URL `http://link.aps.org/doi/10.1103/PhysRev.136.B864`.

[40] Walter Kohn and Lu J. Sham. Self-consistent equations including exchange and correlation effects. Phys. Rev., 140:A1133, 1965. doi: 10.1103/PhysRev.140.A1133. URL `http://link.aps.org/doi/10.1103/PhysRev.140.A1133`.

[41] R. McWeeny. Some recent advances in density matrix theory. Rev. Mod. Phys., 32:334, 1960.

[42] A. Puzder, M.Y. Chou, and R.Q. Hood. Exchange and correlation in the si atom: A quantum monte carlo study. Phys. Rev. A, 64:22501, 2001.

[43] Mark S. Hybertsen and Steven G. Louie. Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies. Phys. Rev. B, 34:5390, 1986.

[44] Pierre Hohenberg and Walter Kohn. Inhomogeneous electron gas. Phys. Rev., 136:B864, 1964. doi: 10.1103/PhysRev.136.B864. URL `http://link.aps.org/doi/10.1103/PhysRev.136.B864`.

[45] Mel Levy. Universal variational functionals of electron densities, first order density matrixes and natural spin-orbitals and solution of the v-representability problem. Proc. Nat'l Acad. Sci. USA, 76:6062, 1979. URL `http://www.pnas.org/content/76/12/6062.abstract`.

[46] M. Levy. Electron densities in search of hamiltonians. Phys. Rev. A, 26:1200, 1982.

[47] Elliott H. Lieb. Density functionals for coulomb systems. Int. J. Quantum Chem., 24(3):243–277, 1983. ISSN 1097-461X. doi: 10.1002/qua.560240302. URL `http://dx.doi.org/10.1002/qua.560240302`.

[48] J. Harris and R.O. Jones. The surface energy of a bounded electron gas. J. Phys. F: Met. Phys., 4:1170, 1974.

[49] David C. Langreth and John P. Perdew. The exchange-correlation energy of a metallic surface. Sol. St. Commun., 17:1425, 1975.

[50] Olle Gunnarsson and Bengt I. Lundquist. Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism. Phys. Rev. B, 13:4274, 1976.

[51] K. Schwarz. Optimization of the statistical exchange parameter α for the free atoms h through nb. Phys. Rev. B, 5:2466, 1971.

[52] I. Lindgren and K. Schwarz. Analysis of the electronic exchange in atoms. Phys. Rev. A, 5:542, 1972.

[53] D.M. Ceperley and B.J. Alder. Ground state of the electron gas by a stochastic method. Phys. Rev. Lett., 45:566, 1980.

[54] David Bohm and David Pines. A collective description of electron interactions: Iii. coulomb interactions in a degenerate electron gas. Phys. Rev., 92:609, 1953.

[55] Robert O. Jones and Olle Gunnarsson. The density functional formalism, its applications and prospects. Rev. Mod. Phys., 61:689, 1989.

[56] A. D. Becke. Density-functional exchange energy with correct asymptotic behavior. Phys. Rev. A, 38:3098, 1988.

[57] John P. Perdew, Robert G. Parr, Mel Levy, and Jose L. Balduz. Density-functional theory for fractional particle number: Derivative discontinuities of the energy. Phys. Rev. Lett., 49:1691–1694, Dec 1982. doi: http://dx.doi.org/10.1103/PhysRevLett.49.1691. URL `http://link.aps.org/doi/10.1103/PhysRevLett.49.1691`.

[58] A. Klamt, V. Jonas, T. Bürger, and J.W. Lohrenz. Refinement and paerameterization of cosmo-rs. J. Phys. Chem. A, 102:5074, 1998.

[59] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist. Van der waals density functional for general geometries. Phys. Rev. Lett, 92:246401, 2004.

[60] E.U. Condon. The theory of complex spectra. Phys. Rev., 36:1121, 1930.

[61] Attila Szabo and Neil S. Ostlund. Modern Quantum Chemistry. McGraw-Hill, New York, 1989. ISBN 9780070627390. URL `http://store.doverpublications.com/0486691861.html`.

[62] G.D. Mahan. Many Particle Physics. Kluwer academic Publishers, 2000.

[63] R.P. Feynman. The theory of positrons. Phys. Rev., 76:749, 1949.

[64] R.P. Feynman. Space time approach to quantum electrodynamics. Phys. Rev., 76:769, 1949.

[65] A. A. Abrikosow, L. P. Gorkov, and I. E. Dzyaloshinski. Methods of Quantum Field Theory in Statistical Physics. Prentice-Hall, Englewood Cliffs, NJ, 1964.

[66] J. Campbell. Proc. Lond. Math. Soc., 1897.

[67] J. Campbell. Proc. Lond. Math. Soc., 1898.

[68] H. Baker. Proc.Lond. Math. Soc, 34, 1902. Baker-Campbell-Hausdorff theorem.

[69] H. Baker. Proc.Lond. Math. Soc, 35, 1903. Baker-Campbell-Hausdorff theorem.

[70] H. Baker. Proc.Lond. Math. Soc (Ser.2), 3:24, 1905. Baker-Campbell-Hausdorff theorem.

[71] F. Hausdorff. Ber.Verh. Sächs. Akad. Wiss. Leipzig, 58:19, 1909.

[72] L. G. Molinari. Another proof of gell-mann and low's theorem. J. Math. Phys., 48:52113, 2007.

[73] G.C. Wick. The evaluation of the collision matrix. Phys. Rev., 80:268, 1950.

[74] J. Goldstone. Derivation of the brueckner many-body theory. Proc. Roy. Soc. (London), A239:267, 1957. doi: doi:10.1098/rspa.1957.0037. URL `http://rspa.royalsocietypublishing.org/content/239/1217/267`.

[75] Lars Hedin. New method for calculating the one-particle green's function with application to the electron-gas problem. Phys. Rev., 139:A796–A823, Aug 1965. doi: 10.1103/PhysRev.139.A796. URL `http://link.aps.org/doi/10.1103/PhysRev.139.A796`.

[76] M. Gell-Mann and K.Brueckner. Correlation energy of an electron gas at high density. Phys. Rev., 106:364, 1957.

[77] S. Doniach and E.H. Sondheimer. Green's Functions for Solid State Physicists. The Benjamin/Cummings Publishing Company, INC, Reading, Massachusetts. Reprinted 1998 by Imperial college press, 1974.

[78] Martin C. Gutzwiller. Effect of correlation on the ferromagnetism of transition metals. Phys. Rev. Lett., 10(5):159, 1963. doi: 10.1103/PhysRevLett.10.159.

[79] E.H. Lieb and F.Y. Wu. Absence of mott transition in an exact solution of the short-range, one-band model in one dimension. Phys. Rev. Lett, 20:1445, 1968.

[80] P.W. Anderson. Localized magnetic states in metals. Phys. Rev., 124:41, 1961.

[81] F. Bloch. Zur theorie des ferromagnetismus. Z. Phys., 61:206, 1930.

[82] D. Khomskii. Classifying multiferroics: Mechanisms and effects. Physics, 2:20, 2009.

[83] U. Schollwoeck. The density-matrix renormalization group: a short introduction. Phil. Trans. R. Soc. A, 369:2643, 2011.

[84] M. Baer. Topological effects in molecular systems: an attempt towards a complete theory. Chem. Phys., 259:123, 2000.

[85] C. Wittig. The landau-zener formula. J. Phys. Chem., 109:8428, 2005.

[86] C. Zener. Non-adiabatic crossing of energy levels. Proc. Roy. Soc. A, 137:696, 1932.

[87] Roberto Car and Michele Parrinello. Unified approach for molecular dynamics and density-functional theory. Phys. Rev. Lett, 55:2471, 1985. doi: 10.1103/PhysRevLett.55.2471. URL `http://link.aps.org/doi/10.1103/PhysRevLett.55.2471`.

[88] L. Verlet. Computer "experiments" on classical fluids i. thermodynamical properties of lennard-jones molecules. Phys. Rev., 159:98, 1967.

[89] Sanwu Wang. Generalization of the thomas-reiche-kuhn and the bethe sum rules. Phys. Rev. A, 60:262–266, Jul 1999. doi: 10.1103/PhysRevA.60.262. URL `http://link.aps.org/doi/10.1103/PhysRevA.60.262`.

[90] A. Messiah. Quantum Mechanics. Dover, 2000.

[91] I.N. Bronstein and K.A. Semendjajew. Taschenbuch der Mathematik. BSB B.G. eubner Verlagsgesellschaft, Leipzig, 1983.

[92] E.N. Economou. Green's Functions in Quantum Mechanics. Springer, 1979.

[93] F.J Dyson. The radiation theories of tomonaga, schwinger and feynman. Phys. Rev., 75:486, 1949.

[94] Ryuzo Abe. Diagrammatic expansion in condensed matter physics. Prog. Theor. Phys., 8:19, 1984.

[95] T. Matsubara. A new approach to quantum-statistical physics. Prog. Theor. Phys., 14:351, 1955.

[96] Claude Bloch and Cyrano de Dominicis. Un développement du potentiel de gibbs d'un systéme quantique composié d'un grand nombre de particules. Nucl. Phys., 7:459, 1958.

[97] M. Gaudin. Une démonstration simplifiié du théoréme de wick en mécanique statistique. Nucl. Phys., 15:89, 1960.

[98] W. Magnus. On the exponential solution of differential equations for a linear operator. Comm. Pure Appl. Math., 7:649, 1954.

Index