

Solid State Theory

Spring Semester 2010

Manfred Sigrist
Institut für Theoretische Physik HIT K23.8
Tel.: 044-633-2584
Email: sigrist@itp.phys.ethz.ch

Literature:

- N.W. Ashcroft and N.D. Mermin: *Solid State Physics*, HRW International Editions, 1976.
- C. Kittel: *Einführung in die Festkörperphysik*, R. Oldenburg Verlag, 1983.
- C. Kittel: *Quantentheorie der Festkörper*, R. Oldenburg, 1970.
- O. Madelung: *Introduction to solid-state theory*, Springer 1981; auch in Deutsch in drei Bänden: *Festkörperphysik I-III*, Springer.
- J.M. Ziman: *Prinzipien der Festkörpertheorie*, Verlag Harri Deutsch, 1975.
- M.P. Marder: *Condensed Matter Physics*, John Wiley & Sons, 2000.
- G. Grosso & G.P. Parravicini: *Solid State Physics*, Academic Press, 2000.
- G. Czychol: *Theoretische Festkörperphysik*, Springer 2004.
- P.L. Taylor & O. Heinonen, *A Quantum Approach to Condensed Matter Physics*, Cambridge Press 2002.
- numerous specialized books.

Contents

Introduction	5
1 Electrons in the periodic crystal	8
1.1 Bloch states of electrons in the periodic crystal	8
1.1.1 Crystal symmetry	8
1.1.2 Bloch's theorem	9
1.2 Nearly free electron approximation	11
1.3 Symmetry properties of the band structure	13
1.4 $\mathbf{k} \cdot \mathbf{p}$ -expansion - effective masses	17
1.5 Band structures - approximate methods	19
1.5.1 Pseudo-potential	19
1.5.2 APW-method (augmented plane wave)	21
1.6 Tightly bound electrons and Wannier functions	23
1.7 Semi-classical description of band electrons	24
1.7.1 Semi-classical equations of motion	24
1.7.2 Current densities	25
1.8 Metals and semiconductors	26
2 Semiconductors	28
2.1 The band structure of the elements in group IV	29
2.1.1 Crystal structure and band structure	29
2.1.2 $\mathbf{k} \cdot \mathbf{p}$ - approximation and effective mass	31
2.2 Electronic properties and elementary excitations	32
2.2.1 Electron-hole excitations	33
2.2.2 Excitons	33
2.2.3 Optical properties	36
2.3 Doping semiconductors	37
2.3.1 Impurity state	37
2.3.2 Carrier concentration	39
2.4 Semiconductor devices	39
2.4.1 pn-contacts as diodes	39
2.4.2 Semiconductor diodes and light	40
2.4.3 MOSFET	41
3 Metals - properties of interacting electrons	44
3.1 The Jellium model of the metallic state	44
3.2 Charge excitations and dielectric function	47
3.2.1 Response and Lindhard function	48
3.2.2 Collective excitation - plasma resonance	50
3.2.3 Screening	52
3.3 Lattice vibrations - phonons in metals	54
3.3.1 Vibration of a isotropic continuous medium	54
3.3.2 Phonons in metals	56

3.3.3	Peierls instability in one dimension	57
3.3.4	Phonons and the dielectric function	60
3.4	Fermi Surfaces - the de Haas-van Alphen effect	62
3.4.1	Landau levels	62
3.4.2	Oscillatory behavior of the magnetization	63
3.4.3	Onsager equation	64
3.5	Quantum Hall Effect	66
3.5.1	Hall effect of the two-dimensional electron gas	67
3.5.2	Integer Quantum Hall Effect	68
3.5.3	Fractional Quantum Hall Effect	73
4	Landau's Theory of Fermi Liquids	76
4.1	Life-Time of quasiparticles	76
4.2	Phenomenological Theory of Fermi Liquids	79
4.2.1	Specific heat and density of states	80
4.2.2	Compressibility	81
4.2.3	Spin susceptibility	82
4.2.4	Effective mass and Galilei invariance	83
4.2.5	Stability of the Fermi liquid	84
4.3	Microscopic considerations	85
4.3.1	Landau parameters	86
4.3.2	Distribution function	88
5	Transport properties of metals	90
5.1	Electrical conductivity	90
5.2	Transport equations and relaxation time	91
5.2.1	The Boltzmann equation	91
5.2.2	The Drude form	93
5.2.3	The relaxation time	96
5.3	Impurity scattering	97
5.3.1	Potential scattering	97
5.3.2	Resonant scattering – Kondo effect	98
5.4	Electron-phonon interaction	99
5.5	Electron-electron scattering	102
5.6	Matthiessen's rule and the Ioffe-Regel limit	102
5.7	General transport coefficients	104
5.7.1	Generalized Boltzmann equation and the Wiedemann-Franz law	104
5.7.2	Thermoelectric effect	106
5.8	Transport in one dimension – Anderson localization	107
5.8.1	Landauer Formula	107
5.8.2	Anderson localization	109
6	Magnetism in metals	111
6.1	Stoner instability	111
6.1.1	Stoner model within the mean field approximation	112
6.1.2	Stoner criterion	113
6.1.3	Spin susceptibility for $T > T_C$	115
6.2	General spin susceptibility and magnetic instabilities	116
6.2.1	General dynamic spin susceptibility	116
6.2.2	Instability with finite wave vector \mathbf{q}	118
6.2.3	Influence of the band structure	119
6.3	Stoner excitations	121

7	Mott insulators and the magnetism of localized moments	123
7.1	Mott transition	124
7.1.1	Hubbard model	124
7.1.2	Insulating state	125
7.1.3	The metallic state	126
7.1.4	Fermi liquid properties of the metallic state	128
7.2	The Mott insulator as a quantum spin system	130
7.2.1	The effective Hamiltonian	130
7.2.2	Mean field approximation of the anti-ferromagnet	131
7.3	Collective modes – spin wave excitations	132

Introduction

Solid state physics (or condensed matter physics) is one of the most active and versatile branches of modern physics that have developed in the wake of the discovery of quantum mechanics. It deals with problems concerning the properties of materials and, more generally, systems with many degrees of freedom, ranging from fundamental questions to technological applications. This richness of topics has turned solid state physics into the largest subfield of physics; furthermore, it has arguably contributed most to technological development in industrialized countries.

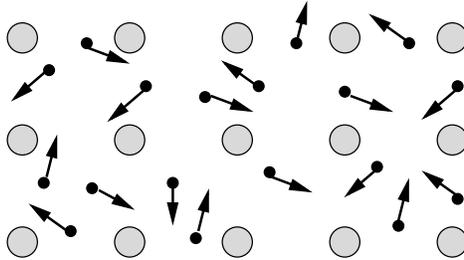


Figure 1: Atom cores and the surrounding electrons.

Condensed matter (solid bodies) consists of atomic nuclei (ions), usually arranged in a regular (elastic) lattice, and of electrons (see Fig. 1). As the macroscopic behavior of a solid is determined by the dynamics of these constituents, the description of the system requires the use of quantum mechanics. Thus, we introduce the Hamiltonian describing nuclei and electrons,

$$\hat{H} = \hat{H}_e + \hat{H}_n + \hat{H}_{n-e}, \quad (1)$$

with

$$\begin{aligned} \hat{H}_e &= \sum_i \frac{\hat{\mathbf{p}}_i^2}{2m} + \frac{1}{2} \sum_{i \neq i'} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_{i'}|}, \\ \hat{H}_n &= \sum_j \frac{\hat{\mathbf{P}}_j^2}{2M_j} + \frac{1}{2} \sum_{j \neq j'} \frac{Z_j Z_{j'} e^2}{|\mathbf{R}_j - \mathbf{R}_{j'}|}, \\ \hat{H}_{n-e} &= - \sum_{i,j} \frac{Z_j e^2}{|\mathbf{r}_i - \mathbf{R}_j|}, \end{aligned} \quad (2)$$

where \hat{H}_e (\hat{H}_n) describes the dynamics of the electrons (nuclei) and their mutual interaction and \hat{H}_{n-e} includes the interaction between ions and electrons. The parameters appearing are

m	free electron mass	$9.1094 \times 10^{-31} \text{kg}$
e	elementary charge	$1.6022 \times 10^{-19} \text{As}$
M_j	mass of j -th nucleus	$\sim 10^3 - 10^4 \times m$
Z_j	atomic (charge) number of j -th nucleus	

The characteristic scales known from atomic and molecular systems are

Length: Bohr radius	$a_B = \hbar^2/m_e^2 \approx 0.5 \times 10^{-10} \text{m}$
Energy: Hartree	$e^2/a_B = m_e^4/\hbar^2 = m_e c^2 \alpha^2 \approx 27 \text{eV} = 2 \text{Ry}$

with the fine structure constant $\alpha = e^2/\hbar c = 1/137$. The energy scale of one Hartree is much less than the (relativistic) rest mass of an electron ($\sim 0.5 \text{MeV}$), which in turn is considered small in particle physics. In fact, in high-energy physics even physics at the Planck scale is considered,

at least theoretically. The Planck scale is an energy scale so large that even gravity is thought to be affected by quantum effects, as

$$E_{\text{Planck}} = c^2 \sqrt{\frac{\hbar c}{G}} \sim 10^{19} \text{ GeV}, \quad l_{\text{Planck}} = \sqrt{\frac{\hbar G}{c^3}} \sim 1.6 \times 10^{-35} \text{ m}, \quad (3)$$

where $G = 6.673 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2}$ is the gravitational constant. This is the realm of the GUT (grand unified theory) and string theory. The goal is not to provide a better description of electrons or atomic cores, but to find the most fundamental theory of physics.

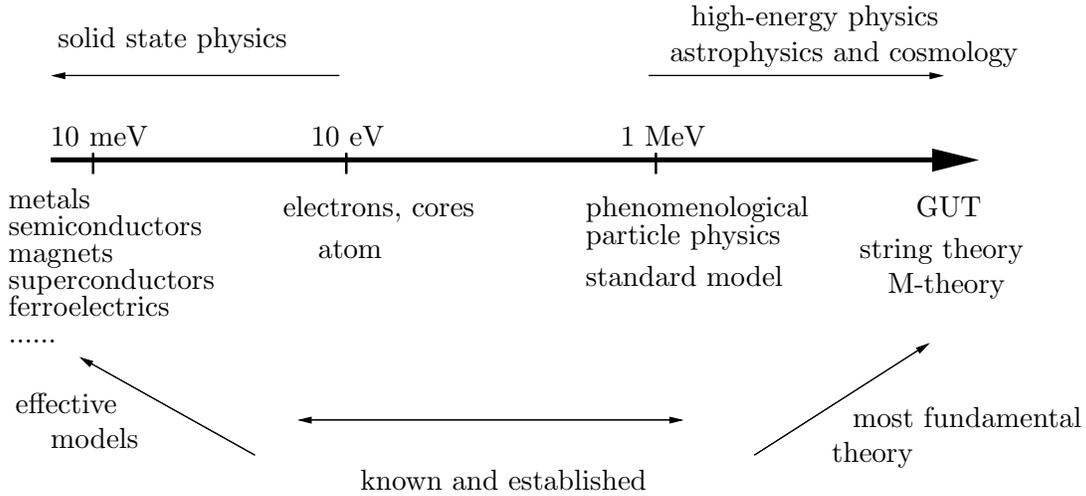


Figure 2: Energy scales in physics.

In contrast, in solid state physics we are dealing with phenomena occurring at room temperature ($T \sim 300\text{K}$) or below, i.e., at characteristic energies of about $E \sim k_B T \sim 0.03\text{eV} = 30\text{meV}$, which is even much smaller than the energy scale of one Hartree. Correspondingly, the important length scales are given by the extension of the system or of the electronic wave functions. The focus is thus quite different from the one of high-energy physics.

There, a highly successful phenomenological theory for low energies, the so-called standard model, exists, whereas the underlying theory for higher energies is unknown. In solid state physics, the situation is reversed. The Hamiltonian (1, 3) describes the known 'high-energy' physics (on the energy scale of Hartree), and one aims at describing the low-energy properties using reduced (effective, phenomenological) theories. Both tasks are far from trivial.

Among the various states of condensed matter that solid state theory seeks to describe are metals, semiconductors, and insulators. Furthermore, there are phenomena such as magnetism, superconductivity, ferroelectricity, charge ordering, and the quantum Hall effect. All of these states share a common origin: Electrons interacting among themselves and with the ions through the Coulomb interaction. More often than not, the microscopic formulation in (1) is too complicated to allow an understanding of the low-energy behavior from first principles. Consequently, the formulation of effective (reduced) theories is an important step in condensed matter theory. On the one hand, characterizing the ground state of a system is an important goal in itself. However, measurable quantities are determined by excited states, so that the concept of 'elementary excitations' takes on a central role. Some celebrated examples are Landau's quasi-particles for Fermi liquids, the phonons connected to lattice vibrations, and magnons in ferromagnets. The idea is to treat the ground state as an effective vacuum in the sense of second quantization, with the elementary excitations as particles on that vacuum. Depending on the system, the vacuum may be the Fermi sea or some state with a broken symmetry, like a ferromagnet, a superconductor, or the crystal lattice itself.

According to P. W. Anderson¹, the description of the properties of materials rests on two principles: The principle of *adiabatic continuity* and the principle of *spontaneously broken symmetry*. By adiabatic continuity we mean that complicated systems may be replaced by simpler systems that have the same essential properties in the sense that the two systems may be adiabatically deformed into each other without changing qualitative properties. Arguably the most impressive example is Landau's Fermi liquid theory mentioned above. The low-energy properties of strongly interacting electrons are the same as those of non-interacting fermions with renormalized parameters. On the other hand, phase transitions into states with qualitatively different properties can often be characterized by broken symmetries. In magnetically ordered states the rotational symmetry and the time-reversal invariance are broken, whereas in the superconducting state the global gauge symmetry is. In many cases the violation of a symmetry is a guiding principle which helps to simplify the theoretical description considerably. Moreover, in recent years some systems have been recognized as having topological order which may be considered as a further principle to characterize low-energy states of matter. A famous example for this is found in the context of the Quantum Hall effect.

The goal of these lectures is to introduce these basic concepts on which virtually all more elaborate methods are building up. In the course of this, we will cover a wide range of frequently encountered ground states, starting with the theory of metals and semiconductors, proceeding with magnets, Mott insulators, and finally superconductors.

¹P.W. Anderson: *Basic Notions of Condensed Matter Physics*, Frontiers in Physics Lecture Notes Series, Addison-Wesley (1984).

Chapter 1

Electrons in the periodic crystal

In this chapter we discuss the properties of extended electron states in a regular lattice of ions. Due to the presence of the lattice, the spectrum of the electrons is modified as compared to the one of free particles, leading to separate energy bands, which determine the qualitative properties of a solid. In particular, the structure of the electron bands can be used to distinguish in a most basic way metals, insulators, and semiconductors.

In the following considerations we will initially neglect the interactions among the electrons as well as the dynamics of the ions. This simplification leads to a single particle description, to which Bloch's theorem can be applied.

1.1 Bloch states of electrons in the periodic crystal

1.1.1 Crystal symmetry

We consider a perfect crystal formed by a periodic array of positively charged ions. All crystals can be characterized by their *space group* \mathcal{R} . In three dimensions, there are 230 different space groups (cf. Table 1.1), each consisting of translations, rotations, inversions and their combinations. Translations are represented by a basic set of primitive translation vectors $\{\mathbf{a}_i\}$, which leave the lattice invariant. A translation by one of these vectors shifts a unit cell of the lattice to a neighboring cell. Any translation that maps the lattice onto itself is a linear combination of the $\{\mathbf{a}_i\}$ with integer coefficients,

$$\mathbf{a} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3. \quad (1.1)$$

A general symmetry transformation including the other elements of the space group may be written in the notation due to Wigner,

$$\mathbf{r}' = g\mathbf{r} + \mathbf{a} = \{g|\mathbf{a}\}\mathbf{r}, \quad (1.2)$$

with g a rotation, reflection or inversion. The elements g form the so-called generating *point group* \mathcal{P} . In three dimensions there are 32 point groups. The different types of symmetry operations involve

basic translations	$\{E \mathbf{a}\}$,
rotations, reflections, inversions	$\{g \mathbf{0}\}$,
screw axes, glide planes	$\{g \mathbf{a}\}$,

where E is the unit element of \mathcal{P} . A screw axis is a symmetry operation of a rotation followed by a translation along the rotation axis. A glide plane is a symmetry operation with reflection

at the plane followed by a translation along the plane. The symmetry operations form a group with

$$\begin{aligned}
&\text{unit element} && \{E|\mathbf{0}\}, \\
&\text{inverse} && \{g|\mathbf{a}\}^{-1} = \{g^{-1}|-g^{-1}\mathbf{a}\}, \\
&\text{associative multiplication} && \{g|\mathbf{a}\}\{g'|\mathbf{a}'\} = \{gg'|\mathbf{ga}' + \mathbf{a}\}.
\end{aligned} \tag{1.3}$$

In general, these groups are non-Abelian, i.e., the group elements do not commute with each other. However, there is an Abelian subgroup, the group of translations $\{E|\mathbf{a}\}$. The elements $g \in \mathcal{P}$ do not necessarily form a subgroup, because some of these elements (e.g., screw axes or glide planes) leave the lattice invariant only in combination with a translation. Nevertheless,

$$\{g|\mathbf{a}\}\{E|\mathbf{a}'\}\{g|\mathbf{a}\}^{-1} = \{E|\mathbf{ga}'\} \quad \text{and} \quad \{g|\mathbf{a}\}^{-1}\{E|\mathbf{a}'\}\{g|\mathbf{a}\} = \{E|g^{-1}\mathbf{a}'\} \tag{1.4}$$

always holds. If \mathcal{P} is a subgroup of \mathcal{R} , then \mathcal{R} is said to be *symmorphic*. In this case, the space group contains only primitive translations $\{E|\mathbf{a}\}$ (no screw axes nor glide planes). The 14 Bravais lattices are symmorphic. Among the 230 space groups 73 are symmorphic and 157 are non-symmorphic.

crystal system (# point groups, # space groups)	point groups Schönflies symbols	space group numbers international tables
triclinic (2, 2)	$C_1, C_{\bar{1}}$	1 - 2
monoclinic (3, 13)	C_2, C_s, C_{2h}	3 - 15
orthorhombic (3, 59)	D_2, C_{2v}, D_{2h}	16 - 74
tetragonal (7, 68)	$C_4, S_4, C_{4h}, D_4, C_{4v}, D_{2d}, D_{4h}$	75 - 142
trigonal (5, 25)	$C_3, S_6, D_3, C_{3v}, D_{3d}$	143 - 167
hexagonal (7, 27)	$C_6, C_{3h}, C_{6h}, D_6, C_{6v}, D_{3h}, D_{6h}$	168 - 194
cubic (5, 36)	T, T_h, O, T_d, O_h	195 - 230

Table 1.1: Table of point and space groups.

1.1.2 Bloch's theorem

We consider a discrete set of lattice translations $\{E|\mathbf{a}\}$ which leave the Hamiltonian invariant. This (discrete) translational invariance is induced by the periodic ionic potential and means that the corresponding translation operator $\hat{T}_{\mathbf{a}}$ on the Hilbert space commutes with the Hamiltonian $\mathcal{H}_e + \mathcal{H}_{ie}$ (purely electronic Hamiltonian \mathcal{H}_e , interaction between electrons and ions \mathcal{H}_{ie}),

$$[\hat{T}_{\mathbf{a}}, \mathcal{H}_e + \mathcal{H}_{ie}] = 0, \tag{1.5}$$

where

$$\mathcal{H}_{ie} = \sum_s \int d^3r V(\mathbf{r}) \hat{\Psi}_s^\dagger(\mathbf{r}) \hat{\Psi}_s(\mathbf{r}) \tag{1.6}$$

and

$$V(\mathbf{r}) = \sum_j V_{\text{ion}}(\mathbf{r} - \mathbf{R}_j), \quad (1.7)$$

where \mathbf{R}_j is the position of the j -th ion and $V_{\text{ion}}(\mathbf{r})$ is the potential of a single ion. We use $\widehat{\Psi}_s(\mathbf{r})$ as the electron field operator in the second quantized formalism. For all lattice translations \mathbf{a} , $V(\mathbf{r} + \mathbf{a}) = V(\mathbf{r})$ holds.

Neglecting the interactions between the electrons, which is contained in the general \mathcal{H}_e , we are left with a single particle problem

$$\mathcal{H}_e + \mathcal{H}_{\text{ie}} \quad \rightarrow \quad \mathcal{H}_0 = \frac{\widehat{\mathbf{p}}^2}{2m} + V(\widehat{\mathbf{r}}). \quad (1.8)$$

\mathcal{H}_0 commutes with $\widehat{T}_{\mathbf{a}}$. Bloch's theorem states that the extended eigenstates of \mathcal{H}_0 are simultaneous eigenstates of $\widehat{T}_{\mathbf{a}}$, with eigenvalues on the unit circle of the complex plane. We use the Bloch ansatz

$$\psi_{n,\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} e^{i\mathbf{k}\cdot\mathbf{r}} u_{n,\mathbf{k}}(\mathbf{r}), \quad (1.9)$$

with

$$\widehat{T}_{\mathbf{a}} \psi_{n,\mathbf{k}}(\mathbf{r}) = \psi_{n,\mathbf{k}}(\mathbf{r} - \mathbf{a}) = e^{-i\mathbf{k}\cdot\mathbf{a}} \psi_{n,\mathbf{k}}(\mathbf{r}), \quad (1.10)$$

which means that the Bloch function $u_{n,\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r} + \mathbf{a})$ is periodic (volume Ω). The energy eigenvalues follow from

$$\mathcal{H}_0 \psi_{n,\mathbf{k}}(\mathbf{r}) = \epsilon_{n,\mathbf{k}} \psi_{n,\mathbf{k}}(\mathbf{r}), \quad (1.11)$$

where n is a quantum number called band index and \mathbf{k} is the pseudo-momentum (wave vector). Note that the eigenvalue of $\widehat{T}_{\mathbf{a}}$, $e^{-i\mathbf{k}\cdot\mathbf{a}}$, implies periodicity in \mathbf{k} -space; there are reciprocal lattice vectors \mathbf{G} for which $e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{a}} = e^{i\mathbf{k}\cdot\mathbf{a}}$ holds. A possible basis of the reciprocal lattice vectors follows from the relation

$$e^{i\mathbf{G}_j\cdot\mathbf{a}_i} = 1 \quad \iff \quad \mathbf{G}_j \cdot \mathbf{a}_i = 2\pi\delta_{ij}. \quad (1.12)$$

This defines the first Brillouin zone: One draws lines joining $\mathbf{k} = 0$ and the neighboring reciprocal lattice points (spanned by $\{\mathbf{G}_i\}$). The Brillouin zone is the smallest cell bounded by the planes that intersect these lines in their middle and which are orthogonal to them. In the one dimensional simple periodic lattice this defines the interval $[-\pi/a, \pi/a]$ (lattice constant a).

The Bloch equation is the wave equation for the periodic function $u_{\mathbf{k}}$,

$$\left\{ \frac{(\widehat{\mathbf{p}} + \hbar\mathbf{k})^2}{2m} + V(\mathbf{r}) \right\} u_{\mathbf{k}}(\mathbf{r}) = \epsilon_{\mathbf{k}} u_{\mathbf{k}}(\mathbf{r}), \quad (1.13)$$

where we have suppressed the band index to simplify the notation. This equation follows from the relation

$$\widehat{\mathbf{p}} e^{i\mathbf{k}\cdot\mathbf{r}} = e^{i\mathbf{k}\cdot\mathbf{r}} (\widehat{\mathbf{p}} + \hbar\mathbf{k}), \quad (1.14)$$

which can be used for more complex forms of the Hamiltonian as well. ¹

¹ \mathcal{H}_0 may be extended to contain spin-orbit coupling, a relativistic correction which leads to the additional term

$$\mathcal{H}'_0 = \frac{\widehat{\mathbf{p}}^2}{2m} + V(\widehat{\mathbf{r}}) + \frac{\hbar}{4m^2c^2} \{ \boldsymbol{\sigma} \times \nabla V(\widehat{\mathbf{r}}) \} \cdot \widehat{\mathbf{p}}, \quad (1.15)$$

where $\boldsymbol{\sigma}$ denotes the Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.16)$$

The Bloch equation in this case is given by

$$\left\{ \frac{(\widehat{\mathbf{p}} + \hbar\mathbf{k})^2}{2m} + V(\mathbf{r}) + \frac{\hbar}{4m^2c^2} (\boldsymbol{\sigma} \times \nabla V(\mathbf{r})) \cdot (\widehat{\mathbf{p}} + \hbar\mathbf{k}) \right\} u_{\mathbf{k}}(\mathbf{r}) = \epsilon_{\mathbf{k}} u_{\mathbf{k}}(\mathbf{r}). \quad (1.17)$$

1.2 Nearly free electron approximation

We can compute $\epsilon_{n,\mathbf{k}}$ numerically rather efficiently by means of various clever methods. In order to reach an understanding of some of the most essential aspects of the band structure of electrons in a crystal, we introduce here a simple analytical approach, the so-called nearly free electron approximation. We start by noting that the periodic potential can be expanded as

$$V(\mathbf{r}) = \sum_{\mathbf{G}} V_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}}, \quad V_{\mathbf{G}} = \frac{1}{\Omega_{\text{UC}}} \int_{\text{UC}} d^3r V(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}}, \quad (1.19)$$

where the sum runs over all reciprocal lattice vectors and the domain of integration is the unit cell (UC) with volume Ω_{UC} . We assume that the lattice is invariant under inversion, i.e., $V(\mathbf{r}) = V(-\mathbf{r})$, so that $V_{\mathbf{G}} = V_{-\mathbf{G}}$. Note that the uniform component V_0 may be set to zero, as it corresponds to an (irrelevant) energy shift.

The Bloch function can be expanded in the same way,

$$u_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}}, \quad (1.20)$$

where the coefficients $c_{\mathbf{G}} = c_{\mathbf{G}}(\mathbf{k})$ are functions of \mathbf{k} . Inserting this ansatz and (1.19) into the Bloch equation, (1.13), we obtain a system of coupled linear eigenvalue equations for the band energies $\epsilon_{\mathbf{k}}$,

$$\left(\frac{\hbar^2}{2m} (\mathbf{k} + \mathbf{G})^2 - \epsilon_{\mathbf{k}} \right) c_{\mathbf{G}} + \sum_{\mathbf{G}'} V_{\mathbf{G}-\mathbf{G}'} c_{\mathbf{G}'} = 0. \quad (1.21)$$

The solution requires the determination of the eigenvalues of an infinite dimensional matrix. The resulting band energies $\epsilon_{\mathbf{k}}$ include corrections to the parabolic dispersion $\epsilon_{\mathbf{k}}^{(0)} = \hbar^2 \mathbf{k}^2 / 2m$ due to the potential.

The problem simplifies under the assumption that the periodic modulation of the potential is weak. Here, we consider two limits for the wave vector \mathbf{k} which are typically of interest. First, we choose \mathbf{k} small, i.e., near the center of the Brillouin zone. A solution of the equation is then given by $c_0 \approx 1$ with the energy eigenvalue $\epsilon_{\mathbf{k}} \approx \hbar^2 \mathbf{k}^2 / 2m$ corresponding to the original parabolic band. For the other coefficients of the wave function we find

$$c_{\mathbf{G}} \approx - \frac{2mV_{\mathbf{G}}}{\hbar^2 \{ (\mathbf{k} + \mathbf{G})^2 - \mathbf{k}^2 \}} \ll 1 \quad \text{for } \mathbf{G} \neq \mathbf{0}. \quad (1.22)$$

Note that this form of $c_{\mathbf{G} \neq \mathbf{0}}$ resembles the lowest order correction in the Rayleigh-Schrödinger perturbation theory. This example corresponds to the lowest branch of the band structure within this approach.

Next we consider the case when the denominator of the expression in Eq.(1.22) is small, i.e., \mathbf{k} is in a range of the Brillouin zone where $\mathbf{k}^2 \approx (\mathbf{k} + \mathbf{G})^2$ for some reciprocal \mathbf{G} . This means that the parabolas centered around $\mathbf{0}$ and $-\mathbf{G}$ cross at $\mathbf{k} = -\mathbf{G}/2$. Choosing for \mathbf{G} a primitive vector of the reciprocal lattice, the crossing point lies on the Brillouin zone boundary and represents a point of high symmetry within the Brillouin zone. This situation requires to consider c_0 and $c_{\mathbf{G}}$ on an equal footing, while other coefficients are still negligible. Therefore, we consider the coupled equations for these two coefficients,

The energy eigenstates are no longer spin eigenstates. Instead, they are of pseudo spinor form

$$u_{\mathbf{k},\pm}(\mathbf{r}) = \chi_{\mathbf{k},\pm\uparrow}(\mathbf{r}) |\uparrow\rangle + \chi_{\mathbf{k},\pm\downarrow}(\mathbf{r}) |\downarrow\rangle, \quad (1.18)$$

where $\sigma_z |\uparrow\rangle = +|\uparrow\rangle$ und $\sigma_z |\downarrow\rangle = -|\downarrow\rangle$. Upon adiabatically switching off spin-orbit coupling, the states with index $+/-$ turn into the usual spin eigenfunctions $|\uparrow\rangle$ and $|\downarrow\rangle$.

$$\begin{aligned} \left\{ \frac{\hbar^2 \mathbf{k}^2}{2m} - \epsilon_{\mathbf{k}} \right\} c_0 + V_{-\mathbf{G}} c_{\mathbf{G}} &= 0, \\ \left\{ \frac{\hbar^2 (\mathbf{k} + \mathbf{G})^2}{2m} - \epsilon_{\mathbf{k}} \right\} c_{\mathbf{G}} + V_{\mathbf{G}} c_0 &= 0. \end{aligned} \quad (1.23)$$

Note that $V_{\mathbf{G}} = V_{-\mathbf{G}}^*$. From Eq.(1.23), the secular equation

$$\det \begin{bmatrix} \frac{\hbar^2 \mathbf{k}^2}{2m} - \epsilon_{\mathbf{k}} & V_{\mathbf{G}}^* \\ V_{\mathbf{G}} & \frac{\hbar^2 (\mathbf{k} + \mathbf{G})^2}{2m} - \epsilon_{\mathbf{k}} \end{bmatrix} = 0 \quad (1.24)$$

follows, which allows us to determine

$$\epsilon_{\mathbf{k}} = \frac{1}{2} \left\{ \frac{\hbar^2}{2m} (\mathbf{k}^2 + (\mathbf{k} + \mathbf{G})^2) \pm \sqrt{\left[\frac{\hbar^2}{2m} (\mathbf{k}^2 - (\mathbf{k} + \mathbf{G})^2) \right]^2 + 4|V_{\mathbf{G}}|^2} \right\}. \quad (1.25)$$

For the symmetry point $\mathbf{k} = -\mathbf{G}/2$ and for $V_{\mathbf{G}} < 0$ we obtain

$$\epsilon_{-\mathbf{G}/2, \pm} = \frac{\hbar^2}{2m} \frac{\mathbf{G}^2}{4} \pm |V_{\mathbf{G}}|, \quad \text{with} \quad u_{\mathbf{k}}(\mathbf{r}) = e^{i \frac{\mathbf{G} \cdot \mathbf{r}}{2}} \begin{cases} \sin \frac{\mathbf{G} \cdot \mathbf{r}}{2} & + \text{ "anti-bonding" }, \\ \cos \frac{\mathbf{G} \cdot \mathbf{r}}{2} & - \text{ "bonding" }. \end{cases} \quad (1.26)$$

This result is equivalent to the splitting of a degenerate level through a symmetry breaking interaction (hybridization). Note that the scheme applied here is quite analogous to Rayleigh-Schrödinger perturbation theory for (nearly) degenerate energy levels.

The band structure can thus be constructed by the superposition of parabolic energy spectra centered around all reciprocal lattice points. At the crossing points of the parabolas we find a "band splitting" due to the periodically modulated potential. This leads to band gaps, i.e., energy ranges where no Bloch states exist. An illustrative and simple band structure of this kind can straightforwardly be constructed in a one-dimensional regular lattice as shown in Fig. 1.1.

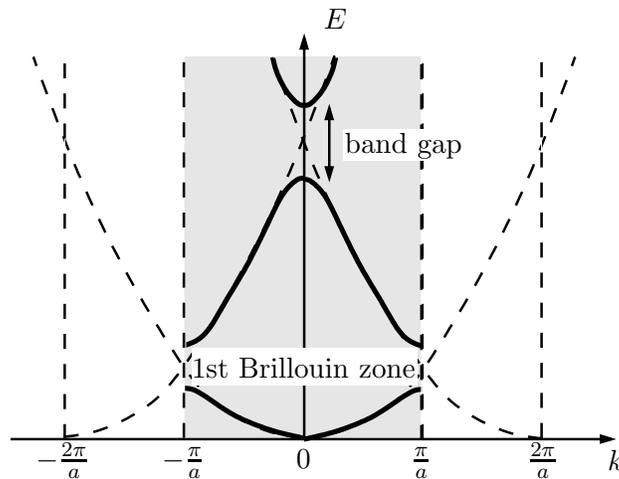


Figure 1.1: Band structure obtained by the nearly free electron approximation for a regular one-dimensional lattice.

1.3 Symmetry properties of the band structure

The symmetry properties are a helpful tool for the analysis of the band structure of crystals, which can be based on the symmetry group (space and point group) of the crystal lattice. Consider the action of an element $\{g|\mathbf{a}\}$ of the space group on a Bloch wave function $\Psi_{\mathbf{k}}(\mathbf{r})$. We denote the corresponding operator as $\widehat{S}_{\{g|\mathbf{a}\}}$, with $[\widehat{S}_{\{g|\mathbf{a}\}}, \mathcal{H}_0] = 0$ and define the operation as²

$$\widehat{S}_{\{g|\mathbf{a}\}}\psi_{\mathbf{k}}(\mathbf{r}) = \psi_{\mathbf{k}}(\{g|\mathbf{a}\}^{-1}\mathbf{r}) = \psi_{\mathbf{k}}(g^{-1}\mathbf{r} - g^{-1}\mathbf{a}). \quad (1.30)$$

Under a pure translation $\widehat{T}_{\mathbf{a}'} = \widehat{S}_{\{E|\mathbf{a}'\}}$ this new wave function transforms like

$$\begin{aligned} \widehat{T}_{\mathbf{a}'}\widehat{S}_{\{g|\mathbf{a}\}}\psi_{\mathbf{k}}(\mathbf{r}) &= \widehat{S}_{\{g|\mathbf{a}\}}\widehat{T}_{g^{-1}\mathbf{a}'}\psi_{\mathbf{k}}(\mathbf{r}) = \widehat{S}_{\{g|\mathbf{a}\}}e^{-i\mathbf{k}\cdot g^{-1}\mathbf{a}'}\psi_{\mathbf{k}}(\mathbf{r}) \\ &= \widehat{S}_{\{g|\mathbf{a}\}}e^{-ig\mathbf{k}\cdot\mathbf{a}'}\psi_{\mathbf{k}}(\mathbf{r}) = e^{-ig\mathbf{k}\cdot\mathbf{a}'}\widehat{S}_{\{g|\mathbf{a}\}}\psi_{\mathbf{k}}(\mathbf{r}), \end{aligned} \quad (1.31)$$

which implies that

$$\widehat{S}_{\{g|\mathbf{a}\}}\psi_{\mathbf{k}}(\mathbf{r}) = \lambda_{\{g|\mathbf{a}\}}\psi_{g\mathbf{k}}(\mathbf{r}), \quad \text{with } |\lambda_{\{g|\mathbf{a}\}}|^2 = 1, \quad (1.32)$$

or, in Dirac notation,

$$\widehat{S}_{\{g|\mathbf{a}\}}|\mathbf{k}\rangle = \lambda_{\{g|\mathbf{a}\}}|g\mathbf{k}\rangle; \quad (1.33)$$

i.e., apart from a phase factor the action of $\{g|\mathbf{a}\}$ corresponds to a rotation of \mathbf{k} to $g^{-1}\mathbf{k}$.³ Then it is easy to see that

$$\epsilon_{g\mathbf{k}} = \langle g\mathbf{k}|\mathcal{H}_0|g\mathbf{k}\rangle = \langle \mathbf{k}|\widehat{S}_{\{g|\mathbf{a}\}}^{-1}\mathcal{H}_0\widehat{S}_{\{g|\mathbf{a}\}}|\mathbf{k}\rangle = \langle \mathbf{k}|\mathcal{H}_0|\mathbf{k}\rangle = \epsilon_{\mathbf{k}}. \quad (1.35)$$

Consequently, there is a *star* of equivalent points $g\mathbf{k}$ with the same band energy (\rightarrow degeneracy) for each \mathbf{k} in the Brillouin zone (cf. Fig. 1.2).

For a general point \mathbf{k} the number of points in the star equals the number of point group elements (without inversion). If \mathbf{k} lies on points or lines of higher symmetry, it is left invariant under a subgroup of the point group. Consequently, the number of “beams” of the star is smaller. The subgroups leaving \mathbf{k} unchanged are called *little group* of \mathbf{k} . If inversion is part of the point group, $-\mathbf{k}$ is always contained in the star of \mathbf{k} . In summary, we have the simple relations

$$\epsilon_{n\mathbf{k}} = \epsilon_{n,g\mathbf{k}}, \quad \epsilon_{n\mathbf{k}} = \epsilon_{n,-\mathbf{k}}, \quad \epsilon_{n\mathbf{k}} = \epsilon_{n,\mathbf{k}+\mathbf{G}}. \quad (1.36)$$

²In Dirac notation we write

$$\psi_{\mathbf{k}}(\mathbf{r}) = \langle \mathbf{r}|\psi_{\mathbf{k}}\rangle \quad (1.27)$$

for the Bloch state with pseudo-momentum \mathbf{k} . The action of the operator $\widehat{S}_{\{g|\mathbf{a}\}}$ on the state $|\mathbf{r}\rangle$ is given by

$$\widehat{S}_{\{g|\mathbf{a}\}}|\mathbf{r}\rangle = |g\mathbf{r} + \mathbf{a}\rangle \quad \text{and} \quad \langle \mathbf{r}|\widehat{S}_{\{g|\mathbf{a}\}} = \langle g^{-1}\mathbf{r} - g^{-1}\mathbf{a}|, \quad (1.28)$$

such that

$$\langle \mathbf{r}|\widehat{S}_{\{g|\mathbf{a}\}}|\psi_{\mathbf{k}}\rangle = \psi_{\mathbf{k}}(g^{-1}\mathbf{r} - g^{-1}\mathbf{a}). \quad (1.29)$$

The same holds for pure translations.

³Symmetry behavior:

$$\begin{aligned} \widehat{S}_{\{g|\mathbf{a}\}}\psi_{\mathbf{k}}(\mathbf{r}) &= \frac{1}{\sqrt{\Omega}}\widehat{S}_{\{g|\mathbf{a}\}}e^{i\mathbf{k}\cdot\mathbf{r}}\sum_{\mathbf{G}}c_{\mathbf{G}}(\mathbf{k})e^{i\mathbf{G}\cdot\mathbf{r}} = \frac{1}{\sqrt{\Omega}}e^{i\mathbf{k}\cdot(g^{-1}\mathbf{r}-g^{-1}\mathbf{a})}\sum_{\mathbf{G}}c_{\mathbf{G}}(\mathbf{k})e^{i\mathbf{G}\cdot(g^{-1}\mathbf{r}-g^{-1}\mathbf{a})} \\ &= \frac{1}{\sqrt{\Omega}}e^{-i(g\mathbf{k})\cdot\mathbf{a}}e^{i(g\mathbf{k})\cdot\mathbf{r}}\sum_{\mathbf{G}}c_{\mathbf{G}}(\mathbf{k})e^{i(g\mathbf{G})\cdot\mathbf{r}} = e^{-i(g\mathbf{k})\cdot\mathbf{a}}\frac{1}{\sqrt{\Omega}}e^{i(g\mathbf{k})\cdot\mathbf{r}}\sum_{\mathbf{G}}c_{g^{-1}\mathbf{G}}(\mathbf{k})e^{i\mathbf{G}\cdot\mathbf{r}} \\ &= e^{-i(g\mathbf{k})\cdot\mathbf{a}}\frac{1}{\sqrt{\Omega}}e^{i(g\mathbf{k})\cdot\mathbf{r}}\sum_{\mathbf{G}}c_{\mathbf{G}}(g\mathbf{k})e^{i\mathbf{G}\cdot\mathbf{r}} = \lambda_{\{g|\mathbf{a}\}}\psi_{g\mathbf{k}}(\mathbf{r}), \end{aligned} \quad (1.34)$$

where we use the fact that $c_{\mathbf{G}} = c_{\mathbf{G}}(\mathbf{k})$ is a function of \mathbf{k} with the property $c_{g^{-1}\mathbf{G}}(\mathbf{k}) = c_{\mathbf{G}}(g\mathbf{k})$ i.e. $\widehat{S}_{\{g|\mathbf{a}\}}u_{\mathbf{k}}(\mathbf{r}) = u_{g\mathbf{k}}(\mathbf{r})$.

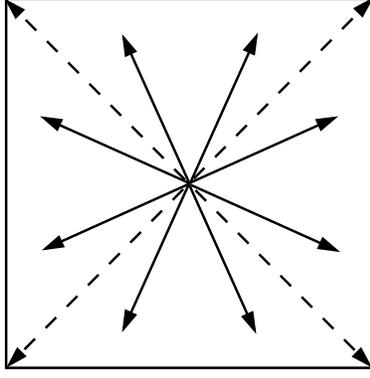


Figure 1.2: Star of \mathbf{k} .

Next, we will consider the energy bands $\epsilon_{n\mathbf{k}}$ on points and along lines of high symmetry in a simple cubic lattice (point group O_h), using the nearly free electron method.

Γ -point: As a first example, we consider the center of the Brillouin zone, usually called the Γ -point (cf. Fig. 1.3). The lowest band at the Γ -point with energy $E_0 = \epsilon_{0\mathbf{k}=0} = 0$ belongs to the parabola around the center of the first Brillouin zone ($\epsilon_{0\mathbf{k}} \approx \hbar^2 \mathbf{k}^2 / 2m$) and is non-degenerate.

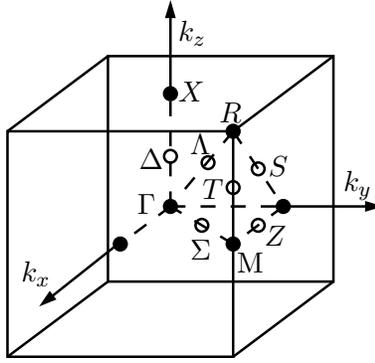


Figure 1.3: Points and lines of high symmetry.

The next higher energy level for free electrons is

$$E_1 = \frac{\hbar^2}{2m} \left(\frac{2\pi}{a} \right)^2 \quad (1.37)$$

and originates from the crossing of the parabolas centered around the six nearest neighbor points of the reciprocal lattice. The reciprocal lattice vectors involved are

$$\begin{aligned} \mathbf{G}_1 &= \frac{2\pi}{a}(1, 0, 0), & \mathbf{G}_2 &= \frac{2\pi}{a}(-1, 0, 0), \\ \mathbf{G}_3 &= \frac{2\pi}{a}(0, 1, 0), & \mathbf{G}_4 &= \frac{2\pi}{a}(0, -1, 0), \\ \mathbf{G}_5 &= \frac{2\pi}{a}(0, 0, 1), & \mathbf{G}_6 &= \frac{2\pi}{a}(0, 0, -1). \end{aligned} \quad (1.38)$$

The relevant basis functions for the expansion of the Bloch function are given by

$$f_n(\mathbf{r}) = e^{i\mathbf{r} \cdot \mathbf{G}_n}, \quad (1.39)$$

with $n = 1, \dots, 6$ and

$$u_{\mathbf{k}=0}(\mathbf{r}) = \sum_{n=1}^6 c_n f_n(\mathbf{r}). \quad (1.40)$$

The secular equation reads

$$\det \begin{bmatrix} E_1 - E & v & u & u & u & u \\ v & E_1 - E & u & u & u & u \\ u & u & E_1 - E & v & u & u \\ u & u & v & E_1 - E & u & u \\ u & u & u & u & E_1 - E & v \\ u & u & u & u & v & E_1 - E \end{bmatrix} = 0 \quad (1.41)$$

with $v = V_{2\mathbf{G}_n}$ and $u = V_{\mathbf{G}_n + \mathbf{G}_{n'}}$ ($n \neq n'$). There are three eigenvalues with corresponding eigenvectors:

Γ	$E = \epsilon_{n\mathbf{k}=\mathbf{G}_1}$	$(c_1, c_2, c_3, c_4, c_5, c_6)$	$u_{\mathbf{k}=0}(\mathbf{r})$	d_Γ
Γ_1^+	$E_1 + v + 4u$	$(1, 1, 1, 1, 1, 1)/\sqrt{6}$	$\phi_0 = \cos Gx + \cos Gy + \cos Gz$	1
Γ_3^+	$E_1 + v - 2u$	$(-1, -1, -1, -1, 2, 2)/2\sqrt{3}$ $(1, 1, -1, -1, 0, 0)/2$	$\phi_{3z^2-r^2} = 2 \cos Gz - \cos Gx - \cos Gy$, $\phi_{\sqrt{3}(x^2-y^2)} = \sqrt{3}(\cos Gx - \cos Gy)$	2
Γ_4^-	$E_1 - v$	$(1, -1, 0, 0, 0, 0)/\sqrt{2}$ $(0, 0, 1, -1, 0, 0)/\sqrt{2}$ $(0, 0, 0, 0, 1, -1)/\sqrt{2}$	$\phi_x = \sin Gx$ $\phi_y = \sin Gy$ $\phi_z = \sin Gz$	3

Here, $G = 2\pi/a$ and Γ denotes the irreducible representations with dimension d_Γ of the point group around the Γ -point (\rightarrow degeneracy). The Γ -point shares the symmetry of the point group of the crystal, which in this case is the cubic group O_h .⁴ A set of even and odd irreducible representations belongs to this group. An irreducible representation can be specified by a vector space of functions of the vector (x, y, z) or the pseudo-vector (s_x, s_y, s_z) that is left invariant by symmetry operations of the group (see Table 1.2). Note that each eigenvalue of the above

even	basis function	odd	basis function
Γ_1^+	$1, x^2 + y^2 + z^2$	Γ_1^-	$xyz(x^2 - y^2)(y^2 - z^2)(z^2 - x^2)$
Γ_2^+	$(x^2 - y^2)(y^2 - z^2)(z^2 - x^2)$	Γ_2^-	xyz
Γ_3^+	$\{2z^2 - x^2 - y^2, \sqrt{3}(x^2 - y^2)\}$	Γ_3^-	$xyz\{2z^2 - x^2 - y^2, \sqrt{3}(x^2 - y^2)\}$
Γ_4^+	$\{s_x, s_y, s_z\}$	Γ_4^-	$\{x, y, z\}$
Γ_5^+	$\{yz, zx, xy\}$	Γ_5^-	$xyz(x^2 - y^2)(y^2 - z^2)(z^2 - x^2)\{yz, zx, xy\}$

Table 1.2: Irreducible representations and representative basis functions of the corresponding vector spaces for the point group O_h .

secular equation belongs to one of the irreducible representations. The corresponding wave functions of the eigenstates form a vector space and transform according to the properties of the representation under symmetry operations.

Δ -line: Now we will investigate the evolution of the band energies when we move \mathbf{k} away from the Γ -point and keep $\mathbf{k} \parallel (0, 0, 1)$. Some of the degeneracies at the Γ -point are lifted because the allowed symmetry operations leaving \mathbf{k} unchanged are restricted to a subgroup of O_h , the little group of \mathbf{k} . In the case at hand, this subgroup is isomorphic to C_{4v} , which is part of the tetragonal crystal system. Note that the inversion acts as $\mathbf{k} \rightarrow -\mathbf{k}$ and is not an element of the little group. The group C_{4v} has four one-dimensional and one two-dimensional representations. As the line along the $(0, 0, 1)$ -axis in the Brillouin zone is called Δ , we denote the representations by $\Delta_1, \dots, \Delta_5$ (cf. Table 1.3).

The degeneracies of the states at the Γ -point are partially lifted for \mathbf{k} along the Δ -line (cf. Table 1.4). It follows that five bands emanate from the three energy levels at the Γ -point, one of which

⁴Literature on point groups: Landau & Lifschitz: Vol. III Chapt. XII; Dresselhaus, Dresselhaus & Jorio, *Group Theory - Applications to the Physics of Condensed Matter*, Springer; Koster et al., *Properties of the thirty-two point groups*, MIT Press (Table book).

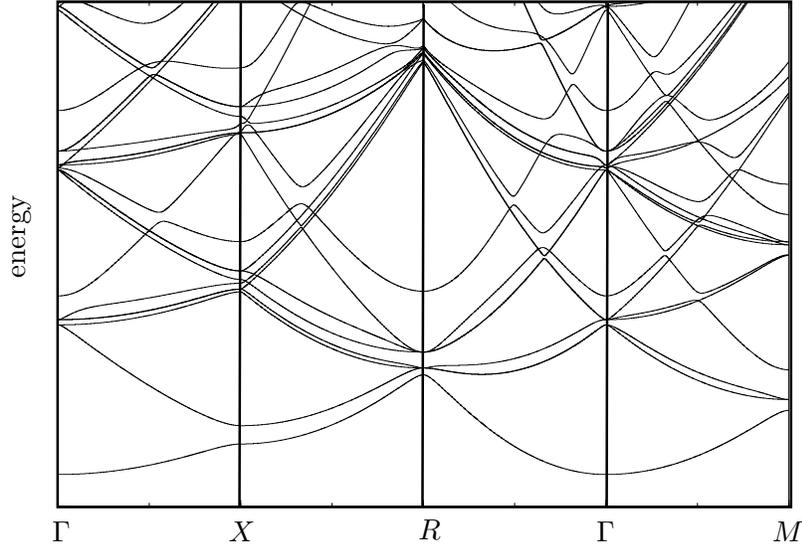


Figure 1.4: The band structure of the simple cubic lattice.

representation	base function
Δ_1	$1, z$
Δ_2	$xy(x^2 - y^2)$
Δ_3	$x^2 - y^2$
Δ_4	xy
Δ_5	$\{x, y\}$

Table 1.3: Irreducible representations of C_{4v} and their basis functions.

O_h	C_{4v}
Γ_1^+	Δ_1
Γ_3^+	$\Delta_1 \oplus \Delta_3$
Γ_4^-	$\Delta_1 \oplus \Delta_5$

Table 1.4: Lifting of degeneracy along the Δ -line.

is two-fold degenerate (Fig. 1.4).

X-point: Once we reach the Brillouin zone boundary at the X-point, the symmetry is larger than on the Δ -line, namely D_{4h} , the full tetragonal point group which for both parities has five irreducible representations, four of them one-dimensional, the remaining one two-dimensional (cf. Table 1.5). Note that C_{4v} is a subgroup of D_{4h} as well as D_{4h} is a subgroup of O_h . Furthermore, the inversion is an element of D_{4h} , as for the X-point \mathbf{k} is equivalent to $-\mathbf{k}$ ($\mathbf{k} - (-\mathbf{k}) = 2\mathbf{k} = \mathbf{G}$ is a reciprocal lattice vector).

The set of states with the lowest energy is equivalent to the problem discussed above in equations (1.23), (1.24) and (1.26). We consider $\mathbf{G}_1 = 0$ and $\mathbf{G}_2 = 2\pi(0, 0, 1)/a$ with energy $(\hbar^2/2m)(\pi/a)^2$ at the X-point. The levels are split into an (even) bonding state and an (odd) anti-bonding state

$$\begin{aligned}
 X_1^+ : E &= \frac{\hbar^2}{2m} \left(\frac{\pi}{a}\right)^2 - |V_{\mathbf{G}_2}|, & e^{iG_2z/2} \cos\left(\frac{G_2z}{2}\right), \\
 X_2^- : E &= \frac{\hbar^2}{2m} \left(\frac{\pi}{a}\right)^2 + |V_{\mathbf{G}_2}|, & e^{iG_2z/2} \sin\left(\frac{G_2z}{2}\right).
 \end{aligned} \tag{1.42}$$

even	base function	odd	base function
X_1^+	1	X_1^-	$xyz(x^2 - y^2)$
X_2^+	$xy(x^2 - y^2)$	X_2^-	z
X_3^+	$x^2 - y^2$	X_3^-	xyz
X_4^+	xy	X_4^-	$z(x^2 - y^2)$
X_5^+	$\{zx, zy\}$	X_5^-	$\{x, y\}$

Table 1.5: Irreducible representations of D_{4h} and their basis functions.

The next higher states are centered around $E = (\hbar^2/2m)(\sqrt{5}\pi/a)^2$ and belong to the next-to-nearest neighbors of the X -point in the reciprocal lattice. There are eight such points, namely

$$\begin{aligned}
\mathbf{G}_1 &= \frac{2\pi}{a}(1, 0, 0), & \mathbf{G}_2 &= \frac{2\pi}{a}(1, 0, 1), & \mathbf{G}_3 &= \frac{2\pi}{a}(-1, 0, 0), & \mathbf{G}_4 &= \frac{2\pi}{a}(-1, 0, 1), \\
\mathbf{G}_5 &= \frac{2\pi}{a}(0, 1, 0), & \mathbf{G}_6 &= \frac{2\pi}{a}(0, 1, 1), & \mathbf{G}_7 &= \frac{2\pi}{a}(0, -1, 0), & \mathbf{G}_8 &= \frac{2\pi}{a}(0, -1, 1).
\end{aligned} \tag{1.43}$$

To find the splitting of the energy levels, we project the base functions in (1.43) onto those of the irreducible representations and find the results displayed in Table 1.6. This analysis shows

representation	$u_{\mathbf{k}=\pi(0,0,1)/a}(\mathbf{r})$	degeneracy
X_1^+	$(\cos(Gx) + \cos(Gy))e^{iGz/2} \cos(Gz/2)$	1
X_3^+	$(\cos(Gx) - \cos(Gy))e^{iGz/2} \cos(Gz/2)$	1
X_5^+	$\{\sin(Gx)e^{-iGz/2} \sin(Gz/2), \sin(Gy)e^{iGz/2} \sin(Gz/2)\}$	2
X_2^-	$(\cos(Gx) + \cos(Gy))e^{iGz/2} \sin(Gz/2)$	1
X_4^-	$(\cos(Gx) - \cos(Gy))e^{iGz/2} \sin(Gz/2)$	1
X_5^-	$\{\sin(Gx)e^{iGz/2} \cos(Gz/2), \sin(Gy)e^{-iGz/2} \cos(Gz/2)\}$	2

Table 1.6: Projections of the base functions 1.43 onto the ones of the irreducible representations at the X -point ($G = 2\pi/a$).

that there are six energy levels where two of them are two-fold degenerate.

This kind of analysis can be applied to all symmetry lines, so that a good qualitative picture of the symmetries of the bands can be obtained. For more quantitative information, knowledge of the specific form of the periodic potential is necessary and also more advanced techniques beyond the nearly free electron approach are required. Nevertheless, the nearly free electron method can give important qualitative insights into the symmetry related properties of the band structure (see Fig. 1.4 for a full band structure).

1.4 $\mathbf{k} \cdot \mathbf{p}$ -expansion - effective masses

Near points of high symmetry in the Brillouin zone (such as the Γ -point), energy bands can be approximated by a quadratic dependence on k_μ in the general form

$$\epsilon_{\mathbf{k}} = \epsilon_0 + \frac{\hbar^2}{2m} \sum_{\mu, \nu} \left(\frac{m}{m^*} \right)_{\mu\nu} k_\mu k_\nu + \dots \tag{1.44}$$

We consider this expansion in a perturbative formulation. We expand the Hamiltonian (1.17) around $\mathbf{k} = 0$ (Γ -point) up to second order in \mathbf{k} and split it into the three parts

$$\begin{aligned}\mathcal{H}_0 &= \frac{\hat{\mathbf{p}}^2}{2m} + \hat{V}, \\ \mathcal{H}_1 &= \frac{\hbar}{m} \hat{\mathbf{p}} \cdot \mathbf{k} = \frac{\hbar}{m} \hat{\boldsymbol{\pi}} \cdot \mathbf{k}, \\ \mathcal{H}_2 &= \frac{\hbar^2 \mathbf{k}^2}{2m},\end{aligned}\tag{1.45}$$

where $\hat{\boldsymbol{\pi}}$ may, in general, have a more complicated form than in this example. We assume that the Hamiltonian \mathcal{H}_0 can be solved exactly and that \mathcal{H}_1 and \mathcal{H}_2 are small perturbations (small \mathbf{k}). Note that the latter is not an operator, but simply a \mathbf{k} -dependent contribution to the energy. For \mathcal{H}_0 we have

$$\mathcal{H}_0|n0\rangle = \epsilon_n|n0\rangle,\tag{1.46}$$

where $|n0\rangle$ are states at $\mathbf{k} = 0$ with the band index (quantum number) n . For simplicity, we take these states to be non-degenerate, so that Rayleigh-Schrödinger perturbation theory yields the perturbed energy

$$E_{\mathbf{k}} = \epsilon_n + \frac{\hbar^2 \mathbf{k}^2}{2m} + \frac{\hbar^2}{m^2} \sum_{n' \neq n} \sum_{\mu, \nu} \frac{\langle n0 | \hat{\pi}_\mu | n'0 \rangle \langle n'0 | \hat{\pi}_\nu | n0 \rangle}{\epsilon_n - \epsilon_{n'}} k_\mu k_\nu,\tag{1.47}$$

which can be translated into a mass-tensor of the form,

$$\left(\frac{m}{m^*}\right)_{\mu\nu} = \delta_{\mu\nu} + \frac{2}{m} \sum_{n' \neq n} \frac{\langle n0 | \hat{\pi}_\mu | n'0 \rangle \langle n'0 | \hat{\pi}_\nu | n0 \rangle}{\epsilon_n - \epsilon_{n'}}.\tag{1.48}$$

Thus, the electronic band structure in the vicinity of the Γ -point can be expressed by a mass-tensor. This approximation is valid for other symmetry points, too. Later, we will find this approximation very convenient when dealing with problems for which states around the upper or lower band edges are important which are often, but not always, high-symmetry points. Note that, at the band edges, all eigenvalues of the mass tensor have the same sign. There are other symmetry points (usually located at the boundary of the Brillouin zone) where the mass tensor has both positive and negative eigenvalues. These are called saddle points, which play an important role in connection with van Hove singularities in the density of states.

Note that, at symmetry points, the energy shift is linear in \mathcal{H}_1 , as

$$\langle n0 | \hat{\boldsymbol{\pi}} | n0 \rangle = 0.\tag{1.49}$$

This is because of parity and $\hat{\boldsymbol{\pi}}$ being a rank one tensor operator, as can be easily verified by noting that $\hat{\boldsymbol{\pi}} \cdot \mathbf{k}$ should be a scalar in Eq.(1.45).⁵

The resulting selection rules are important for the states $|n'0\rangle$ appearing in the matrix elements of second-order corrections, too. The eigenstates can also be approximated using the Rayleigh-Schrödinger method, resulting in

$$|n\mathbf{k}\rangle = e^{i\mathbf{k}\cdot\mathbf{r}} \left\{ |n0\rangle + \frac{\hbar}{m} \sum_{n' \neq n} |n'0\rangle \frac{\langle n'0 | \hat{\boldsymbol{\pi}} \cdot \mathbf{k} | n0 \rangle}{\epsilon_n - \epsilon_{n'}} \right\}.\tag{1.51}$$

⁵In this case, $\hat{P}\hat{\boldsymbol{\pi}}\hat{P} = -\hat{\boldsymbol{\pi}}$ holds for the parity operator \hat{P} . But $\hat{P}|n0\rangle = \pm|n0\rangle$ ($|n0\rangle$ is a parity eigenstate whenever the system has inversion symmetry, which carries over to the little group of $\mathbf{k} = 0$). Then,

$$\langle n0 | \hat{\boldsymbol{\pi}} | n0 \rangle = -\langle n0 | \hat{P}\hat{\boldsymbol{\pi}}\hat{P} | n0 \rangle = -\langle n0 | \hat{\boldsymbol{\pi}} | n0 \rangle,\tag{1.50}$$

so that the matrix element vanishes.

Finally, we discuss the case of degenerate levels at the Γ -point. As an example, we consider a three-fold degenerate level corresponding to the irreducible representation Γ_4^- with $|n_\mu 0\rangle$ ($\mu = x, y, z$). The Rayleigh-Schrödinger perturbation theory leads to the problem of diagonalizing the 3×3 -matrix

$$H_{\mu\nu} = \frac{1}{m^2 \Delta E} \langle n_\mu 0 | \hat{\pi} \cdot \mathbf{k} | n 0 \rangle \langle n 0 | \hat{\pi} \cdot \mathbf{k} | n_\nu 0 \rangle. \quad (1.52)$$

Here, we take into account only one virtual state $|n 0\rangle$ belonging to Γ_1 . It is easy to see, that the matrix has the form $H_{\mu\nu} = A k_\mu k_\nu$. The solution of the secular equation

$$\det(H_{\mu\nu} - E \delta_{\mu\nu}) = 0 \quad (1.53)$$

yields $E = 0$ (two-fold degenerate) and $E = A \mathbf{k}^2$ (non-degenerate), i.e.,

$$\epsilon_{\mathbf{k}} = \epsilon_0 + \begin{cases} \frac{\hbar^2}{2m} \mathbf{k}^2 + A \mathbf{k}^2, \\ \frac{\hbar^2}{2m} \mathbf{k}^2. \end{cases} \quad (1.54)$$

Out of the three-fold degenerate level bands with different effective masses evolve, two of which are degenerate. By means of this simple consideration we can determine the effective masses at the band edges and see how degeneracies are lifted in a given direction \mathbf{k} .

1.5 Band structures - approximate methods

While the approximation of nearly free electrons gives a qualitatively reasonable picture of the band structure, it rests on the assumption that the periodic potential is weak, and thus may be treated as a small perturbation. However, in reality the ionic potential is strong compared to the electrons' kinetic energy. This leads to strong modulations of the wave function around the ions, which is not well described by slightly perturbed plane waves.

1.5.1 Pseudo-potential

In order to overcome this weakness of the plane wave solution, we would have to superpose a very large number of plane waves, a task which is not easily put into practice. Alternatively, we can divide the electronic states into the ones corresponding to filled low-lying energy states, which are concentrated around the ionic core (core states), and into extended (and more weakly modulated) states, which form the valence and conduction bands. The core electron states may be approximated by atomic orbitals of isolated atoms. For a metal such as aluminum (Al: $1s^2 2s^2 2p^6 3s^2 3p$) the core electrons correspond to the $1s$ -, $2s$ -, and $2p$ -orbitals, whereas the $3s$ - and $3p$ -orbitals contribute dominantly to the extended states of the valence- and conduction bands. We will focus on the latter, as these determine the low-energy physics of the electrons. The core electrons are deeply bound and can be considered inert.

We introduce the core electron states as $|\phi_j\rangle$, with $\mathcal{H}|\phi_j\rangle = E_j|\phi_j\rangle$. The remaining states have to be orthogonal to these core states, so that we make the ansatz

$$|\phi_{n,\mathbf{k}}\rangle = |\chi_{n\mathbf{k}}\rangle - \sum_j |\phi_j\rangle \langle \phi_j | \chi_{n,\mathbf{k}} \rangle, \quad (1.55)$$

with $|\chi_{n,\mathbf{k}}\rangle$ an orthonormal set of states. Then, $\langle \phi_{n,\mathbf{k}} | \phi_j \rangle = 0$ holds for all j . We are free to choose plane waves for the $|\chi_{n\mathbf{k}}\rangle$. The resulting $|\phi_{n,\mathbf{k}}\rangle$ are then called orthogonalized plane waves (OPW). The Bloch functions are superpositions of these OPW,

$$|\psi_{n,\mathbf{k}}\rangle = \sum_{\mathbf{G}} b_{\mathbf{k}+\mathbf{G}} |\phi_{n,\mathbf{k}+\mathbf{G}}\rangle, \quad (1.56)$$

where the coefficients $b_{\mathbf{k}+\mathbf{G}}$ converge rapidly, such that, hopefully, only a small number of OPWs is needed for a good description.

In a first step, we consider an arbitrary $|\chi_{n\mathbf{k}}\rangle$ and insert it into the eigenvalue equation $\mathcal{H}|\phi_{n\mathbf{k}}\rangle = E_{n\mathbf{k}}|\phi_{n\mathbf{k}}\rangle$,

$$\begin{aligned} \Rightarrow \quad \mathcal{H}|\chi_{n\mathbf{k}}\rangle - \sum_j \mathcal{H}|\phi_j\rangle\langle\phi_j|\chi_{n,\mathbf{k}}\rangle &= E_{n\mathbf{k}} \left\{ |\chi_{n\mathbf{k}}\rangle - \sum_j |\phi_j\rangle\langle\phi_j|\chi_{n,\mathbf{k}}\rangle \right\} \\ \Rightarrow \quad \mathcal{H}|\chi_{n\mathbf{k}}\rangle + \sum_j [E_{n\mathbf{k}} - E_j]|\phi_j\rangle\langle\phi_j|\chi_{n,\mathbf{k}}\rangle &= E_{n\mathbf{k}}|\chi_{n\mathbf{k}}\rangle. \end{aligned} \quad (1.57)$$

We introduce the operator $\widehat{V}' = \sum_j [E_{n\mathbf{k}} - E_j]|\phi_j\rangle\langle\phi_j|$ (an integral operator in real space), describing a non-local and energy-dependent potential. With this operator we can rewrite the eigenvalue equation in the form

$$(\mathcal{H} + \widehat{V}_{\text{ps}})|\chi_{n,\mathbf{k}}\rangle = (\mathcal{H}_0 + \widehat{V} + \widehat{V}')|\chi_{n,\mathbf{k}}\rangle = E_{n\mathbf{k}}|\chi_{n\mathbf{k}}\rangle. \quad (1.58)$$

This is an eigenvalue equation for the so-called pseudo-wave function (or pseudo-state) $|\chi_{n\mathbf{k}}\rangle$, instead of the Bloch state $|\psi_{n\mathbf{k}}\rangle$, where the modified potential

$$\widehat{V}_{\text{ps}} = \widehat{V} + \widehat{V}' \quad (1.59)$$

is called *pseudo-potential*. The attractive core potential $\widehat{V} = V(\widehat{\mathbf{r}})$ is always negative. On the other hand, $E_{n\mathbf{k}} > E_j$, such that \widehat{V}' is positive. It follows that \widehat{V}_{ps} is weaker than both \widehat{V} and \widehat{V}' .

An arbitrary number of core states $\sum_j a_j|\psi_j\rangle$ may be added to $|\chi_{n\mathbf{k}}\rangle$ without violating the orthogonality condition (1.55). Consequently, neither the pseudo-potential nor the pseudo-states are uniquely determined and may be optimized variationally w.r.t. the $\{a_j\}$ in order to optimally reduce the spatial modulation of either the pseudo-potential or the wave-function.

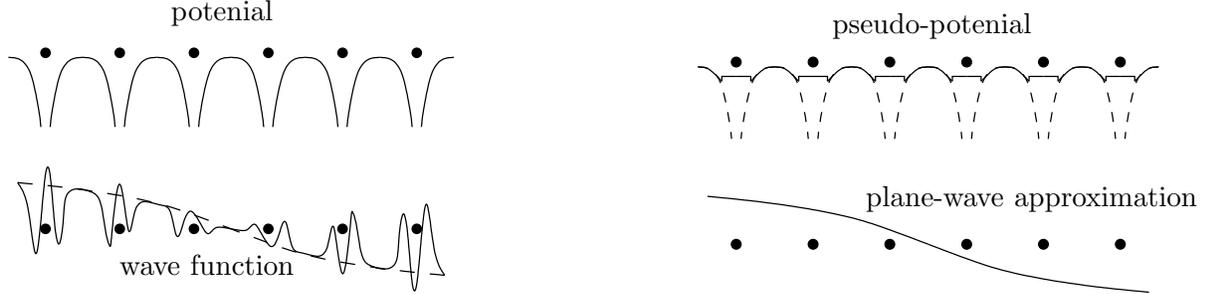


Figure 1.5: Illustration of the pseudo-potential.

If we are only interested in states inside a small energy window, the energy dependence of the pseudo-potential can be neglected, and V_{ps} may be approximated by a standard potential (see Fig. 1.5). Such a simple ansatz is exemplified by the atomic pseudo-potential, proposed by Ashcroft, Heine and Abarenkov (AHA). The potential of a single ion is assumed to be of the form

$$v_{\text{ps}}(r) = \begin{cases} V_0 & r < R_c, \\ -\frac{Z_{\text{ion}}e^2}{r} & r > R_c, \end{cases} \quad (1.60)$$

where Z_{ion} is the charge of the ionic core and R_c its effective radius (determined by the core electrons). The constants R_c and V_0 are chosen such that the energy levels of the outermost electrons are reproduced correctly for the single-atom case. For example, the $1s$ -, $2s$ -, and $2p$ -electrons of Na form the ionic core. R_c and V_0 are adjusted such that the one-particle problem

$\mathbf{p}^2/2m + v_{\text{ps}}(r)$ leads to the correct ionization energy of the 3s-electron. More flexible approaches allow for the incorporation of more experimental input into the pseudo-potential.

The full pseudo-potential can be constructed from the contribution of the individual atoms,

$$V_{\text{ps}}(\mathbf{r}) = \sum_n v_{\text{ps}}(\mathbf{r} - \mathbf{R}_n), \quad (1.61)$$

where \mathbf{R}_n is the lattice vector. For the method of nearly free electrons we need the Fourier transform of the potential evaluated at the reciprocal lattice vectors,

$$V_{\text{ps},\mathbf{G}} = \frac{1}{\Omega} \int d^3r V_{\text{ps}}(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}} = \frac{N}{\Omega} \int d^3r v_{\text{ps}}(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}}. \quad (1.62)$$

For the AHA form, this is given by

$$V_{\text{ps},\mathbf{G}} = -\frac{4\pi Z_{\text{ion}} e^2}{G^2} \left[\cos(GR_c) + \frac{V_0}{Z_{\text{ion}} e^2 G} \{ (R_c^2 G^2 - 2) \cos(GR_c) + 2 - 2R_c G \sin(GR_c) \} \right]. \quad (1.63)$$

For small reciprocal lattice vectors, the zeroes of the trigonometric functions on the RHS of (1.63) reduce the strength of the potential. For large G , the pseudo-potential decreases in any case. It is thus clear that the pseudo-potential is always weaker than the original potential.

For complex unit cells containing more than one atom the pseudo-potential may be written as

$$V_{\text{ps}}(\mathbf{r}) = \sum_{n\alpha} v_{\alpha}[\mathbf{r} - (\mathbf{R}_n + \mathbf{R}_{\alpha})], \quad (1.64)$$

where \mathbf{R}_{α} denotes the position of the α -th base atom in the unit cell. Here, v_{α} is the pseudo-potential of the α -th ion. In reciprocal space,

$$\begin{aligned} V_{\text{ps},\mathbf{G}} &= \frac{N}{\Omega} \sum_{\alpha} e^{-i\mathbf{G}\cdot\mathbf{R}_{\alpha}} \int d^3r v_{\alpha}(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}} \\ &= \sum_{\alpha} e^{-i\mathbf{G}\cdot\mathbf{R}_{\alpha}} F_{\alpha,\mathbf{G}}. \end{aligned} \quad (1.65)$$

The form factor $F_{\alpha,\mathbf{G}}$ contains the information of the base atoms and may be calculated or fitted to experiments.

1.5.2 APW-method (augmented plane wave)

Next, we consider a method introduced by Slater in 1937. It is an extension of the so-called Wigner-Seitz cell method (1933) and consists of approximating the crystal potential by a so-called muffin-tin potential. This is a periodic potential, which is taken to be spherically symmetric and position dependent around each atom up to a distance r_s , and constant for larger distances. The spheres of radius r_s are taken to be non-overlapping and are contained completely in the Wigner-Seitz cell (Fig. 1.6).⁶ It is the advantage of this decomposition that the problem can be solved using a divide-and-conquer strategy. Inside the muffin-tin radius we solve the spherically symmetric problem, while the solutions on the outside are given by plane waves; the remaining task is to match the solutions at the boundaries.

The spherically symmetric problem for $|\mathbf{r}| < r_s$ is solved with the standard ansatz

$$\varphi(\mathbf{r}) = \frac{u_l(r)}{r} Y_{lm}(\theta, \phi), \quad (1.66)$$

⁶The Wigner-Seitz cell is the analogue of the Brillouin zone in real space. One draws planes cutting each line joining two atoms in the middle, and orthogonal to them. The smallest cell bounded by these planes is the Wigner-Seitz cell.

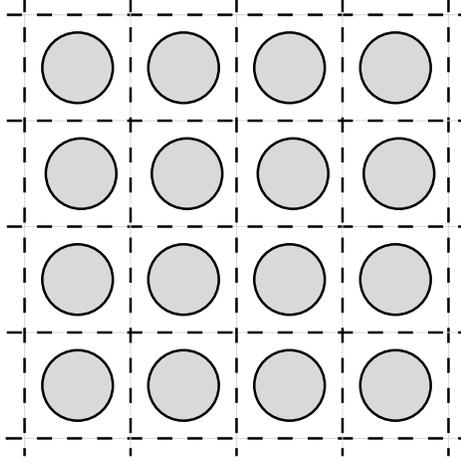


Figure 1.6: Muffin-tin potential.

where the radial part obeys the differential equation

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dr^2} + \frac{\hbar^2 l(l+1)}{2mr^2} + V(r) - E \right] u_l(r, E) = 0. \quad (1.67)$$

We define an augmented plane wave (APW) $A(\mathbf{k}, \mathbf{r}, E)$, which, for \mathbf{r} outside the Muffin-tin sphere, is a pure plane wave with wave vector \mathbf{k} . For this, we employ the representation of plane waves by spherical harmonics,

$$e^{i\mathbf{k}\cdot\mathbf{r}} = 4\pi \sum_{l,m} i^l j_l(kr) Y_{lm}^*(\hat{\mathbf{k}}) Y_{lm}(\hat{\mathbf{r}}), \quad (1.68)$$

where $j_l(x)$ is the l -th spherical Bessel function. We parametrize

$$A(\mathbf{k}, \mathbf{r}, E) = \begin{cases} \frac{4\pi}{\sqrt{\Omega_{\text{UC}}}} \sum_{l,m} i^l j_l(kr_s) \frac{r_s u_l(r, E)}{r u_l(r_s, E)} Y_{lm}^*(\hat{\mathbf{k}}) Y_{lm}(\hat{\mathbf{r}}), & r < r_s, \\ \frac{4\pi}{\sqrt{\Omega_{\text{UC}}}} \sum_{l,m} i^l j_l(kr) Y_{lm}^*(\hat{\mathbf{k}}) Y_{lm}(\hat{\mathbf{r}}), & r > r_s, \end{cases} \quad (1.69)$$

where Ω_{UC} is the volume of the unit cell. Note that the wave function is continuous at $r = r_s$, but that its derivatives are not continuous in general. We can use an expansion similar to the one in the nearly free electron approximation,

$$\psi_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} a_{\mathbf{G}}(\mathbf{k}) A(\mathbf{k} + \mathbf{G}, \mathbf{r}, E), \quad (1.70)$$

where the \mathbf{G} are reciprocal lattice vectors. The unknown coefficients can be determined variationally by solving the system of equations

$$\sum_{\mathbf{G}} \langle A_{\mathbf{k}}(E) | \mathcal{H} - E | A_{\mathbf{k}+\mathbf{G}}(E) \rangle a_{\mathbf{G}}(\mathbf{k}) = 0, \quad (1.71)$$

where

$$\langle A_{\mathbf{k}}(E) | \mathcal{H} - E | A_{\mathbf{k}'}(E) \rangle = \left(\frac{\hbar^2 \mathbf{k} \cdot \mathbf{k}'}{2m} - E \right) \Omega_{\text{UC}} \delta_{\mathbf{k}, \mathbf{k}'} + V_{\mathbf{k}, \mathbf{k}'} \quad (1.72)$$

with

$$V_{\mathbf{k},\mathbf{k}'} = 4\pi r_s^2 \left\{ - \left(\frac{\hbar^2 \mathbf{k} \cdot \mathbf{k}'}{2m} - E \right) \frac{j_1(|\mathbf{k} - \mathbf{k}'| r_s)}{|\mathbf{k} - \mathbf{k}'|} + \sum_{l=0}^{\infty} \frac{\hbar^2}{2m} (2l+1) P_l(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}}') j_l(k r_s) j_l(k' r_s) \frac{u'_l(r_s, E)}{u_l(r_s, E)} \right\}. \quad (1.73)$$

Here, $P_l(z)$ is the l -th Legendre polynomial and $u' = du/dr$. The solution of (1.71) yields the energy bands. The most difficult parts are the approximation of the crystal potential by the muffin-tin potential and the computation of the matrix elements in (1.71). The rapid convergence of the method is its big advantage: just a few dozens of \mathbf{G} -vectors are needed and the largest angular momentum needed is roughly $l \approx 5$. Another positive aspect is the fact that the APW-method allows to interpolate between the two extremes of extended, weakly bound electronic states and tightly bound states.

1.6 Tightly bound electrons and Wannier functions

If the electrons in the valence and conduction bands are strongly bound to the ions, another very efficient approximation to the band structure exists. In this case, it is easier to approach the problem in real space instead of reciprocal space. This leads to the so-called tight-binding model.

We introduce the Wannier functions as 'Fourier transforms' of the Bloch functions,

$$\psi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_j e^{i\mathbf{k} \cdot \mathbf{R}_j} w(\mathbf{r} - \mathbf{R}_j), \quad (1.74)$$

where $w(\mathbf{r} - \mathbf{R}_j)$ is the Wannier function centered around the j -th atom. There is a Wannier function for each atomic orbital. For the sake of simplicity, we restrict ourselves to the case of one orbital per atom. The Wannier function obeys the orthogonality relation

$$\int d^3r w^*(\mathbf{r} - \mathbf{R}_j) w(\mathbf{r} - \mathbf{R}_l) = \delta_{jl}. \quad (1.75)$$

We may assume the one-particle Hamiltonian to be of the form $\mathcal{H} = -\hbar^2 \nabla^2 / 2m + V(\mathbf{r})$, with a periodic potential $V(\mathbf{r})$. Then,

$$\epsilon_{\mathbf{k}} = \int d^3r \psi_{\mathbf{k}}^*(\mathbf{r}) \mathcal{H} \psi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{N} \sum_{j,l} e^{-i\mathbf{k} \cdot (\mathbf{R}_j - \mathbf{R}_l)} \int d^3r w^*(\mathbf{r} - \mathbf{R}_j) \mathcal{H} w(\mathbf{r} - \mathbf{R}_l), \quad (1.76)$$

with

$$\epsilon_0 = \int d^3r w^*(\mathbf{r} - \mathbf{R}_j) \mathcal{H} w(\mathbf{r} - \mathbf{R}_j), \quad (1.77)$$

$$t_{jl} = \int d^3r w^*(\mathbf{r} - \mathbf{R}_j) \mathcal{H} w(\mathbf{r} - \mathbf{R}_l) \quad \text{for } j \neq l. \quad (1.78)$$

It follows immediately that the band energy may be written as a discrete sum,

$$\epsilon_{\mathbf{k}} = \epsilon_0 + \frac{1}{N} \sum_{j,l} t_{jl} e^{-i\mathbf{k} \cdot (\mathbf{R}_j - \mathbf{R}_l)} = \epsilon_0 + \sum_l t_{0l} e^{i\mathbf{k} \cdot \mathbf{R}_l}, \quad (1.79)$$

where $\mathbf{R}_0 = 0$ is assumed. It is obvious that $\epsilon_{\mathbf{k}+\mathbf{G}} = \epsilon_{\mathbf{k}}$. The quantities t_{jl} are called *hopping matrix elements*. It is possible to construct an effective Hamiltonian based on the above findings, which describes the band structure of independent electrons, as

$$\mathcal{H} = \sum_{i,j} \sum_s t_{ij} c_{is}^\dagger c_{js}, \quad (1.80)$$

where c_{is} (c_{is}^\dagger) annihilates (creates) an electron with spin s at the lattice point i . The Hamiltonian describes the hopping of electrons from site j to site i . This formulation is advantageous, if the hopping matrix elements fall off rapidly with the distance between the lattice points. This should be the case for electronic states which are tightly bound to the ions.

Consider a simple cubic lattice, assuming that $t_{jl} = -t$ for nearest neighbors and zero otherwise. The band energy follows from a Fourier transform and is given by

$$\epsilon_{\mathbf{k}} = \epsilon_0 - 2t\{\cos k_x a + \cos k_y a + \cos k_z a\}, \quad (1.81)$$

where a is the lattice constant. The same can be applied to more complicated lattices and systems with several relevant orbitals per atom.

1.7 Semi-classical description of band electrons

In quantum mechanics, the Ehrenfest theorem shows that the expectation values of the position and momentum operators obey equations similar to the equation of motion in Newtonian mechanics. An analogous formulation holds for electrons in a periodic potential, where we assume that the electron may be described as a wave packet of the form

$$\psi_{\mathbf{k}}(\mathbf{r}, t) = \sum_{\mathbf{k}'} g_{\mathbf{k}}(\mathbf{k}') e^{i\mathbf{k}' \cdot \mathbf{r} - i\epsilon_{\mathbf{k}'} t}, \quad (1.82)$$

where $g_{\mathbf{k}}(\mathbf{k}')$ is centered around \mathbf{k} in reciprocal space and has a width of Δk . Δk should be much smaller than the size of the Brillouin zone for the ansatz to make sense, i.e., $\Delta k \ll 2\pi/a$, such that the wave packet is spread over many unit cells of the lattice since Heisenberg's uncertainty principle $\Delta k \Delta x > 1$ implies $\Delta x \gg a/2\pi$. In this way, the pseudo-momentum \mathbf{k} of the wave packet remains well defined. Furthermore, the applied electric and magnetic fields have to be small enough not to induce transitions between different bands. The latter condition is not very restrictive in practice.

1.7.1 Semi-classical equations of motion

We introduce the rules of the semi-classical motion of electrons with applied electric and magnetic fields without proof:

- The band index of an electron is conserved, i.e., there are no transitions between the bands.
- The equations of motion read

$$\dot{\mathbf{r}} = \mathbf{v}_n(\mathbf{k}) = \frac{1}{\hbar} \frac{\partial \epsilon_n \mathbf{k}}{\partial \mathbf{k}}, \quad (1.83)$$

$$\hbar \dot{\mathbf{k}} = -e\mathbf{E}(\mathbf{r}, t) - \frac{e}{c} \mathbf{v}_n(\mathbf{k}) \times \mathbf{H}(\mathbf{r}, t).$$

- All electronic states have a wave vector that lies in the first Brillouin zone, as \mathbf{k} and $\mathbf{k} + \mathbf{G}$ label the same state for all reciprocal lattice vectors \mathbf{G} .
- In thermal equilibrium, the electron density in the n -th band in the volume element $d^3 k$ around \mathbf{k} is given by

$$n_F[\epsilon_n(\mathbf{k})] 2 \frac{d^3 k}{(2\pi)^3} = \frac{1}{e^{[\epsilon_n(\mathbf{k}) - \mu]/k_B T} + 1} 2 \frac{d^3 k}{(2\pi)^3}, \quad (1.84)$$

where the factor 2 is due to the two spin states of an electron. Each state of given \mathbf{k} and spin can be occupied only once (Pauli principle).

Note that $\hbar\mathbf{k}$ is not the momentum of the electron, but the so-called lattice momentum or pseudo momentum in the Bloch theory of bands. It is connected with the eigenvalue of the translation operator on the state. Consequently, the right hand side of the second equation in (1.83) is not the force that acts on the electron, as the forces exerted by the periodic lattice potential is not included. The latter effect is contained implicitly through the form of the band energy $\epsilon(\mathbf{k})$, which governs the first equation.⁷

A plausibility argument concerning the conservation of energy leads to the second equation in (1.83). The time derivative of the energy (kinetic and potential)

$$E = \epsilon_n(\mathbf{k}(t)) - e\phi(\mathbf{r}(t)) \quad (1.87)$$

has to vanish, i.e.,

$$0 = \frac{dE}{dt} = \frac{\partial\epsilon_n(\mathbf{k})}{\partial\mathbf{k}} \cdot \dot{\mathbf{k}} - e\nabla\phi \cdot \dot{\mathbf{r}} = \mathbf{v}_n(\mathbf{k}) \cdot \left\{ \hbar\dot{\mathbf{k}} - e\nabla\phi \right\}. \quad (1.88)$$

From this, Eq. (1.83) follows directly for the electric field $\mathbf{E} = -\nabla\phi$ and the Lorentz force is allowed because the force is always perpendicular to the velocity \mathbf{v}_n .

1.7.2 Current densities

Later, we will see that homogenous steady (current carrying) states of electron systems can be described by the momentum distribution $n(\mathbf{k})$. The current density follows from

$$\mathbf{j} = -2e \int_{\text{BZ}} \frac{d^3k}{(2\pi)^3} \mathbf{v}(\mathbf{k}) n(\mathbf{k}) = -2e \int_{\text{BZ}} \frac{d^3k}{(2\pi)^3} n(\mathbf{k}) \frac{1}{\hbar} \frac{\partial\epsilon(\mathbf{k})}{\partial\mathbf{k}}, \quad (1.89)$$

where the integral extends over all \mathbf{k} in the Brillouin zone (BZ) and the factor 2 once again originates in the two spin states of the electrons. Note that for a finite current density $n(\mathbf{k})$ has to deviate from the equilibrium Fermi-Dirac distribution in Eq. (1.88). It is obvious that for an empty band the current density vanishes. The same holds true for a completely filled band, as $n(\mathbf{k}) = 1$ for all \mathbf{k} implies

$$\mathbf{j} = -2e \int_{\text{BZ}} \frac{d^3k}{(2\pi)^3} \frac{1}{\hbar} \frac{\partial\epsilon(\mathbf{k})}{\partial\mathbf{k}} = 0 \quad (1.90)$$

because $\epsilon(\mathbf{k})$ is periodic in the Brillouin zone, i.e., $\epsilon(\mathbf{k} + \mathbf{G}) = \epsilon(\mathbf{k})$ when \mathbf{G} is a reciprocal lattice vector. Thus, neither empty nor completely filled bands can carry currents.

An interesting aspect of band theory is the picture of holes. We compute the current density for a partially filled band in the framework of the semi-classical approximation,

$$\begin{aligned} \mathbf{j} &= -e \int_{\text{BZ}} \frac{d^3k}{4\pi^3} n(\mathbf{k}) \mathbf{v}_n(\mathbf{k}) = -e \left\{ \int_{\text{BZ}} \frac{d^3k}{4\pi^3} \mathbf{v}(\mathbf{k}) - \int_{\text{BZ}} \frac{d^3k}{4\pi^3} [1 - n(\mathbf{k})] \mathbf{v}(\mathbf{k}) \right\} \\ &= +e \int_{\text{BZ}} \frac{d^3k}{4\pi^3} [1 - n(\mathbf{k})] \mathbf{v}(\mathbf{k}). \end{aligned} \quad (1.91)$$

⁷*Bloch oscillation:* The fact that the band energy is a periodic function of \mathbf{k} leads to a strange oscillatory behavior. As a one-dimensional example we consider the band energy $\epsilon_k = -2 \cos ka$, which leads to the following solution of the semi-classical equations (1.83) in the presence of a homogenous electric field E ,

$$\hbar\dot{k} = -eE \quad \Rightarrow \quad k = -\frac{eEt}{\hbar} \quad \Rightarrow \quad \dot{x} = -\frac{2a}{\hbar} \sin\left(\frac{eEat}{\hbar}\right), \quad (1.85)$$

and it follows that the position x of the electron oscillates,

$$x(t) = \frac{1}{eE} \cos\left(\frac{eEat}{\hbar}\right). \quad (1.86)$$

This behavior is called Bloch oscillation and implies that the electron oscillates around its initial position rather than moving in one direction when subjected to an electric field. This effect can be observed under very special conditions only, it is easily destroyed by damping or scattering.

This suggests that we can determine the current density from the electrons in filled states with charge $-e$ or from the 'holes', missing electrons carrying positive charge in the unoccupied states. Both descriptions are equivalent in band theory. However, usually it is easier to work with holes if a band is almost filled, and with electrons if the filling is small.

1.8 Metals and semiconductors

Each state $|\psi_{n,\mathbf{k}}\rangle$ can be occupied by two electrons, one with spin state $|\uparrow\rangle$ and $|\downarrow\rangle$. In the ground state, all states up to the Fermi energy are filled. The nature of the ground state of electrons in a solid depends on the number of electrons per atom. Usually, this number is an integer, so that in the simplest cases we distinguish only two different situations: Firstly, the bands can be either completely filled or empty if the number of electrons per atom is even. In this case, the Fermi energy lies in a band gap (cf. Fig. 1.7), and a finite energy is needed to add or remove an electron, or to excite electrons. If the band gap $\Delta \ll$ bandwidth, we call the material a *semiconductor*, for $\Delta \sim$ bandwidth, it is an *insulator*. In both cases, for temperatures $k_B T \ll \Delta$, by the application of small electric voltages no electron motion is possible (no electric transport). The highest filled band is called *valence band*, the lowest empty band *conduction band*. Note that we will later encounter another form of an insulator, the Mott insulator, whose insulating behavior is not governed by a band structure effect (kinetic energy), but by a correlation effect through strong Coulomb interaction. Secondly, if the number of electrons per atom is odd, one band is half filled (see Fig. 1.7). Then the system is a *metal*, as charges can be moved without overcoming a band gap and electrons can be excited with arbitrarily small energy. The electrons remain mobile down to arbitrarily low temperatures. The prime example of a metal are the Alkali metals in the first column of the periodic table (Li, Na, K, Rb, Cs), as all of them have the configuration [noble gas] $(ns)^1$, i.e., one mobile electron per ion.

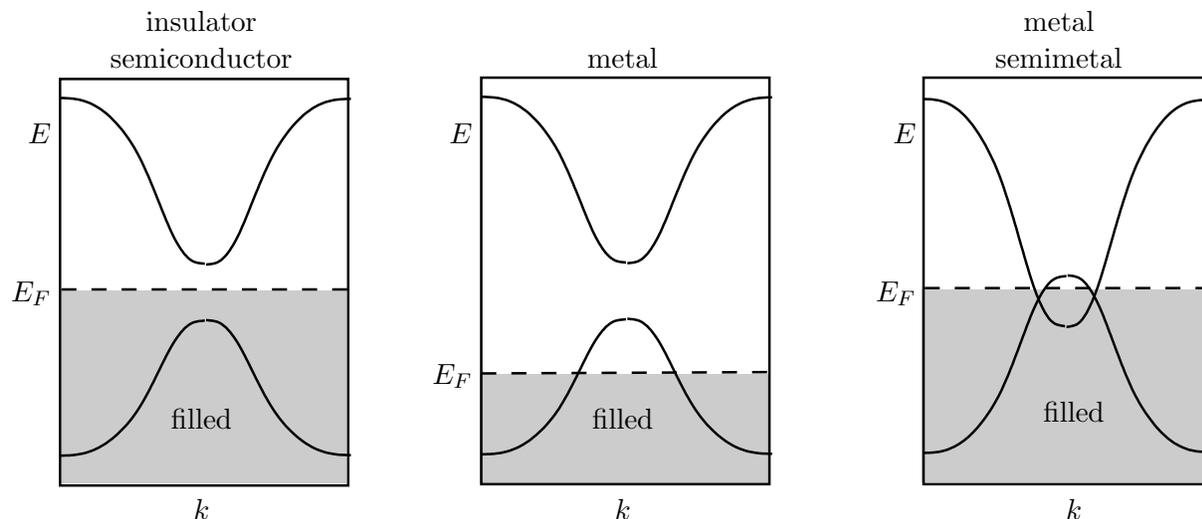


Figure 1.7: Material classes according to band filling: left panel: insulator or semiconductor (Fermi level in band gap); center panel: metal (Fermi level inside band); right panel: metal or semimetal (Fermi level inside two overlapping bands).

In general, band structures are more complex. Different bands need not be separated by energy gaps, but can overlap instead. In particular, this happens if different orbitals are involved in the structure of the bands. In these systems bands can have any fractional filling (not just filled or half-filled). The earth alkaline metals are an example for this (second column of the periodic table, Be, Mg, Ca, Sr, Ba), which are metallic in spite of having two (n, s) -electrons per unit cell. In cases where two bands overlap at the Fermi energy but the overlap is small, we call it a

semi-metal. An extreme case is graphite, where valence and conduction band touch in isolated points, so that there are no electrons at the Fermi energy, but the band gap is zero.

Chapter 2

Semiconductors

The technological relevance of semiconductors can hardly be overstated. In this chapter, we review some of their basic properties. Regarding the electric conductivity, semiconductors are placed in between metals and insulators. Normal metals are good conductors at all temperatures, and the conductivity usually increases with decreasing temperature. On the other hand, for semiconductors and insulators the conductivity decreases upon cooling (see Fig. 2.1).

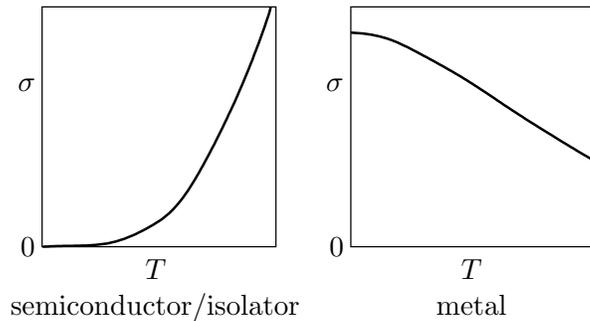


Figure 2.1: Temperature dependence of the electric conductivity for semiconductors and metals.

Below, we will see that the conductivity may be written as

$$\sigma = \frac{ne^2\tau}{m}, \quad (2.1)$$

where n is the density of (mobile) electrons, τ is the average time between two scattering events of the electrons, and m and e are the electronic mass and charge, respectively. In metals, n is independent of temperature, whereas τ decreases with increasing temperature. The latter thus determines the temperature dependence of the conductivity. In insulators and semiconductors, there are no mobile charges at $T = 0$. At finite temperature, they are induced by thermal excitations from the valence band to the conduction band, yielding

$$n \sim 10^{20} \text{cm}^{-3} \left(\frac{T}{300\text{K}} \right)^{3/2} e^{-E_g/2k_B T}, \quad (2.2)$$

where E_g is the band gap.¹ For insulators, the energy gap is huge, e.g., 5.5 eV for diamond. Consequently, the charge carrier density is around $n \sim 10^{-73} \text{cm}^{-3}$ at room temperature (300K).

¹Actually, one has to count both the excited electrons in the conduction band and the resulting holes in the valence band, as both contribute to the current,

$$\mathbf{j} = (\sigma_+ + \sigma_-)\mathbf{E}, \quad \text{with} \quad \sigma_{\pm} = \frac{n_{\pm}e^2\tau_{\pm}}{m_{\pm}}, \quad (2.3)$$

where $+$ and $-$ stand for holes and electrons, respectively, and $n_+ = n_-$ holds for thermal excitation. Note that, in general, the effective masses and scattering times are not the same for the valence and conduction bands.

By contrast, for semiconductors with an energy gap of 0.5 eV - 1 eV, the density is in the range $n \sim 10^3 \text{cm}^{-3}$ - 10^{11}cm^{-3} , resulting in a conductivity larger by orders of magnitude. However, the conductivities of both are dwarfed by the metallic conductors ($n \sim 20^{23} \text{cm}^{-3}$ - 10^{24}cm^{-3}). Adding a small amount of impurities, a process called doping with acceptors or donators, the conductivity of semiconductors can be engineered in various ways, rendering them indispensable as components in innumerable applications.

2.1 The band structure of the elements in group IV

2.1.1 Crystal structure and band structure

The most important semiconductor for applications is silicon (Si) that - like carbon (C), germanium (Ge) and tin (Sn) - belongs to the group IV of the periodic table. These elements have four electrons in their outermost shell in the configuration $(ns)^2(np)^2$ ($n = 2$ for C, $n = 3$ for Si, $n = 4$ for Ge, and $n = 5$ for Sn). All elements form crystals with a diamond structure (cf. Fig. reffig:diamond), i.e., a face-centered cubic lattice with a unit cell containing two atoms at $(0, 0, 0)$ and $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ (for Sn this is called α -Sn). The crystal structure is stabilized by hybridization of the four valence electrons, leading to covalent bonding of oriented orbitals,

$$\begin{aligned} |\psi_1\rangle &= |ns\rangle + |np_x\rangle + |np_y\rangle + |np_z\rangle, & |\psi_2\rangle &= |ns\rangle + |np_x\rangle - |np_y\rangle - |np_z\rangle, \\ |\psi_3\rangle &= |ns\rangle - |np_x\rangle + |np_y\rangle - |np_z\rangle, & |\psi_4\rangle &= |ns\rangle - |np_x\rangle - |np_y\rangle + |np_z\rangle. \end{aligned} \quad (2.4)$$

Locally, the neighbors of an atom form a tetrahedron around it, which leads to the diamond structure of the lattice.

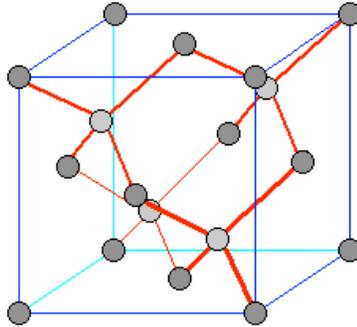


Figure 2.2: Crystal structure of diamond (face-centered cubic).

A simplified picture of the band structure around the Γ -point can be obtained by applying the free-electron approximation discussed in the last chapter. There is a parabolic band centered around the center of the Brillouin zone $(0, 0, 0)$ (the Brillouin zone of the FCC lattice is illustrated in Fig. 2.3). The corresponding representation is Γ_1 . The next multiplet with an energy of $\epsilon = 6\pi^2\hbar^2/ma^2$ derives from the parabolic bands emanating from the neighboring Brillouin zones, with $\mathbf{G} = (2\pi/a)(\pm 1, \pm 1, \pm 1)$. Note that the reciprocal lattice of a face-centered cubic lattice is body-centered cubic. The eight states are split into $\Gamma_1 \oplus \Gamma_2 \oplus \Gamma_4 \oplus \Gamma_5$. The order of the resulting energies can be obtained from band structure calculations, yielding $\epsilon_{\Gamma_5} < \epsilon_{\Gamma_4} < \epsilon_{\Gamma_2} < \epsilon_{\Gamma_1}$. The essential elements of the low-energy band structure of C and Si are shown in Fig. 2.4.

There are eight electrons per unit cell. It follows that the bands belonging to Γ_1 (non-degenerate) and Γ_5 (threefold degenerate) are completely filled. The maximum of the valence band is located at the Γ -point and belongs to Γ_5 . Because of the existence of an energy gap between valence

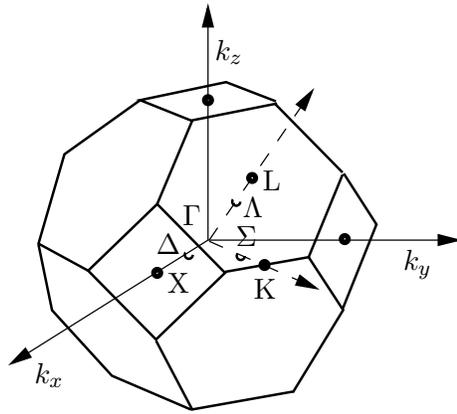


Figure 2.3: Brillouin zone of the face-centered cubic lattice.

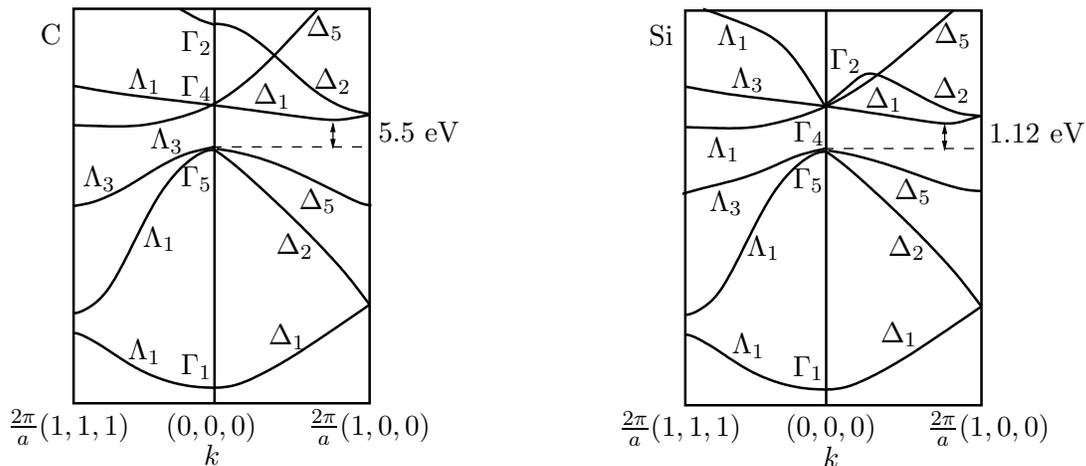


Figure 2.4: Band structure of C and Si.

band and conduction band, the system is a semiconductor. The gap is *indirect*, meaning that the minimum of the conduction band and the maximum of the valence band lie at different points in the Brillouin zone, i.e., the gap is minimal between the Γ -point of the valence band and some finite momentum $\hbar\mathbf{k}_0$ along the $[100]$ -direction of the conduction band.²

Here are some facts about semiconductors:

- Carbon has an energy gap of around 5.5eV in the diamond structure. The large energy gap causes the transparency of diamond in the visible range ($\sim 1.5\text{eV} - 3.5\text{eV}$), as the electromagnetic energy in this range cannot be absorbed by the electrons.
- The energy gap of silicon is 1.12eV and thus much smaller; furthermore, it is indirect.
- Germanium has an indirect gap of 0.67eV.
- GaAs is another important semiconductor, composed of one element of the third and fifth group, respectively. In contrast to C, Si, and Ge, the energy gap is direct.

In the following, we illustrate the fact that C, Si and Ge are semiconductors by investigating the bonding between neighboring atoms. All oriented bonds in the diamond structure are

²Energy gaps in semiconductors and insulators are said to be *direct* if the wave-vector connecting the maximum of the valence band and the minimum of the conduction band vanishes. Otherwise a gap is called *indirect* (see Fig. 2.5).

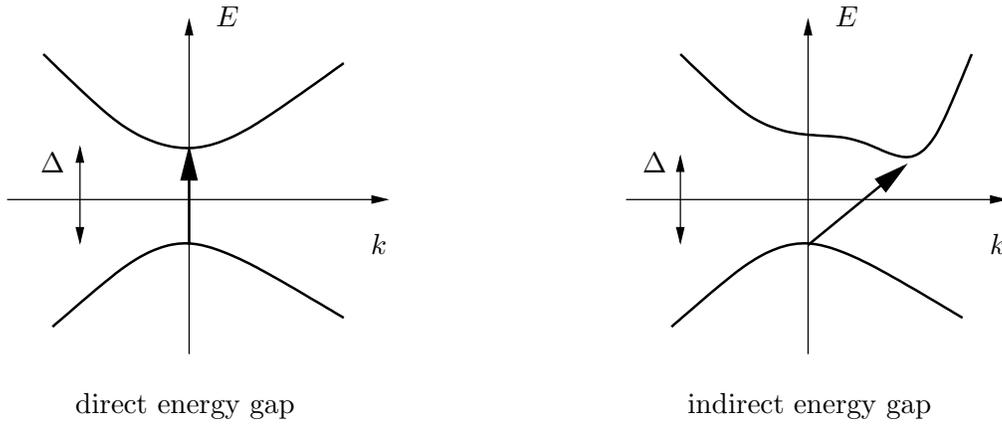


Figure 2.5: Illustration of direct and indirect band gaps.

covalent and consist of two electrons (cf. Fig. 2.6). The bonding can be described by molecular (Hund-Mullikan) orbitals, i.e., as a superposition of two orbitals $\psi_A(\mathbf{r})$ and $\psi_B(\mathbf{r})$, belonging to neighboring atoms (A and B). They correspond to the oriented orbitals in (2.4), and are essentially Wannier wave-functions. For the molecular orbitals it follows that

$$\Psi_{\pm}(\mathbf{r}) = \frac{1}{\sqrt{2}} \{ \psi_A(\mathbf{r}) \pm \psi_B(\mathbf{r}) \}. \quad (2.5)$$

HOLPRIGER ABSATZ

The two electrons in the molecular orbital form a spin-singlet in the bonding state $\Psi_+(\mathbf{r})$ with lower energy than the anti-bonding state $\Psi_-(\mathbf{r})$, which remains empty. The electrons in the bonding state remain localized: in order to be mobile an electron has to occupy the anti-bonding states, as all bonding states are occupied. The energy gap derives from the energy cost to populate the anti-bonding state. The ratio of the kinetic energy gain of a mobile electron to the energy difference between the bonding and anti-bonding state determines the size of the gap, which may even vanish.

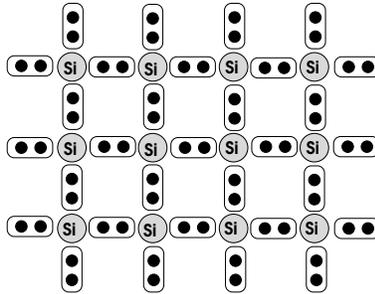


Figure 2.6: Schematic electronic structure of a semiconductor with covalent bonds.

2.1.2 $\mathbf{k} \cdot \mathbf{p}$ - approximation and effective mass

The band structure in the vicinity of the band edges can be very well described using the $\mathbf{k} \cdot \mathbf{p}$ method, as we show now for silicon. First, we consider the maximum of the valence band at $\mathbf{k} = 0$ (Γ -point), with electronic states

$$\{|yz\rangle, |zx\rangle, |xy\rangle\} \quad (2.6)$$

belonging to the representation Γ_5^+ . By symmetry considerations we obtain, in second order perturbation theory, the secular equation for the degenerate subspace,

$$\det \begin{bmatrix} ak_x^2 + b(k_y^2 + k_z^2) - E & ck_x k_y & ck_x k_z \\ ck_x k_y & ak_y^2 + b(k_x^2 + k_z^2) - E & ck_y k_z \\ ck_x k_z & ck_y k_z & ak_z^2 + b(k_x^2 + k_y^2) - E \end{bmatrix} = 0. \quad (2.7)$$

The general form of the eigenvalues is complicated, but it can be shown that the threefold degeneracy of the energies is lifted when moving away from the Γ -point. On the Δ - (C_{4v}) and Λ -lines (C_{3v}), which have higher symmetry, there is one twofold degenerate and one non-degenerate band (cf. Fig. 2.4):

$$\Delta\text{-line: } \mathbf{k} = k(1, 0, 0), \quad E_1(\Delta_2) = ak^2, \quad E_{2,3}(\Delta_5) = bk^2, \quad (2.8)$$

$$\Lambda\text{-line: } \mathbf{k} = \frac{k}{\sqrt{3}}(1, 1, 1), \quad E_1(\Lambda_1) = (a + 2b + 2c)k^2, \quad E_{2,3}(\Lambda_3) = (a + 2b - c)k^2,$$

where Δ_i and Λ_i are irreducible representations of the point group C_{4v} and C_{3v} , respectively.³ The minimum of the conduction band is located on the Δ -line at $\mathbf{k}_0 = k_0(1, 0, 0)$ with $k_0 \approx 0.8\overline{\Gamma X}$. Apart from spin, the corresponding band is non-degenerate. It follows that the $\mathbf{k} \cdot \mathbf{p}$ -approximation is given by

$$E_{\mathbf{k}} = a'(k_x - k_0)^2 + b'(k_y^2 + k_z^2), \quad (2.9)$$

due to the symmetry around $\mathbf{k}_0 = k_0(1, 0, 0)$. The electronic properties are determined by the states close to the band edges, so that these approximations play an important role in the physics of semiconductors.

2.2 Electronic properties and elementary excitations

We consider a simple two-band model to illustrate the most basic properties of the excitation spectrum of a semiconductor. The Hamiltonian is given by

$$\mathcal{H} = \sum_{\mathbf{k}, s} \epsilon_{V, \mathbf{k}} \hat{c}_{V, \mathbf{k}s}^\dagger \hat{c}_{V, \mathbf{k}s} + \sum_{\mathbf{k}, s} \epsilon_{C, \mathbf{k}} \hat{c}_{C, \mathbf{k}s}^\dagger \hat{c}_{C, \mathbf{k}s}, \quad (2.10)$$

where $\epsilon_{V, \mathbf{k}}$ and $\epsilon_{C, \mathbf{k}}$ are the band energies of the valence band and conduction band, respectively. The operator $c_{n\mathbf{k}s}^\dagger$ ($c_{n\mathbf{k}s}$) creates (annihilates) an electron with (pseudo-)momentum \mathbf{k} and spin s in the band n , $n = V, C$. In the ground state $|\Phi_0\rangle$,

$$|\Phi_0\rangle = \prod_{\mathbf{k}, s} \hat{c}_{V, \mathbf{k}s}^\dagger |0\rangle, \quad (2.11)$$

the valence band is completely filled, whereas the conduction band is empty. The product on the right hand side runs over all wave vectors in the first Brillouin zone. The ground state energy is given by

$$E_0 = 2 \sum_{\mathbf{k}} \epsilon_{V, \mathbf{k}}. \quad (2.12)$$

The total momentum and spin of the ground state vanish.

³Spin-orbit coupling has been neglected so far. Including the spin degrees of freedom would lead to a splitting of the energies at $k = 0$ into a two-fold degenerate level (Γ_6^+) and a four-fold degenerate one (Γ_8^+).

2.2.1 Electron-hole excitations

A simple excitation of the system consists of removing an electron (i.e., creating a hole) from the valence band and putting it into the conduction band, i.e.,

$$|C, \mathbf{k} + \mathbf{q}, s; V, \mathbf{k}, s'\rangle = c_{C, \mathbf{k} + \mathbf{q}, s}^\dagger c_{V, \mathbf{k}, s'} |\Phi_0\rangle, \quad (2.13)$$

where the possibility of changing the spin s of the electron to s' and of shifting the wave vector of conduction electrons by \mathbf{q} is included ($|C, \mathbf{k} + \mathbf{q}, s; V, \mathbf{k}, s'\rangle$ is assumed to be normalized). The electron-hole pair may either be in a spin-singlet (pure charge excitation) or a spin-triplet state (spin excitation). Apart from spin, the state is characterized by the wave vectors \mathbf{k} and \mathbf{q} . The excitation energy is given by

$$E = \epsilon_{C, \mathbf{k} + \mathbf{q}} - \epsilon_{V, \mathbf{k}}. \quad (2.14)$$

The spectrum of such an electron-hole excitation with given \mathbf{q} is determined by the spectral function

$$I(\mathbf{q}, E) = \sum_{\mathbf{k}} |\langle C, \mathbf{k} + \mathbf{q}, s; V, \mathbf{k}, s' | c_{C, \mathbf{k} + \mathbf{q}, s}^\dagger c_{V, \mathbf{k}, s'} |\Phi_0\rangle|^2 \delta(E - (\epsilon_C(\mathbf{k} + \mathbf{q}) - \epsilon_V(\mathbf{k}))). \quad (2.15)$$

Excitations exist for all pairs ω and \mathbf{q} for which $I(\mathbf{q}, \omega)$ does not vanish and, consequently, only above a \mathbf{q} -dependent threshold, which is minimal for $\mathbf{q} = \mathbf{k}_0$, where $\mathbf{k}_0 = 0$ ($\mathbf{k}_0 \neq 0$) for a direct (indirect) energy gap. As \mathbf{k} is not fixed, there is a continuum of excited states above the threshold for each \mathbf{q} (see Fig. 2.7).

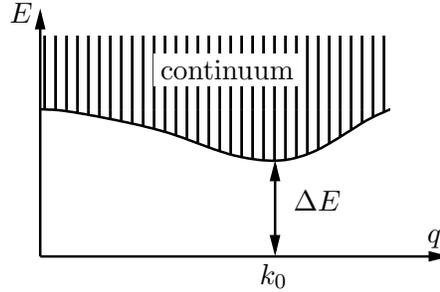


Figure 2.7: Electron-hole excitation spectrum. Excitations exist in the shaded region, where $I(\mathbf{q}, E) \neq 0$.

For the electron-hole excitations considered here, interactions are irrelevant, and the electrons involved are treated as non-interacting particles. Note the analogy with the Dirac-sea in relativistic quantum mechanics: The electron-hole excitations of a semiconductor correspond to electron-positron pair creation in the Dirac theory.

2.2.2 Excitons

Taking into account the Coulomb interaction between the electrons, there is another class of excitations called excitons. In order to discuss them, we extend the Hamiltonian (2.10) by the Coulomb interaction,

$$\hat{V} = \sum_{s, s'} \int d^3 r d^3 r' \hat{\Psi}_s^\dagger(\mathbf{r}) \hat{\Psi}_{s'}^\dagger(\mathbf{r}') \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} \hat{\Psi}_{s'}(\mathbf{r}') \hat{\Psi}_s(\mathbf{r}), \quad (2.16)$$

where the field operators are defined by

$$\hat{\Psi}_s(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{n=V, C} \sum_{\mathbf{k}} u_{n, \mathbf{k}}(\mathbf{r}) e^{i\mathbf{k} \cdot \mathbf{r}} \hat{c}_{n, \mathbf{k}, s}, \quad (2.17)$$

where $u_{n,\mathbf{k}}(\mathbf{r})$ are the Bloch functions of the band $n = C, V$. Now, we consider a general particle-hole state,

$$|\Phi_{\mathbf{q}}\rangle = \sum_{\mathbf{k}} A(\mathbf{k}) \hat{c}_{C,\mathbf{k}+\mathbf{q},s}^\dagger \hat{c}_{V,\mathbf{k},s'} |\Phi_0\rangle = \sum_{\mathbf{k}} A(\mathbf{k}) |C, \mathbf{k} + \mathbf{q}, s; V, \mathbf{k}, s'\rangle, \quad (2.18)$$

and demand that it satisfies the stationary Schrödinger equation $(\mathcal{H} + \hat{V})|\Phi_{\mathbf{q}}\rangle = E|\Phi_{\mathbf{q}}\rangle$. This two-body problem can be expressed as

$$\sum_{\mathbf{k}'} \langle C, \mathbf{k} + \mathbf{q}, s; V, \mathbf{k}, s' | \mathcal{H} + \hat{V} | C, \mathbf{k}' + \mathbf{q}, s; V, \mathbf{k}', s' \rangle A(\mathbf{k}') = EA(\mathbf{k}). \quad (2.19)$$

The matrix elements are given by

$$\langle C, \mathbf{k} + \mathbf{q}, s; V, \mathbf{k}, s' | \mathcal{H} | C, \mathbf{k}' + \mathbf{q}, s; V, \mathbf{k}', s' \rangle = \delta_{\mathbf{k},\mathbf{k}'} \{ \epsilon_{C,\mathbf{k}+\mathbf{q}} - \epsilon_{V,\mathbf{k}} \} \quad (2.20)$$

and

$$\begin{aligned} \langle C, \mathbf{k} + \mathbf{q}, s; V, \mathbf{k}, s' | \hat{V} | C, \mathbf{k}' + \mathbf{q}, s; V, \mathbf{k}', s' \rangle = \\ \frac{2\delta_{S,0}}{\Omega^2} \int d^3r d^3r' u_{C,\mathbf{k}+\mathbf{q}}^*(\mathbf{r}) u_{V,\mathbf{k}}(\mathbf{r}) u_{C,\mathbf{k}'+\mathbf{q}}(\mathbf{r}') u_{V,\mathbf{k}'}^*(\mathbf{r}') e^{-i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')} \frac{e^2}{|\mathbf{r}-\mathbf{r}'|} \\ - \frac{1}{\Omega^2} \int d^3r d^3r' u_{C,\mathbf{k}+\mathbf{q}}^*(\mathbf{r}) u_{V,\mathbf{k}}(\mathbf{r}') u_{C,\mathbf{k}'+\mathbf{q}}(\mathbf{r}) u_{V,\mathbf{k}'}^*(\mathbf{r}') e^{i(\mathbf{k}'-\mathbf{k})\cdot(\mathbf{r}-\mathbf{r}')} \frac{e^2}{|\mathbf{r}-\mathbf{r}'|}, \end{aligned} \quad (2.21)$$

where $\delta_{S,0} = 1$ if the excitation is a spin-singlet and $\delta_{S,0} = 0$ otherwise. The first term is the exchange term, and the second term the direct term of the Coulomb interaction. Now we consider a semiconductor with a direct energy gap at the Γ -point. Thus, the most important wave vectors are those around $\mathbf{k} = 0$. We approximate

$$u_{n,\mathbf{k}'}^*(\mathbf{r}) u_{n,\mathbf{k}}(\mathbf{r}) \approx \frac{1}{\Omega} \int d^3r u_{n,\mathbf{k}'}^*(\mathbf{r}) u_{n,\mathbf{k}}(\mathbf{r}) = \frac{1}{\Omega} \langle u_{n,\mathbf{k}'} | u_{n,\mathbf{k}} \rangle \approx 1, \quad (2.22)$$

which is reasonable for $\mathbf{k} \approx \mathbf{k}'$. In the same manner, we see that

$$u_{C,\mathbf{k}+\mathbf{q}}^*(\mathbf{r}) u_{V,\mathbf{k}}(\mathbf{r}) \approx \frac{1}{\Omega} \langle u_{C,\mathbf{k}+\mathbf{q}} | u_{V,\mathbf{k}} \rangle \approx \frac{1}{\Omega} \langle u_{C,\mathbf{k}} | u_{V,\mathbf{k}} \rangle = 0. \quad (2.23)$$

Note that the semiconductor is a dielectric medium with a dielectric constant ϵ ($\mathbf{D} = \epsilon \mathbf{E}$). Classical electrodynamics states that

$$\nabla \cdot \mathbf{E} = \frac{4\pi\rho}{\epsilon}, \quad (2.24)$$

i.e., the Coulomb potential is partially screened due to dielectric polarization. Including this effect in the Schrödinger equation phenomenologically, the matrix element (2.21) takes on the form

$$-\frac{4\pi e^2}{\Omega \epsilon |\mathbf{k} - \mathbf{k}'|^2}. \quad (2.25)$$

Thus, we can write the stationary equation (2.19) as

$$(\epsilon_{C,\mathbf{k}+\mathbf{q}} - \epsilon_{V,\mathbf{k}} - E) A(\mathbf{k}) - \frac{1}{\Omega} \sum_{\mathbf{k}'} \frac{4\pi e^2}{\epsilon |\mathbf{k} - \mathbf{k}'|^2} A(\mathbf{k}') = 0. \quad (2.26)$$

We include the band structure using the $\mathbf{k} \cdot \mathbf{p}$ - approximation which, for a direct energy gap, leads to

$$\epsilon_{C,\mathbf{k}} = \frac{\hbar^2 \mathbf{k}^2}{2m_C} + E_0 + E_g \quad \text{and} \quad \epsilon_{V,\mathbf{k}} = E_0 - \frac{\hbar^2 \mathbf{k}^2}{2m_V}, \quad (2.27)$$

where E_0 denotes the energy of the valence band top. We define a so-called envelope function $F(\mathbf{r})$ by

$$F(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{k}} A(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}}. \quad (2.28)$$

This function satisfies the differential equation

$$\left[-\frac{\hbar^2 \nabla^2}{2\mu_{\text{ex}}} + \frac{\hbar^2}{2i} \left(\frac{1}{m_V} - \frac{1}{m_C} \right) \mathbf{q} \cdot \nabla - \frac{e^2}{\varepsilon|\mathbf{r}|} \right] F(\mathbf{r}) = \left\{ E - E_g - \frac{\hbar^2 \mathbf{q}^2}{2\mu_{\text{ex}}} \right\} F(\mathbf{r}), \quad (2.29)$$

where μ_{ex} is the reduced mass, i.e., $\mu_{\text{ex}}^{-1} = m_V^{-1} + m_C^{-1}$. The term linear in ∇ can be eliminated by the transformation

$$F(\mathbf{r}) = F'(\mathbf{r}) \exp \left(\frac{i}{2} \frac{m_V - m_C}{m_V + m_C} \mathbf{q} \cdot \mathbf{r} \right), \quad (2.30)$$

and after some algebraic manipulations we obtain

$$\left[-\frac{\hbar^2 \nabla^2}{2\mu_{\text{ex}}} - \frac{e^2}{\varepsilon|\mathbf{r}|} \right] F'(\mathbf{r}) = \left\{ E - E_g - \frac{\hbar^2 \mathbf{q}^2}{2M_{\text{ex}}} \right\} F'(\mathbf{r}), \quad (2.31)$$

where $M_{\text{ex}} = m_V + m_C$.

The stationary equation (2.31) is equivalent to the Schrödinger equation of a hydrogen atom. The energy levels then are given by

$$E_{\mathbf{q}} = E_g - \frac{\mu_{\text{ex}} e^4}{2\varepsilon^2 \hbar^2 n^2} + \frac{\hbar^2 \mathbf{q}^2}{2M_{\text{ex}}}, \quad (2.32)$$

which implies that there are excitations below the particle-hole continuum, corresponding to particle-hole bound states. This excitation spectrum is discrete and there is a well-defined relation between energy and momentum (\mathbf{q}), which is the wave vector corresponding to the center of mass of the particle-hole pair. This non-trivial quasi-particle is called *exciton*. In the present approximation it takes on the form of a simple two-particle state. In fact, however, it may be viewed as a collective excitation, as the dielectric constant includes the polarization by all electrons. When the screening is neglected, the excitonic states would not make sense as their energies would not be within the band gap but much below. For the case of weak binding considered above, the excitation is called a *Wannier exciton*. The typical binding energy is

$$E_b \sim \frac{\mu_{\text{ex}}}{m\varepsilon^2} Ry. \quad (2.33)$$

Typical values of the constants on the right hand side are $\varepsilon \sim 10$ and $\mu_{\text{ex}} \sim m/10$, so that the binding energy is in the meV range. This energy is much smaller than the energy gap, such that the excitons are inside the gap, as shown schematically in Fig. 2.8.

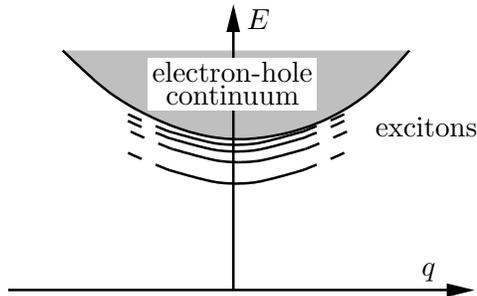


Figure 2.8: Qualitative form of the exciton spectrum below the electron-hole continuum.

The exciton levels are dispersive and their spectrum becomes increasingly dense with increasing energy, similar to the hydrogen atom. When they merge with the particle-hole continuum the bound state is ‘ionized’, i.e., the electron and the hole dissociate and behave like independent particles.

Strongly bound excitons are called *Frenkel excitons*. In the limit of strong binding, the pair is almost local, so that the excitation is restricted to a single atom rather than involving the whole semiconductor band structure.

Excitons are mobile, but they carry no charge, as they consist of an electron and a hole with opposite charges. For small densities they approximately obey Bose-Einstein statistics, as they are made from two fermions. In special cases, Bose-Einstein condensation of excitons can be observed experimentally.

2.2.3 Optical properties

Excitation in semiconductors can occur via the absorption of electromagnetic radiation. The energy and momentum transferred by a photon is $\hbar\omega$ and $\hbar\mathbf{q}$, respectively. With the relation $\omega = c|\mathbf{q}|$ we may estimate this momentum transfer in a semiconductor with $E_g \sim 1\text{eV} \sim e^2/a$ (speed of light c , lattice constant $a \approx 10a_B$) to be

$$k = \frac{\omega}{c} = \frac{\hbar\omega}{\hbar c} 2\pi \sim \frac{e^2}{\hbar c} \frac{2\pi}{a} = \alpha \frac{2\pi}{a} \ll \frac{2\pi}{a}, \quad (2.34)$$

which shows that momentum transfer can be ignored.

For semiconductors with a direct energy gap (e.g., GaAs) the photo-induced electron-hole excitation is most easy and yields absorption rates with the characteristics

$$\Gamma_{\text{abs}}(\omega) \propto \begin{cases} (\hbar\omega - E_g)^{1/2}, & \text{dipole-allowed,} \\ (\hbar\omega - E_g)^{3/2}, & \text{dipole-forbidden.} \end{cases} \quad (2.35)$$

Here, the terms “dipole-allowed” and “dipole-forbidden” have a similar meaning as in the excitation of atoms regarding whether matrix elements of the type $\langle u_{V,\mathbf{k}} | \mathbf{r} | u_{C,\mathbf{k}} \rangle$ are finite or vanish, respectively. Obviously, dipole-allowed transitions occur at a higher rate for photon energies immediately above the energy gap E_g , than for dipole-forbidden transitions.

For semiconductors with indirect energy gap (e.g., Si and Ge), the lowest energy transition connecting the top of the valence band to the bottom of the conduction band is not allowed without the help of phonons which contribute little energy but much momentum transfer, as $\hbar\omega_{\mathbf{Q}} \ll \hbar\omega$ with $\omega_{\mathbf{Q}} = c_s|\mathbf{Q}|$ and the sound velocity $c_s \ll c$. The requirement of a phonon assisting in the transition reduces the transition rate to

$$\Gamma_{\text{abs}}(\omega) \propto c_+(\hbar\omega + \hbar\omega_{\mathbf{Q}} - E_g)^2 + c_-(\hbar\omega - \hbar\omega_{\mathbf{Q}} - E_g)^2, \quad (2.36)$$

where \mathbf{Q} corresponds to the wave vector of the phonon connecting the top of the valence band and the bottom of the conduction band. There are two relevant processes: either the phonon is absorbed (c_+ -process) or it is emitted (c_- -process) (see Fig. 2.9).

In addition, absorption processes including exciton states exist. This leads to discrete absorption peaks below the absorption continuum. In Fig. 2.10, we show the situation for a direct-gap semiconductor.

Naturally, the recombination of electrons and holes is important as well; in particular, if it is a radiative recombination, i.e., leads to the emission of a photon. Additionally, other recombination channels such as recombination at impurities, interfaces and through Auger processes are possible. The radiative recombination for the direct-gap semiconductors is most relevant for applications. There, the photon emission rate follows the approximate law

$$\Gamma_{\text{em}}(\omega) \propto [N_\gamma(\omega) + 1](\hbar\omega - E_g)^{1/2} e^{-\hbar\omega/k_B T}, \quad (2.37)$$

with the photon density $N_\gamma(\omega)$. This yields the dominant rate for $\hbar\omega$ very close to E_g .

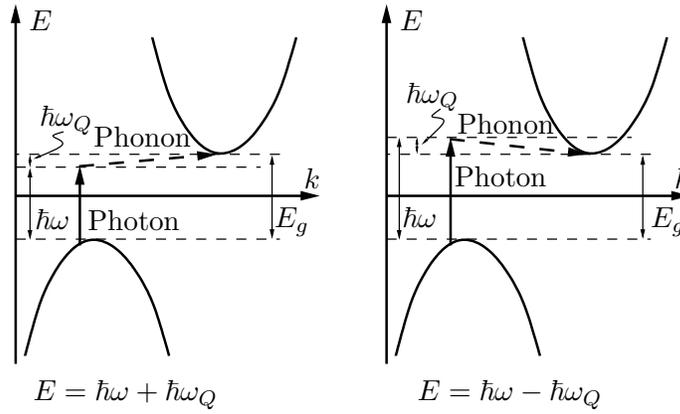


Figure 2.9: Phonon-assisted photon absorption in a semiconductor with indirect gap: phonon absorption (left panel) and phonon emission (right panel).

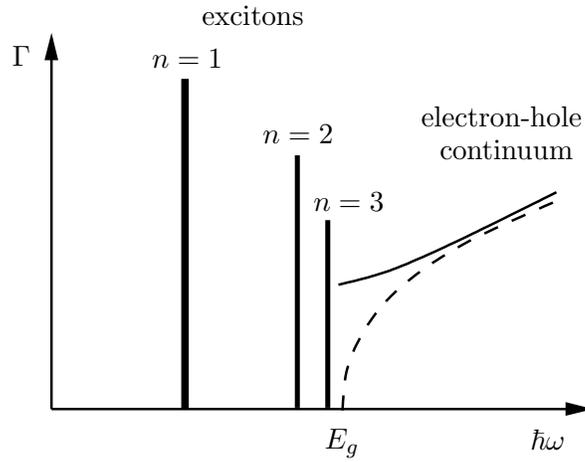


Figure 2.10: Absorption spectrum including the exciton states for a direct-gap semiconductor with dipole-allowed transitions. The exciton states appear as sharp lines below the electron-hole continuum starting at $\hbar\omega = E_g$.

2.3 Doping semiconductors

Let us replace a Si atom in a Si semiconductor by aluminium Al (group III) or phosphorus P (group V), which then act as impurities in the crystal lattice. Both Al and P are in the same row of the periodic table, and their electron configurations are given by

$$\text{Al: } \{(1s)^2(2s)^2(2p)^6\}\underline{(3s)^2(3p)}, \quad \text{P: } \{(1s)^2(2s)^2(2p)^6\}\underline{(3s)^2(3p)^3}.$$

Al (P) has one electron less (more) than Si.

2.3.1 Impurity state

We consider the case of a P-impurity contributing an additional electron whose dynamics is governed by the conduction band of the semiconductor. For the sake of simplicity, we describe the conduction band by a single isotropic band with effective mass m^* ,

$$\epsilon_{\mathbf{k}} = \frac{\hbar^2 \mathbf{k}^2}{2m^*}. \quad (2.38)$$

The P-ion represents a positively charged center, which attracts the additional electron. In the simplest model, this situation is described by the so-called Wannier equation

$$\left\{ -\frac{\hbar^2 \nabla^2}{2m^*} - \frac{e^2}{\varepsilon|\vec{r}|} \right\} F(\mathbf{r}) = EF(\mathbf{r}), \quad (2.39)$$

where ε is the dielectric constant in which the screening of the ionic potential is contained. Analogous to the discussion of the exciton states, $F(\mathbf{r})$ is an envelope wave function of the electron; again, the problem resembles that of a hydrogen atom. Therefore, the low energy states of the additional electron are bound states around the P-ion. The electron may become mobile when this “hydrogen atom” is ionized. The binding energy relative to the minimum of the conduction band given by

$$E_n = -\frac{m^* e^4}{2\hbar^2 \varepsilon^2 n^2} = -\frac{m^*}{m \varepsilon^2 n^2} \text{Ry}, \quad (2.40)$$

and the effective radius of the lowest bound state by

$$r_1 = \frac{\hbar^2 \varepsilon}{m^* e^2} = \frac{\varepsilon m}{m^*} a_B, \quad (2.41)$$

where $a_B = 0.53\text{\AA}$ is the Bohr radius. For Si, $m^* \approx 0.2m$ and $\varepsilon \approx 12$, such that

$$E_1 \approx -20\text{meV} \quad \text{and} \quad r_1 \approx 30\text{\AA}. \quad (2.42)$$

Thus, the resulting states are weakly bound, with energies inside the band gap.

We conclude that the net effect of the P-impurities is to introduce additional electrons into the crystal that can be easily transferred to the conduction band by thermal excitation (ionization). One speaks of an *n-doped* semiconductor (n: negative charge). In full analogy one can consider Al-impurities, thereby replacing electrons with holes: An Al-atom introduces an additional hole into the lattice which is weakly bound to the Al-ion (its energy is slightly above the band edge of the valence band) and may dissociate from the impurity by thermal excitation. This case is called *p-doping* (p: positive charge). In both cases, the chemical potential is tied to the dopand levels, i.e., it lies between the dopand level and the valence band for p-doping and between the dopand level and the conduction band in case of n-doping (Fig. 2.11).

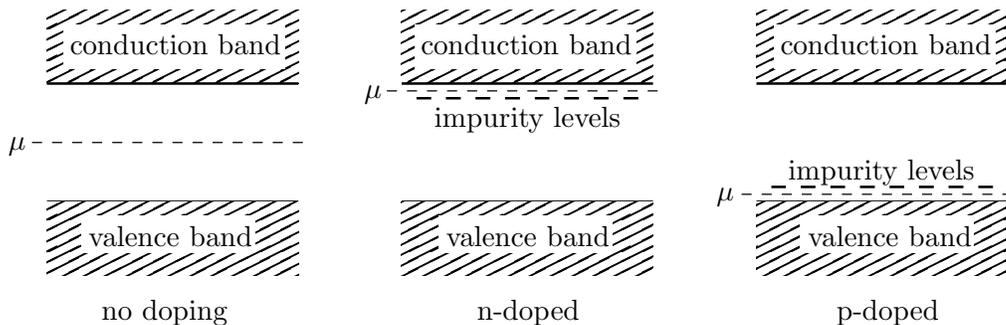


Figure 2.11: Position of the chemical potential in semiconductors.

The electric conductivity of semiconductors (in particular at room temperature) can be tuned strongly by doping with so-called ‘donators’ (n-doping) and ‘acceptors’ (p-doping). Practically all dopand atoms are ionized, with the electrons/holes becoming mobile. Combining differently doped semiconductors, the possibility to engineer electronic properties is enhanced even more. This is the basic reason for the semiconductors being ubiquitous in modern electronics.

2.3.2 Carrier concentration

Let us briefly consider the carrier concentration in semiconductors at room temperature. Carriers are always created in form of electron hole pairs; following the “reaction formula”

$$e + h \leftrightarrow \gamma, \quad (2.43)$$

where γ denotes a photon which is absorbed (e - h -creation) or emitted (e - h -recombination) and accounts for the energy balance. The carrier concentration is described by a mass action law of the form,

$$n_e n_h = n(T)^2 = n_0 \left(\frac{T}{T_0} \right)^3 e^{-E_b/k_B T}, \quad (2.44)$$

where T_0 , n_0 and E_g are parameters specific to the semiconductor. In the case of Si $n_e n_h \approx 10^{20} \text{cm}^{-3}$ at $T = 300\text{K}$. Thus, for the undoped semiconductor, $n_e = n_h \approx 10^{10} \text{cm}^{-3}$. On the other hand, for n-doped Si with a typical donor concentration of $n_D \approx 10^{17} \text{cm}^{-3}$ we can safely assume that most of the donors are ionized at room temperature such that $n_e \approx n_D$ and

$$n_h = \frac{n(T)^2}{n_e} \approx 10^3 \text{cm}^{-3}. \quad (2.45)$$

Therefore, the vast majority of mobile carriers are electrons, while the hole carriers are negligible. The opposite is true for p-doped Si.

2.4 Semiconductor devices

Semiconductors are among the most important components of current high-technology. In this section, we consider a few basic examples of semiconductor devices.

2.4.1 pn-contacts as diodes

pn-junctions—made by combining a p-doped and an n-doped version of the same semiconductor—are used as rectifiers⁴. When contacting the two types of doped semiconductors the chemical potential, which is pinned by the dopand (impurity) levels, is important for the behavior of the electrons at the interface. In electrostatic equilibrium, the chemical potential is constant across the interface. This is accompanied by a “band bending” leading to the ionization of the impurity levels in the interface region (see Fig. 2.12). Consequently, these ions produce an electric dipole layer which induces an electrostatic potential shift across the interface. Additionally, the carrier concentration is strongly reduced in the interface region (depletion layer).

In the absence of a voltage over the junction, the net current flow vanishes and the dipole is in electrostatic equilibrium. This can also be interpreted as the equilibrium of two oppositely directed currents, called the drift current and the diffusion current. From the point of view of the electrons, the dipole field exerts a force pulling the electrons from the p-side to the n-side. This leads to the drift current J_{drift} . On the other hand, the electron concentration gradient leads to the diffusion current J_{diff} from the n-side to the p-side. The diffusion current is directed against the potential gradient, so that the diffusing electrons have to overcome a potential step. The equilibrium condition is given by

$$0 = J_{\text{tot}} = J_{\text{diff}} + J_{\text{drift}} \propto C_1(T)e^{-E_g/k_B T} - C_2(T)e^{-E_g/k_B T} = 0, \quad (2.46)$$

where $C_1 = C_2 = C$. Both currents are essentially determined by the factor $C(T)e^{-E_g/k_B T}$. For the drift current, this dependence stems from the dependence of the current on the concentration of mobile charge carriers (electrons and holes on the p-side and n-side, respectively), which are

⁴dt. Gleichrichter

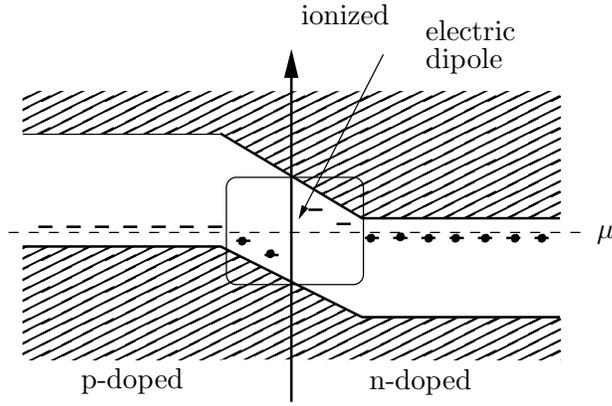


Figure 2.12: Occupation of the impurity levels of a pn-junction.

created by thermal excitation. The magnitude of the drift current depends on the concentration because, in regions where the dipole field is large, all electrons on the p-side are pulled to the n-side (and vice versa for holes). Applying a voltage does not change this contribution significantly. For the diffusion current, the factor $C(T)e^{-E_g/k_B T}$ describes the thermal activation over the dipole barrier. This contribution strongly depends on the applied voltage, as the height of the barrier directly depends on the voltage. For zero voltage, the height of the barrier is essentially given by the energy gap $E_b \approx E_g$. With an applied voltage, this is modified according to $E_b \approx E_g + eU$, where $eU = \mu_n - \mu_p$. From these considerations, the well-known current-voltage characteristic of the pn-junctions follows directly as

$$J_{\text{tot}}(U) = C(T)e^{-E_g/k_B T} \left\{ e^{eU/k_B T} - 1 \right\}. \quad (2.47)$$

For $U > 0$, the current is rapidly enhanced with increasing voltage. This is called forward bias. By contrast, charge transport is suppressed for $U < 0$ (reverse bias), leading to small currents only. The current-voltage characteristics $J(U)$ (see Fig. 2.13) shows a clearly asymmetric behavior, which can be used to rectify ac-currents. Rectifiers (or diodes) are an important component of many integrated circuits.

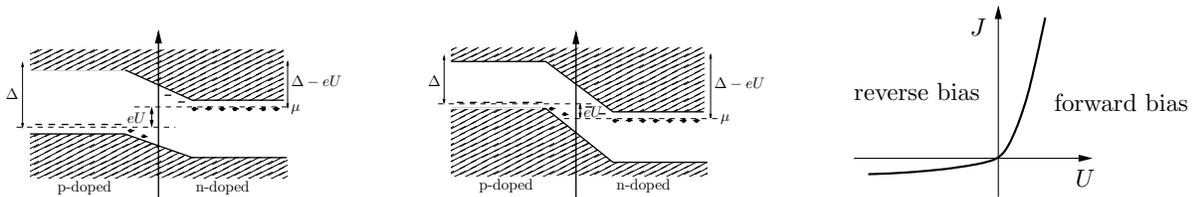


Figure 2.13: The pn-junction with an applied voltage and the resulting J - U characteristics.

2.4.2 Semiconductor diodes and light

LED (Light Emitting Diode): As mentioned above, the recombination of electrons and holes can lead to the emission of photons (radiating recombination) with a rather well-defined frequency essentially corresponding to the energy gap E_g . Excess electron-hole pairs can be produced in pn diodes by running a current in forward direction. Using semiconductors with different energy gaps allows the tuning of the color of the emitted light. Direct-gap semiconductors are most suitable for this kind of devices. Well-known are the semiconductors of the GaAs-GaN series (see table 2.1).

semiconductor	GaAs	GaAs _{0.6} P _{0.4}	GaAs _{0.4} P _{0.6}	GaP	GaN
wave length (nm)	940	660	620	550	340
color	infrarot	rot	gelb	grün	ultraviolet

Table 2.1: Materials used for LEDs and their properties.

There are, however, certain problems concerning the emission of light by semiconductors. In particular, the difference in refractive indices inside and outside the device leads to large reflective losses ($n_{\text{SC}} \approx 3$ und $n_{\text{air}} \approx 1$). Thus, the efficiency, defined as $N(\text{photons emitted})/N(\text{p-h pairs created})$, is small, but still better than the efficiency of conventional light bulbs.

Solar cell: The population of charge carriers can be changed by the absorption of light. Suppose that the n-side of a diode is exposed to irradiation by light, which leads to excess hole carriers (minority charge carriers). Some of these holes can diffuse towards the pn-interface and will be drawn to the p-side by the dipole field. In this way, they induce an additional current J_L modifying the current-voltage characteristics to

$$J_{\text{tot}} = J_{pn} - J_L = J_s(e^{eU/k_B T} - 1) - J_L. \quad (2.48)$$

It is important for the successful migration of the holes to the interface dipole that they do not recombine too quickly. When $J_{\text{tot}} = 0$, the voltage drop across the diode is U_L . The maximum efficiency is reached by applying an external voltage $U_c < U_L$ such that the product $J_c \times U_c$ is maximized (cf. Fig. 2.14).

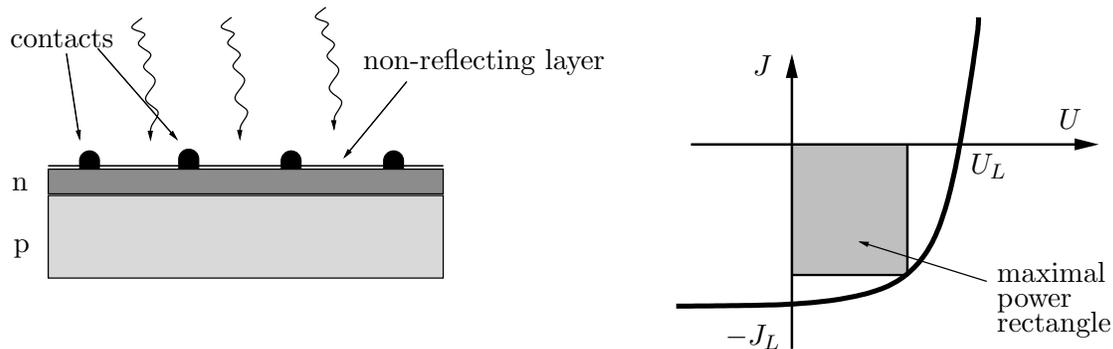


Figure 2.14: Solar cell design and shifted current-voltage characteristics. The efficiency is maximal for a maximal area of the power rectangle.

2.4.3 MOSFET

The arguably most important application of semiconductors is the transistor, an element existing with different architectures. Here we consider the MOSFET (Metal-Oxide-Semiconductor-Field-Effect-Transistor). A transistor is a switch allowing to control the current through the device by switching a small control voltage. In the MOSFET, this is achieved by changing the charge carrier concentration in a p-doped semiconductor using a metallic gate. The basic design of a MOSFET is as follows (see Fig. 2.15): A thin layer of SiO_2 is deposited on the surface of a p-type semiconductor. SiO_2 is a good insulator that is compatible with the lattice structure of Si. Next, a metallic layer, used as a gate electrode, is deposited on top of the insulating layer. The voltage between the Si semiconductor and the metal electrode is called gate voltage U_G . The insulating SiO_2 layer ensures that no currents flow between the electrode and the semiconductor when a gate voltage is applied. The switchable currents in the MOSFET flow between the source

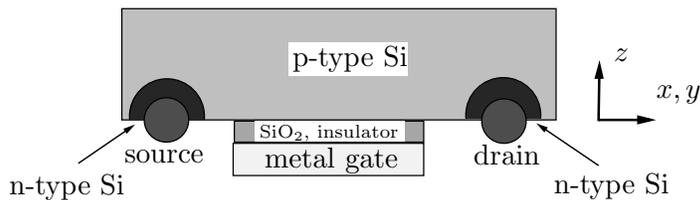


Figure 2.15: Schematic design of a MOSFET device.

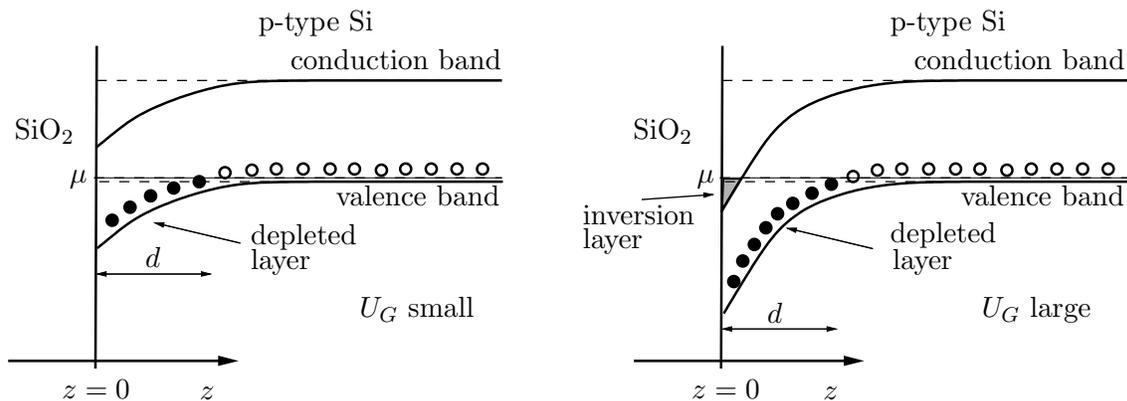


Figure 2.16: Depletion layer at SiO₂-semiconductor interface: $E_g > U_G > 0$ (left panel); $U_G > E_g$ inversion layer (right panel).

and the drain which are heavily n-doped semiconductor regions. We consider the two typical switching states:

“off” The gate voltage $U_G = 0$. Virtually no current flows, as the conduction band of the p-doped semiconductor is almost empty.

“on” The gate voltage is positive, $U_G > 0$. In this case, the energy of the Si bands is lowered, such that in a certain region within the Si the acceptor levels are filled with electrons (or, equivalently, holes are removed). This depletion layer has the extension d . The negative charge of the acceptors leads to a position-dependent potential $\Phi(z)$, where z is the distance from the boundary between SiO₂ and Si. $\Phi(z)$ satisfies the Poisson equation

$$\frac{d^2}{dz^2}\Phi(z) = \frac{4\pi\rho(z)}{\varepsilon}, \quad (2.49)$$

where the charge density originates in the occupied acceptor levels,

$$\rho(z) = \begin{cases} -en_A, & z < d, \\ 0, & z > d, \end{cases} \quad (2.50)$$

and n_A is the density of acceptors. The boundary conditions are

$$\Phi(z=0) = U_G \quad \text{and} \quad \Phi(z=d) = 0. \quad (2.51)$$

Thus, the solution for $z \leq d$ is given by

$$\Phi(z) = \frac{2\pi en_A}{\varepsilon}(z-d)^2, \quad \text{with} \quad d^2 = \frac{\varepsilon U_G}{2\pi en_A}. \quad (2.52)$$

The thickness of the depletion layer increases with increasing gate voltage. When the applied gate voltage is sufficiently large ($U_G > E_g$), a so-called inversion layer is created

(cf. Fig. 2.16). The conduction band is bent down so close to the boundary that its lower edge lies below the chemical potential. The electrons accumulating in this inversion layer provide carriers connecting the n-type source and drain electrodes and allow for a large current. This is the “on” configuration.

Conduction band electrons accumulating in the inversion layer behave like a two-dimensional electron gas. In this system, the quantum Hall effect (QHE) can occur. It is characterized by highly unusual charge transport properties in the presence of a large magnetic field.

Chapter 3

Metals - properties of interacting electrons

The electronic states in a periodic atomic lattice are extended and have an energy spectrum forming energy bands. In the ground state these energy states are filled successively starting at the bottom of the electronic spectrum until the number of electrons is exhausted. A metallic state occurs whenever in this way a band is only partially filled. The fundamental difference that distinguishes metals from insulators and semiconductors is the absence of a gap for electron-hole excitations, so that these can be excited at arbitrarily small energies. This difference has profound phenomenological consequences.

We will consider a simplified model suitable for the description of simple metals like the Alkali metals Li, Na, or K, with an (atomic) electron configuration consisting of closed shell cores and a single valence electron in an ns -orbital. Restricting ourselves to the valence electron and applying the approximation of nearly free electrons, we find that the lowest band around the Γ -point is half-filled. We will then neglect the influence of the periodic lattice potential and consider the problem of a free electron gas subject to the (repulsive) Coulomb interaction.

3.1 The Jellium model of the metallic state

The Jellium model is the probably simplest possible model of a metal that can be used to discuss qualitative and in parts even quantitative aspects of simple metals. The main simplification made is to replace the ionic lattice by a homogeneous positively charged background (called Jellium). The uniform charge density en_{ion} is chosen such that the whole system (ions + electrons) is charge neutral, i.e. $n_{\text{ion}} = n$, where n is the electron density.

We begin with the single-particle wave functions of the free electrons which are given by the usual plane waves

$$\psi_{\mathbf{k},s}(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} e^{i\mathbf{k}\cdot\mathbf{r}} \quad (3.1)$$

where Ω is the volume of the system (needed for normalization of the wave function), \mathbf{k} and $s = \uparrow, \downarrow$ denote the wave vector and spin, respectively. We consider periodic boundary conditions for the wave function by taking the space to be a cube of side length L and demanding that

$$\begin{aligned} \psi_{\mathbf{k},s}(\mathbf{r} + (L, 0, 0)) &= \psi_{\mathbf{k},s}(\mathbf{r} + (0, L, 0)) = \psi_{\mathbf{k},s}(\mathbf{r} + (0, 0, L)) = \psi_{\mathbf{k},s}(\mathbf{r}) \\ \Rightarrow \mathbf{k} &= \frac{2\pi}{L}(n_x, n_y, n_z) \quad \text{mit } n_x, n_y, n_z = 0, \pm 1, \pm 2, \dots \end{aligned} \quad (3.2)$$

The energy of a single particle state is given by $\epsilon_{\mathbf{k}} = \hbar^2 \mathbf{k}^2 / 2m$. The non-interacting ground state is obtained by filling all single particle states up to the Fermi energy with two electrons.

In the language of second quantization the ground state is thus given by

$$|\Psi_0\rangle = \prod_{|\mathbf{k}| \leq k_F, s} \hat{c}_{\mathbf{k},s}^\dagger |0\rangle \quad (3.3)$$

where k_F is the Fermi wave vector (The corresponding Fermi energy is $\epsilon_F = \hbar^2 k_F^2 / 2m$). Counting the filled electronic states, we have

$$n = \frac{1}{\Omega} \sum_{|\mathbf{k}| \leq k_F, s} 1 = 2 \int \frac{d^3 k}{(2\pi)^3} 1 = 2 \frac{4\pi}{3} \frac{k_F^3}{(2\pi)^3} \quad \Rightarrow \quad k_F = \{3\pi^2 n\}^{1/3} \quad (3.4)$$

where k_F is the radius of the Fermi sphere in k -space around $\mathbf{k} = 0$. The operators $\hat{c}_{\mathbf{k},s}^\dagger$ ($\hat{c}_{\mathbf{k},s}$) create (annihilate) an electron with wave vector \mathbf{k} and spin s .

Now we compute the ground state energy of the Jellium system variationally, using the density n as the only variational parameter. Varying the density is equivalent to varying the lattice constant, and accordingly in this way we obtain an understanding of the stability of a metal, i.e. the cohesion of the ion lattice through the mobile electrons (in contrast to the stability of semiconductors due to covalent chemical bonding). We denote the variational state as $|\Psi_0\rangle$ with a given k_F . The second quantized Hamiltonian has the form

$$\begin{aligned} \mathcal{H} &= \mathcal{H}_{\text{kin}} + \mathcal{H}_{ee} + \mathcal{H}_{ei} + \mathcal{H}_{ii} \\ \mathcal{H}_{\text{kin}} &= \sum_{\mathbf{k},s} \epsilon_{\mathbf{k}} \hat{c}_{\mathbf{k},s}^\dagger \hat{c}_{\mathbf{k},s} \\ \mathcal{H}_{ee} &= \frac{1}{2} \sum_{s,s'} \int d^3 r d^3 r' \hat{\Psi}_s^\dagger(\mathbf{r}) \hat{\Psi}_{s'}^\dagger(\mathbf{r}') \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} \hat{\Psi}_{s'}(\mathbf{r}') \hat{\Psi}_s(\mathbf{r}) \\ \mathcal{H}_{ei} &= - \sum_s \int d^3 r d^3 r' \frac{ne^2}{|\mathbf{r} - \mathbf{r}'|} \hat{\Psi}_s^\dagger(\mathbf{r}) \hat{\Psi}_s(\mathbf{r}') \\ \mathcal{H}_{ii} &= \frac{1}{2} \int d^3 r d^3 r' \frac{n^2 e^2}{|\mathbf{r} - \mathbf{r}'|}, \end{aligned} \quad (3.5)$$

where we have used the electron field operator with the definitions

$$\hat{\Psi}_s^\dagger(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{k}} \hat{c}_{\mathbf{k},s}^\dagger e^{-i\mathbf{k}\cdot\mathbf{r}} \quad \text{und} \quad \hat{\Psi}_s(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{k}} \hat{c}_{\mathbf{k},s} e^{i\mathbf{k}\cdot\mathbf{r}} \quad (3.6)$$

The variational energy can be computed from $E_g = \langle \Psi_0 | \mathcal{H} | \Psi_0 \rangle$ and consists of four different contributions: First we have the kinetic energy

$$E_{\text{kin}} = \langle \Psi_0 | \mathcal{H}_{\text{kin}} | \Psi_0 \rangle = \sum_{\mathbf{k},s} \epsilon_{\mathbf{k}} \underbrace{\langle \Psi_0 | \hat{c}_{\mathbf{k},s}^\dagger \hat{c}_{\mathbf{k},s} | \Psi_0 \rangle}_{= n_{\mathbf{k}s}} = 2\Omega \int \frac{d^3 k}{(2\pi)^3} \epsilon_{\mathbf{k}} n_{\mathbf{k}s} = N \frac{3}{5} \epsilon_F \quad (3.7)$$

with

$$n_{\mathbf{k}s} = \begin{cases} 1 & |\mathbf{k}| \leq k_F \\ 0 & |\mathbf{k}| > k_F \end{cases} \quad (3.8)$$

where $N = \Omega n$ is the number of valence electrons. Second there is the energy resulting from the Coulomb repulsion between the electrons,

$$E_{ee} = \frac{1}{2} \int d^3 r d^3 r' \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} \underbrace{\sum_{s,s'} \langle \Psi_0 | \hat{\Psi}_s^\dagger(\mathbf{r}) \hat{\Psi}_{s'}^\dagger(\mathbf{r}') \hat{\Psi}_{s'}(\mathbf{r}') \hat{\Psi}_s(\mathbf{r}) | \Psi_0 \rangle}_{= n^2 - G(\mathbf{r} - \mathbf{r}')} = E_{\text{Hartree}} + E_{\text{Fock}}; \quad (3.9)$$

The third contribution originates in the attractive interaction between ions and electrons,

$$E_{\text{ei}} = - \int d^3r d^3r' \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} \underbrace{n \sum_s \langle \Psi_0 | \widehat{\Psi}_s^\dagger(\mathbf{r}) \Psi_s(\mathbf{r}) | \Psi_0 \rangle}_{=n}; \quad (3.10)$$

Finally we have the repulsive ion-ion interaction

$$E_{\text{ii}} = \langle \Psi_0 | \mathcal{H}_{\text{ii}} | \Psi_0 \rangle = \frac{1}{2} \int d^3r d^3r' \frac{n^2 e^2}{|\mathbf{r} - \mathbf{r}'|}. \quad (3.11)$$

The Coulomb repulsion between the electrons leads to two terms, called the *direct* (or Hartree) term describing the Coulomb energy of a uniformly spread charge distribution, and the *exchange* or Fock term resulting from the exchange hole that follows from the Fermi-Dirac statistics (Pauli exclusion principle). The pair correlation function is ¹

$$\sum_{s,s'} \langle \Psi_0 | \widehat{\Psi}_s^\dagger(\mathbf{r}) \widehat{\Psi}_{s'}^\dagger(\mathbf{r}') \widehat{\Psi}_{s'}(\mathbf{r}') \widehat{\Psi}_s(\mathbf{r}) | \Psi_0 \rangle = n^2 - G(\mathbf{r} - \mathbf{r}') \quad (3.18)$$

with

$$G(\mathbf{r}) = \frac{9n^2}{2} \left(\frac{k_F |\mathbf{r}| \cos k_F |\mathbf{r}| - \sin k_F |\mathbf{r}|}{(k_F |\mathbf{r}|)^3} \right)^2 \quad (3.19)$$

¹Derivation of the pair correlation function: We use Eq.(3.6) and express,

$$\begin{aligned} & \langle \Psi_0 | \widehat{\Psi}_s^\dagger(\mathbf{r}) \widehat{\Psi}_{s'}^\dagger(\mathbf{r}') \widehat{\Psi}_{s'}(\mathbf{r}') \widehat{\Psi}_s(\mathbf{r}) | \Psi_0 \rangle \\ &= \frac{1}{\Omega^2} \sum_{\mathbf{k}, \mathbf{k}', \mathbf{q}, \mathbf{q}'} e^{-i(\mathbf{k}-\mathbf{k}') \cdot \mathbf{r}} e^{-i(\mathbf{q}-\mathbf{q}') \cdot \mathbf{r}'} \langle \Phi_0 | \widehat{c}_{\mathbf{k}s}^\dagger \widehat{c}_{\mathbf{q}s'}^\dagger \widehat{c}_{\mathbf{q}'s'} \widehat{c}_{\mathbf{k}'s} | \Phi_0 \rangle. \end{aligned} \quad (3.12)$$

We distinguish two cases: (1) $s \neq s'$,

$$\langle \Phi_0 | \widehat{c}_{\mathbf{k}s}^\dagger \widehat{c}_{\mathbf{q}s'}^\dagger \widehat{c}_{\mathbf{q}'s'} \widehat{c}_{\mathbf{k}'s} | \Phi_0 \rangle = \delta_{\mathbf{k}\mathbf{k}'} \delta_{\mathbf{q}\mathbf{q}'} n_{\mathbf{k}s} n_{\mathbf{q}s'} \quad (3.13)$$

and (2) $s = s'$,

$$\langle \Phi_0 | \widehat{c}_{\mathbf{k}s}^\dagger \widehat{c}_{\mathbf{q}s}^\dagger \widehat{c}_{\mathbf{q}'s} \widehat{c}_{\mathbf{k}'s} | \Phi_0 \rangle = (\delta_{\mathbf{k}\mathbf{k}'} \delta_{\mathbf{q}\mathbf{q}'} - \delta_{\mathbf{k}\mathbf{q}'} \delta_{\mathbf{q}\mathbf{k}'}) n_{\mathbf{k}s} n_{\mathbf{q}s}, \quad (3.14)$$

which lead to

$$\langle \Psi_0 | \widehat{\Psi}_s^\dagger(\mathbf{r}) \widehat{\Psi}_{s'}^\dagger(\mathbf{r}') \widehat{\Psi}_{s'}(\mathbf{r}') \widehat{\Psi}_s(\mathbf{r}) | \Psi_0 \rangle = \frac{1}{\Omega^2} \sum_{\mathbf{k}, \mathbf{q}} n_{\mathbf{k}s} n_{\mathbf{q},s'} = \frac{n^2}{4} \quad (3.15)$$

for $s \neq s'$ and

$$\langle \Psi_0 | \widehat{\Psi}_s^\dagger(\mathbf{r}) \widehat{\Psi}_s^\dagger(\mathbf{r}') \widehat{\Psi}_s(\mathbf{r}') \widehat{\Psi}_s(\mathbf{r}) | \Psi_0 \rangle = \frac{1}{\Omega^2} \sum_{\mathbf{k}, \mathbf{q}} \left\{ 1 - e^{i(\mathbf{q}-\mathbf{k}) \cdot (\mathbf{r}-\mathbf{r}')} \right\} n_{\mathbf{k}s} n_{\mathbf{q},s} \quad (3.16)$$

for $s = s'$. This leads eventually to Eq.(3.18) with

$$\begin{aligned} G(\mathbf{r}) &= 2 \left\{ \frac{1}{\Omega} \sum_{\mathbf{k}} e^{i\mathbf{k} \cdot \mathbf{r}} n_{\mathbf{k}s} \right\}^2 = 2 \left\{ \int_{|\mathbf{k}| \leq k_F} \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k} \cdot \mathbf{r}} \right\}^2 \\ &= 2 \left\{ \frac{1}{2\pi^2 r} \int_0^{k_F} dk k \sin kr \right\}^2 = 2 \left\{ \frac{1}{2\pi^2} \frac{\sin k_F r - k_F r \cos k_F r}{r^3} \right\}^2 \end{aligned} \quad (3.17)$$

and $n = k_F^3/3\pi^2$ ($k = |\mathbf{k}|$ and $r = |\mathbf{r}|$).

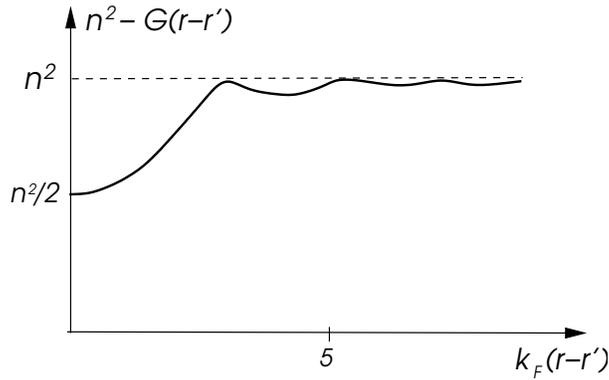


Fig. 3.1: Pair correlation function.

it is easy to verify by direct computation that the $E_{\text{Hartree}} + E_{\text{ei}} + E_{\text{ii}} = 0$ (these are the terms that would occur in a classical electrostatic calculation, implying that the binding energy of metals is a quantum effect), so that only the kinetic energy and the Fock term remain. The latter is negative,

$$E_{\text{Fock}} = -\Omega \frac{9n^2}{4} \int d^3r \frac{e^2}{|\mathbf{r}|} \left\{ \frac{\sin k_F |\mathbf{r}| - k_F |\mathbf{r}| \cos k_F |\mathbf{r}|}{(k_F |\mathbf{r}|)^3} \right\}^2 = -N \frac{3e^2}{4\pi} k_F. \quad (3.20)$$

Thus, the total energy per electron is given by

$$\frac{E_g}{N} = \frac{3}{5} \frac{\hbar^2 k_F^2}{2m} - \frac{3e^2}{4\pi} k_F = \left\{ \frac{2.21}{r_s^2} - \frac{0.916}{r_s} \right\} \text{Ry} \quad (3.21)$$

where we defined dimensionless quantity r_s by

$$n = \frac{3}{4\pi d^3} \Rightarrow r_s = \frac{d}{a_B} = \left(\frac{9\pi}{4} \right)^{1/3} \frac{m e^2}{\hbar^2 k_F}, \quad (3.22)$$

where d is the average radius of the volume occupied by one electron and $1\text{Ry} = e^2/2a_B$. This energy can be minimized with respect to r_s , yielding $r_{s,\text{min}} = 4.83 \Rightarrow d \approx 2.41\text{\AA}$. This corresponds to a lattice constant of $a = (4\pi/3)^{1/3} d \approx 3.9\text{\AA}$. This estimate is roughly in agreement with the actual lattice constants of the Alkali metals (group I, first column of the periodic table): $r_{s,\text{Li}} = 3.22$, $r_{s,\text{Na}} = 3.96$, $r_{s,\text{K}} = 4.86, \dots$. Apparently, the delocalized electrons lead to cohesion of the positive background. The good agreement of this simple estimate with the experimental values is due to the fact that for the Alkali metals have only one valence electron in an s-orbital that is delocalized, whereas the the core electrons are in a noble gas configuration and thus relatively inert.

Note that in the variational approach outlined above correlation effects among the electrons due to the Coulomb repulsion have been neglected. In particular, electrons can be expected to 'avoid' each other not just because of the Pauli principle, but also as a result of the repulsive interaction. However, for the problem under consideration this effect turns out to be small for $r_s \sim r_{s,\text{min}}$:

$$\frac{E_{\text{tot}}}{NRy} = \frac{2.21}{r_s^2} - \frac{0.916}{r_s} + \underbrace{0.062 \ln r_s - 0.096}_{\text{correlation corrections}} + \dots \quad (3.23)$$

which can be obtained by more sophisticated quantum field theoretical analysis.

3.2 Charge excitations and dielectric function

In analogy to semiconductors, the elementary excitations of metallic systems are the electron-hole excitations, which for metals, however, can have arbitrarily small energies. One particularly

drastic consequence of this behavior is the strong screening of the long-ranged Coulomb potential. A negative test charge in a metal decreases the electron density in its vicinity, and the induced cloud of positive charges (relative to uniform charge density) weakens the Coulomb potential:

$$V(r) = \frac{e^2}{r} \quad \rightarrow \quad V'(r) = e^2 \frac{e^{-r/l}}{r} \quad (3.24)$$

i.e. the Coulomb potential is modified into the short-ranged Yukawa potential. In contrast, due to the finite energy gap for electron-hole excitations the charge distribution in semiconductors cannot adapt to perturbations easily, so that even the screened Coulomb potential is still long-ranged. As we have mentioned earlier, the semiconductor acts as a dielectric medium and its screening effects are accounted for by the polarization of localized electric dipoles,

$$V(r) = \frac{e^2}{r} \quad \rightarrow \quad V'(r) = \frac{e^2}{\epsilon r}, \quad (3.25)$$

i.e. the Coulomb potential is renormalized by the dielectric constant ϵ .

3.2.1 Response and Lindhard function

We will now investigate the response of an electron gas to a time- and position-dependent weak external potential $V_a(\mathbf{r}, t)$ in more detail based on the equation of motion. We introduce the Hamiltonian

$$\mathcal{H} = \mathcal{H}_{kin} + \mathcal{H}_V = \sum_{\mathbf{k}, s} \epsilon_{\mathbf{k}} \hat{c}_{\mathbf{k}s}^\dagger \hat{c}_{\mathbf{k}s} + \sum_s \int d^3r V_a(\mathbf{r}, t) \hat{\Psi}_s^\dagger(\mathbf{r}) \hat{\Psi}_s(\mathbf{r}) \quad (3.26)$$

where the second term is considered as a small perturbation. In a first step we consider the linear response of the system to the external potential. On this level we may restrict ourself to one Fourier component in the spatial and time dependence of the potential,

$$V_a(\mathbf{r}, t) = V_a(\mathbf{q}, \omega) e^{i\mathbf{q}\cdot\mathbf{r} - i\omega t} e^{\eta t} \quad \text{with} \quad \eta \rightarrow 0_+, \quad (3.27)$$

which includes the adiabatic switching on of the potential. To linear response this potential would induce a modulation of the electron density of the form $n_{ind}(\mathbf{r}, t) = n_0 + \delta n_{ind}(\mathbf{r}, t)$ with

$$\delta n_{ind}(\mathbf{r}, t) = \delta n_{ind}(\mathbf{q}, \omega) e^{i\mathbf{q}\cdot\mathbf{r} - i\omega t}. \quad (3.28)$$

Using Eq.(3.6) we obtain for the density operator in momentum space,

$$\hat{\rho}_{\mathbf{q}} = \sum_s \int d^3r \hat{\Psi}_s^\dagger(\mathbf{r}) \hat{\Psi}_s(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} = \frac{1}{\Omega} \sum_{\mathbf{k}, s} \hat{c}_{\mathbf{k}+\mathbf{q}s}^\dagger \hat{c}_{\mathbf{k}s} = \frac{1}{\Omega} \sum_{\mathbf{k}, s} \hat{\rho}_{\mathbf{k}, \mathbf{q}, s}, \quad (3.29)$$

which leads to

$$\mathcal{H}_V = \frac{1}{\Omega} \sum_{\mathbf{k}, \mathbf{q}, s} \hat{\rho}_{\mathbf{k}, -\mathbf{q}, s} V_a(\mathbf{q}, \omega) e^{i\mathbf{q}\cdot\mathbf{r} - i\omega t}. \quad (3.30)$$

The relevant density operator to describe the electron density is $\hat{\rho}_{\mathbf{q}}(t)$ in Heisenberg representation. We use now the equation of motion for $\hat{\rho}_{\mathbf{k}, \mathbf{q}, s}(t)$:

$$i\hbar \frac{d}{dt} \hat{\rho}_{\mathbf{k}, \mathbf{q}, s} = [\hat{\rho}_{\mathbf{k}, \mathbf{q}, s}, \mathcal{H}] = [\hat{\rho}_{\mathbf{k}, \mathbf{q}, s}, \mathcal{H}_{kin} + \mathcal{H}_V] \quad (3.31)$$

$$= \{\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}}\} \hat{\rho}_{\mathbf{k}, \mathbf{q}, s} + \left\{ \hat{c}_{\mathbf{k}s}^\dagger \hat{c}_{\mathbf{k}s} - \hat{c}_{\mathbf{k}+\mathbf{q}s}^\dagger \hat{c}_{\mathbf{k}+\mathbf{q}s} \right\} V_q(\mathbf{q}, \omega) e^{-i\omega t} e^{\eta t}.$$

we now take the thermal average $\langle \hat{A} \rangle = Tr(\hat{A} e^{-\beta \mathcal{H}})/Z$, and follow the linear response scheme by taking the same time dependence for $\hat{\rho}_{\mathbf{k}, \mathbf{q}, s}(t)$ as for the potential, so that the equation of motion reads,

$$(\hbar\omega + i\hbar\eta) \langle \hat{\rho}_{\mathbf{k}, \mathbf{q}, s} \rangle = \{\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}}\} \langle \hat{\rho}_{\mathbf{k}, \mathbf{q}, s} \rangle + (n_{0\mathbf{k}, s} - n_{0\mathbf{k}+\mathbf{q}, s}) V_a(\mathbf{q}, \omega) \quad (3.32)$$

where $n_{0\mathbf{k},s} = \langle \hat{c}_{\mathbf{k}s}^\dagger \hat{c}_{\mathbf{k}s} \rangle$ and therefore

$$\delta n_{ind}(\mathbf{q}, \omega) = \frac{1}{\Omega} \sum_{\mathbf{k},s} \langle \hat{\rho}_{\mathbf{k},\mathbf{q},s} \rangle = \frac{1}{\Omega} \sum_{\mathbf{k},s} \frac{n_{0\mathbf{k}+\mathbf{q},s} - n_{0\mathbf{k},s}}{\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}} - \hbar\omega - i\hbar\eta} V_a(\mathbf{q}, \omega). \quad (3.33)$$

This defines the dynamical linear response function $\delta n_{ind}(\mathbf{q}, \omega) = \chi_0(\mathbf{q}, \omega) V_a(\mathbf{q}, \omega)$,

$$\chi_0(\mathbf{q}, \omega) = \frac{1}{\Omega} \sum_{\mathbf{k},s} \frac{n_{0\mathbf{k}+\mathbf{q},s} - n_{0\mathbf{k},s}}{\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}} - \hbar\omega - i\hbar\eta} \quad (3.34)$$

which is known as the Lindhard function.

The density $\delta n_{ind}(\mathbf{r}, t)$ can be considered again as a source for a additional Coulomb potential which can be determined by means of the Poisson equation,

$$\nabla^2 V_{ind}(\mathbf{r}, t) = -4\pi e^2 \delta n_{ind}(\mathbf{r}, t) \quad \Rightarrow \quad V_{ind}(\mathbf{q}, \omega) = \frac{4\pi e^2}{q^2} \delta n_{ind}(\mathbf{q}, \omega). \quad (3.35)$$

Now we go one step beyond simple linear response by saying that the induced charge distribution is not only driven by the external potential V_a but actually by the potential V felt by the electrons in the metal. This is a renormalized potential and consists of V_a and V_{ind} . We determine now V in a self-consistent way,

$$V(\mathbf{q}, \omega) = V_a(\mathbf{q}, \omega) + V_{ind}(\mathbf{q}, \omega) = V_a(\mathbf{q}, \omega) + \frac{4\pi e^2}{q^2} \chi_0(\mathbf{q}, \omega) V(\mathbf{q}, \omega), \quad (3.36)$$

which leads to

$$V(\mathbf{q}, \omega) = \frac{V_a(\mathbf{q}, \omega)}{\varepsilon(\mathbf{q}, \omega)} \quad \text{with} \quad \varepsilon(\mathbf{q}, \omega) = 1 - \frac{4\pi e^2}{q^2} \chi_0(\mathbf{q}, \omega). \quad (3.37)$$

This defines the dynamical dielectric function $\varepsilon(\mathbf{q}, \omega)$ and describes the renormalization of the external potential due to the dynamical response of the electrons in the metal.

We may use this also to define the response function based on Eq.(3.35) and (3.36),²

$$\chi(\mathbf{q}, \omega) = \frac{\chi_0(\mathbf{q}, \omega)}{\varepsilon(\mathbf{q}, \omega)} = \frac{\chi_0(\mathbf{q}, \omega)}{1 - \frac{4\pi e^2}{q^2} \chi_0(\mathbf{q}, \omega)}, \quad (3.39)$$

resulting from the relation,

$$\delta n(\mathbf{q}, \omega) = \chi_0(\mathbf{q}, \omega) V(\mathbf{q}, \omega) = \chi(\mathbf{q}, \omega) V_a(\mathbf{q}, \omega). \quad (3.40)$$

The response function $\chi(\mathbf{q}, \omega)$ contains information not only about the renormalization of potentials, but also on the excitation spectrum of the metal. We may separate χ into its real and imaginary part, $\chi(\mathbf{q}, \omega) = \chi_1(\mathbf{q}, \omega) + i\chi_2(\mathbf{q}, \omega)$. Using the relation

$$\lim_{\eta \rightarrow 0^+} \frac{1}{z - i\eta} = \mathcal{P} \left(\frac{1}{z} \right) + i\pi\delta(z) \quad (3.41)$$

²Eq. (3.39) can be written in the form of a geometric series,

$$\chi(\mathbf{q}, \omega) = \chi_0(\mathbf{q}, \omega) \left[1 + \frac{4\pi e^2}{q^2} \chi_0(\mathbf{q}, \omega) + \left(\frac{4\pi e^2}{q^2} \chi_0(\mathbf{q}, \omega) \right)^2 + \dots \right]. \quad (3.38)$$

From the point of view of perturbation theory, this series corresponds to summing a limited subset of perturbative terms to infinite order. This approximation is called Random Phase Approximation (RPA) and is based on the assumption the phase relation between different particle-hole excitations entering the perturbation series are random such that interference terms vanish on the average. This approximation is used quite frequently, in particular, in the discussion of instabilities of a system towards an ordered phase.

where the Cauchy principal value of the first term has to be taken, we may separate the Lindhard function into

$$\begin{aligned}\chi_{01}(\mathbf{q}, \omega) &= \frac{1}{\Omega} \sum_{\mathbf{k}, s} \mathcal{P} \left\{ \frac{n_{0, \mathbf{k}+\mathbf{q}} - n_{0, \mathbf{k}}}{\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}} - \hbar\omega} \right\} \\ \chi_{02}(\mathbf{q}, \omega) &= \frac{1}{\Omega} \sum_{\mathbf{k}, s} (n_{0, \mathbf{k}+\mathbf{q}} - n_{0, \mathbf{k}}) \delta(\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}} - \hbar\omega)\end{aligned}\tag{3.42}$$

The real part will be important later in the context of instabilities of metals. The excitation spectrum is visible in the imaginary part which relates to the absorption of energy by the electrons subject to a time-dependent external perturbation.³ Note that the $\chi_{02}(\mathbf{q}, \omega)$ corresponds to Fermi's golden rule known from time-dependent perturbation theory, i.e. the transition rate from the ground state to an excited state of energy $\hbar\omega$ and momentum \mathbf{q} .

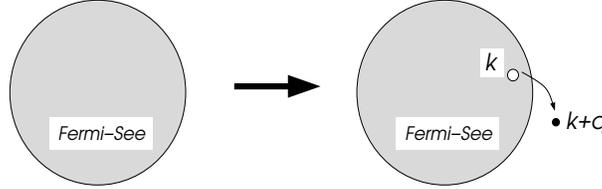


Fig. 3.2: Electron-hole excitation.

The relevant excitations for the Lindhard function are particle-hole excitations. Starting from the ground state of a completely filled Fermi sea, we remove an electron with momentum \mathbf{k} and insert it outside the Fermi sea in a state with momentum $\mathbf{k} + \mathbf{q}$ (see Fig.3.2). The energy difference is

$$E_{\mathbf{k}, \mathbf{q}} = \epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}} > 0.\tag{3.43}$$

In analogy to the semiconducting case, there is a continuum of particle-hole excitations in the energy-momentum plane. This is sketched in Fig. 3.3. Note the absence of an energy gap for excitations.

3.2.2 Collective excitation - plasma resonance

For the bare electron-hole excitations the Coulomb interaction was ignored, such that the bare Lindhard function provides information about spectrum. Including the Coulomb interaction we will show that a new collective excitation exists, the so-called plasma resonance. For a long-ranged interaction like the Coulomb interaction this resonance appears at finite frequency at small momenta \mathbf{q} . We will derive it here using the response function $\chi(\mathbf{q}, \omega)$.

Assuming here $|\mathbf{q}| \ll k_F$ we expand $\chi_0(\mathbf{q}, \omega)$ in \mathbf{q} , starting with

$$\epsilon_{\mathbf{k}+\mathbf{q}} = \epsilon_{\mathbf{k}} + \mathbf{q} \cdot \nabla_{\mathbf{k}} \epsilon_{\mathbf{k}} + \dots \quad \text{und} \quad n_{0, \mathbf{k}+\mathbf{q}} = n_{0, \mathbf{k}} + \frac{\partial n_0}{\partial \epsilon} \mathbf{q} \cdot \nabla_{\mathbf{k}} \epsilon_{\mathbf{k}} + \dots\tag{3.44}$$

Note that $\partial n_0 / \partial \epsilon_{\mathbf{k}} = -\delta(\epsilon_{\mathbf{k}} - \epsilon_F)$ at $T = 0$ and $\nabla_{\mathbf{k}} \epsilon_{\mathbf{k}} = \hbar \mathbf{v}_F = \hbar v_F \mathbf{k} / k$ is the Fermi velocity. This leads to the following approximation:

$$\begin{aligned}\chi_0(\mathbf{q}, \omega) &\approx -2 \int \frac{d^3 k}{(2\pi)^3} \frac{\mathbf{q} \cdot \mathbf{v}_F \delta(\epsilon_{\mathbf{k}} - \mu)}{\mathbf{q} \cdot \mathbf{v}_F - \omega - i\eta} \\ &\approx \frac{2}{(2\pi)^2} \int d \cos \theta \frac{k_F^2}{\hbar v_F} \left\{ \frac{q v_F \cos \theta}{\omega + i\eta} + \left(\frac{q v_F \cos \theta}{\omega + i\eta} \right)^2 + \left(\frac{q v_F \cos \theta}{\omega + i\eta} \right)^3 + \left(\frac{q v_F \cos \theta}{\omega + i\eta} \right)^4 + \dots \right\} \\ &= \frac{k_F^3 q^2}{3\pi^2 m (\omega + i\eta)^2} \left\{ 1 + \frac{3}{5} \frac{v_F^2 q^2}{(\omega + i\eta)^2} \right\} = \frac{n_0 q^2}{m (\omega + i\eta)^2} \left\{ 1 + \frac{3}{5} \frac{v_F^2 q^2}{(\omega + i\eta)^2} \right\}\end{aligned}\tag{3.45}$$

³See Chapter 6 "Linear response theory" of the course "Statistical Physics" FS09.

We now use this to approximate χ ,

$$\begin{aligned}\chi(\mathbf{q}, \omega) &\approx \frac{n_0 q^2 R(q, \omega)^2}{(\omega + i\eta)^2 - \frac{4\pi e^2 n_0^2}{m} R(q, \omega)^2} \\ &= \frac{n_0 q^2 R(q, \omega)}{\omega_p} \left\{ \frac{1}{\omega + i\eta - \omega_p R(q, \omega)} - \frac{1}{\omega + i\eta + \omega_p R(q, \omega)} \right\}\end{aligned}\quad (3.46)$$

with

$$R(q, \omega)^2 = \left(1 + \frac{3v_F^2 q^2}{5\omega^2} \right) \quad \text{and} \quad \omega_p^2 = \frac{4\pi e^2 n_0}{m}. \quad (3.47)$$

Using Eq.(3.41) from Eq.(3.45) we obtain the imaginary part as

$$Im\chi(\mathbf{q}, \omega) \approx \frac{\pi n_0 q^2 R(q, \omega_p)}{\omega_p} [\delta(\omega - \omega_p R(q, \omega_p)) - \delta(\omega + \omega_p R(q, \omega_p))] \quad (3.48)$$

which yields sharp excitation modes,

$$\omega(\mathbf{q}) = \omega_p R(q, \omega_p) = \omega_p \left\{ 1 + \frac{3v_F^2 q^2}{10\omega_p^2} + \dots \right\}, \quad (3.49)$$

which is called plasma resonance with ω_p as the *plasma frequency*.

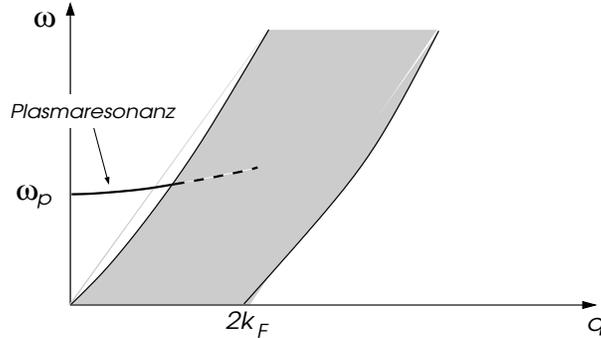


Fig. 3.3: Excitation spectrum in the ω - q -plane. The large shaded region corresponds to the electron-hole continuum and the sharp line outside the continuum represents the plasma resonance which is damped when entering the continuum.

Similar to the exciton, the plasma excitation has a well-defined energy-momentum relation and may consequently be viewed as a quasiparticle (*plasmon*) which has bosonic character. When the plasmon dispersion merges with the electron-hole continuum it is damped (Landau damping) because of the allowed decay into electron-hole excitations. The resulting finite lifetime corresponds to a finite width of the resonance of the collective excitation.

Metall	$\omega_p^{(exp)}$ (eV)	$\omega_p^{(theo)}$ (eV)
Li	7.1	8.5
Na	5.7	6.2
K	3.7	4.6
Mg	10.6	-
Al	15.3	-

Values of the plasma frequency. For the alkali metals a theoretically determined ω_p is given for comparison, using Eq.(3.47) with m the free electron mass and n determined through

$$r_{s, Li} = 3.22, \quad r_{s, Na} = 3.96 \quad \text{and} \quad R_{s, K} = 4.86.$$

It is possible to understand the plasma excitation within a classical picture. Consider negatively charged electrons in a positively charged ionic background. When the electrons are shifted by an amount \mathbf{r} w.r.t. the ions, a polarization $\mathbf{P} = -n_0 e \mathbf{r}$ results. The polarization causes an electric field $\mathbf{E} = 4\pi \mathbf{P}$ which acts as a restoring force. The equation of motion for an individual electron describes harmonic oscillations:

$$m \frac{d^2}{dt^2} \mathbf{r} = -e \mathbf{E} = -4\pi e^2 n_0 \mathbf{r} . \quad (3.50)$$

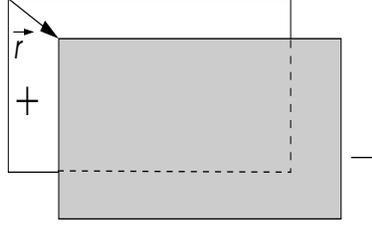


Fig. 3.4: Classical view point for the plasma excitation.

The oscillation frequency is

$$\omega_p^2 = \frac{4\pi e^2 n_0}{m} \quad (3.51)$$

the same as in eq. (3.47).

3.2.3 Screening

Now turn to the situation of a static field ($\omega = 0$). Using the expansion (3.44) we obtain

$$\chi_0(\mathbf{q}, 0) = -\frac{1}{\Omega} \sum_{\mathbf{k}, s} \delta(\epsilon_{\mathbf{k}} - \epsilon_F) = -\frac{1}{\pi^2} \frac{k_F^2}{\hbar v_F} = -\frac{3n_0}{2\epsilon_F} \Rightarrow \frac{1}{\epsilon(\mathbf{q}, 0)} = \frac{1}{1 + \frac{k_{TF}^2}{q^2}} \quad (3.52)$$

with the so-called Thomas-Fermi wave vector $k_{TF}^2 = 6\pi e^2 n_0 / \epsilon_F$. The effect of the modified (renormalized) \mathbf{q} -dependence of the dielectric function can be understood by considering the potential V_a of a point charge:

$$V(\mathbf{q}) = \frac{V_a(\mathbf{q})}{\epsilon(\mathbf{q}, 0)} = \frac{4\pi e^2}{q^2 + k_{TF}^2} \Rightarrow V(\mathbf{r}) = \frac{e^2}{r} e^{-k_{TF} r} . \quad (3.53)$$

The potential is screened by a reorganization of the electrons and turns from the long-ranged Coulomb potential into a Yukawa potential with exponential decay. The screening length is k_{TF}^{-1} , the Thomas-Fermi screening length. In metals k_{TF} is typically of the same order of magnitude as k_F , i.e. the screening length is of order 5\AA which is roughly the distance between neighboring atoms.⁴

Friedel oscillations: We can evaluate the static dielectric function for a system of free electrons, obtaining (after a simple calculation)

$$\epsilon(\mathbf{q}, 0) = 1 + \frac{4e^2 m k_F}{\pi q^2} \left\{ \frac{1}{2} + \frac{4k_F^2 - q^2}{8k_F q} \ln \left| \frac{2k_F + q}{2k_F - q} \right| \right\} . \quad (3.58)$$

⁴Thomas-Fermi approach for electron gas: The Thomas-Fermi theory for the charge distributions slowly varying in space is based on the approximation that we can always assume the electrons form locally a Fermi gas. The potential of $\rho_{ex}(\mathbf{r})$ induces a charge redistribution relative to the uniform density of electrons $n_e(E_F)$ (density of gas with Fermi energy E_F , neutralizing the ionic background charge). Within Thomas-Fermi approximation the induced charge distribution can then be written as

$$\rho_{ind}(\vec{r}) = -e \{ n_e(E_F + e\Phi(\vec{r})) - n_e(E_F) \} \quad \text{mit} \quad n_e(E_F) = \frac{k_F^3}{3\pi^2} = \frac{1}{3\pi^2 \hbar^2} (2mE_F)^{3/2} \quad (3.54)$$

where $E_F = \hbar^2 k_F^2 / 2m$. This approach is justified, if the potential $\Phi(\vec{r})$ change slowly compared to k_F^{-1} , so that classically we may describe the electron gas as filled Fermi sphere of corresponding electron density. The Poisson

Noticably the dielectric function varies little for small $\mathbf{q} \ll k_F$. At $q = \pm 2k_F$ there is a logarithmic singularity. This is a consequence of the sharpness of the Fermi surface in k -space. Consider the induced charge of a point charge at the origin.⁵

$$\delta n(\mathbf{r}) = e \int \frac{d^3q}{(2\pi)^3} \left\{ \frac{1}{\epsilon(q)} - 1 \right\} n_a(\vec{q}, 0) e^{i\mathbf{q}\cdot\mathbf{r}} = -\frac{e}{r} \int_0^\infty g(q) n_a(\vec{q}, 0) \sin qr \, dq \quad (3.60)$$

with

$$g(q) = \frac{q}{2\pi^2} \frac{\epsilon(q) - 1}{\epsilon(q)}. \quad (3.61)$$

Note that $g(q)$ vanishes for both $q \rightarrow 0$ and $q \rightarrow \infty$. Using partial integration twice, we find

$$\delta n(\mathbf{r}) = \frac{e}{r^3} \int_0^\infty g''(q) \sin qr \, dq \quad (3.62)$$

where

$$g'(q) \approx A \ln|q - 2k_F| \quad \text{and} \quad g''(q) \approx \frac{A}{q - 2k_F} \quad (3.63)$$

dominate around $q \sim 2k_F$. Hence

$$\delta n(r) \approx \frac{eA}{r^3} \int_{2k_F-\Lambda}^{2k_F+\Lambda} \frac{\sin[(q - 2k_F)r] \cos 2k_F r + \cos[(q - 2k_F)r] \sin 2k_F r}{q - 2k_F} \, dq \rightarrow \pi e A \frac{\cos 2k_F r}{r^3}. \quad (3.64)$$

with a cutoff $\Lambda \rightarrow \infty$. The induced charge distribution exhibits so-called *Friedel oscillations*.

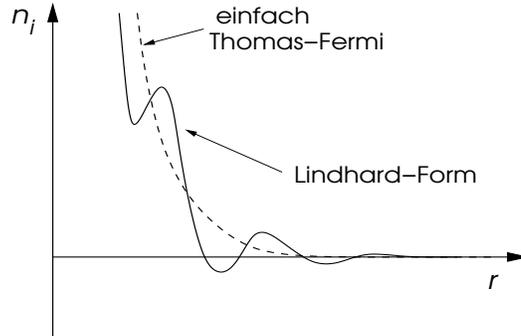


Fig. 3.5: Friedel oscillations of the charge distribution.

equation may now be formulated as

$$\begin{aligned} \vec{\nabla}^2 \Phi(\vec{r}) &= -4\pi[\rho_{ind}(\vec{r}) + \rho_{ex}(\vec{r})] \approx 4\pi e^2 \left. \frac{\partial n_e(E)}{\partial E} \right|_{E=E_F} \Phi(\vec{r}) - 4\pi \rho_{ex}(\vec{r}) \\ &= \frac{1}{l_{TF}^2} \Phi(\vec{r}) - 4\pi \rho_{ex}(\vec{r}) \end{aligned} \quad (3.55)$$

with the Thomas-Fermi screening length l_{TF} defined as,

$$l_{TF}^{-2} = 4\pi e^2 \left. \frac{\partial n_e(E)}{\partial E} \right|_{E=E_F} = \frac{6\pi e^2 n_e}{E_F}. \quad (3.56)$$

with $n_e = n_e(E_F)$ as the uniform electron density. For a point charge of magnitude Q at the origin we obtain,

$$\Phi(\vec{r}) = Q \frac{e^{-r/l_{TF}}}{r}. \quad (3.57)$$

This is the Yukawa potential as obtained above.

⁵The charge distribution can be deduced from the Poisson equation (3.35):

$$\delta n(\mathbf{q}) = \frac{q^2}{4\pi e^2} V_i(\mathbf{q}) = \chi_0(\mathbf{q}, 0) V(\mathbf{q}) = \chi_0(\mathbf{q}, 0) \frac{V_a(\mathbf{q})}{\epsilon(\mathbf{q}, 0)} = \frac{1 - \epsilon(\mathbf{q}, 0)}{\epsilon(\mathbf{q}, 0)} n_a(\vec{q}, 0) \quad (3.59)$$

The charge distribution in real space can be obtained by Fourier transformation.

Dielectric function in various dimensions: Above we have treated the dielectric function for a three-dimensional parabolic band. Similar calculations can be performed for one- and two-dimensional systems. In general, the static susceptibility is given by

$$\chi_0(q, \omega = 0) = \begin{cases} -\frac{1}{2\pi q} \ln \left| \frac{s+2}{s-2} \right|, & \text{1D} \\ -\frac{1}{2\pi} \left\{ 1 - \left(1 - \frac{4}{s^2} \right) \theta(s-2) \right\}, & \text{2D} \\ -\frac{k_F}{2\pi^2} \left\{ 1 - \frac{s}{4} \left(1 - \frac{4}{s^2} \right) \ln \left| \frac{s+2}{s-2} \right| \right\} & \text{3D} \end{cases} \quad (3.65)$$

where $s = q/k_F$. Interestingly $\chi_0(q, 0)$ has a singularity at $q = 2k_F$ for all dimensionalities. The singularity becomes weaker as the dimensionality is increased. In one dimension, there is a logarithmic divergence, in two dimensions there is a kink, and in three dimensions only the derivative diverges. Later we will see that these singularities may lead to instabilities of the metallic state, in particular for the one-dimensional case.

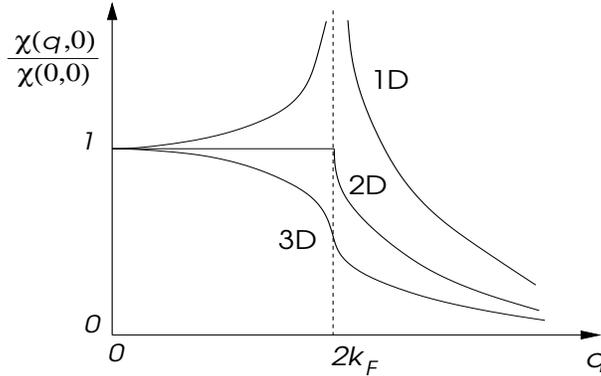


Fig. 3.6: Lindhard functions for different dimensions. The lower the dimension the stronger the singularity at $q = 2k_F$.

3.3 Lattice vibrations - phonons in metals

The atoms in a lattice of a solid can vibrate around their equilibrium positions. We will describe in the following by treating the lattice as a continuous elastic medium. This approximation is sufficient to obtain some of the essential features of the interaction between lattice vibrations and electrons, in particular screening effects. The approach is limited, however, to mono-atomic unit cells because the internal structure of a unit cell is neglected.

3.3.1 Vibration of an isotropic continuous medium

When an elastic medium is deformed an infinitesimal volume element d^3r around the point \mathbf{r} is generally moved to a different point $\mathbf{r}'(\mathbf{r})$. The deformation may be described by defining the displacement field $\mathbf{u}(\mathbf{r}) = \mathbf{r}'(\mathbf{r}) - \mathbf{r}$ at any point of the (undeformed) medium. In general, \mathbf{u} is also a function of time. In the simplest form of an isotropic medium the elastic energy for small deformations is given by

$$E_{el} = \frac{\lambda}{2} \int d^3r \{ \nabla \cdot \mathbf{u}(\mathbf{r}, t) \}^2 \quad (3.66)$$

where λ is the elastic modulus, neglecting shear contributions.⁶ The continuum form above is valid for wavelengths that are much longer than the lattice constant, so that details of the arrangement of atoms in the lattice can be neglected. The kinetic energy of the motion of the medium is

$$E_{kin} = \frac{\rho_0}{2} \int d^3r \left\{ \frac{\partial \mathbf{u}(\mathbf{r}, t)}{\partial t} \right\}^2 \quad (3.68)$$

where $\rho_0 = M_i n_i$ is the mass density (n_i : atom/ion density). Variation of the Lagrangian functional $L[\mathbf{u}] = E_{kin} - E_{el}$ leads to the equation of motion

$$\frac{1}{c_s^2} \frac{\partial^2}{\partial t^2} \mathbf{u}(\mathbf{r}, t) - \nabla(\nabla \cdot \mathbf{u}(\mathbf{r}, t)) = 0 \quad (3.69)$$

where $c_s^2 = \lambda/\rho_0$. The displacement field can be expanded into normal modes,

$$\mathbf{u}(\mathbf{r}, t) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{k}} \mathbf{e}_{\mathbf{k}} \left\{ q_{\mathbf{k}}(t) e^{i\mathbf{k} \cdot \mathbf{r}} + q_{\mathbf{k}}^*(t) e^{-i\mathbf{k} \cdot \mathbf{r}} \right\} \quad \Rightarrow \quad \frac{d^2}{dt^2} q_{\mathbf{k}} + \omega_{\mathbf{k}}^2 q_{\mathbf{k}} = 0 \quad (3.70)$$

where $\omega_{\mathbf{k}} = c_s |\mathbf{k}| = c_s k$. Note that within our simplification for the elastic energy (3.67) all modes are longitudinal wave, i.e. $\nabla \times \mathbf{u}(\mathbf{r}, t) = 0$ with $\mathbf{e}_{\mathbf{k}} \parallel \mathbf{k}$. The energy in terms of the normal modes reads

$$E = \sum_{\mathbf{k}} \rho_0 \omega_{\mathbf{k}}^2 \{ q_{\mathbf{k}}(t) q_{\mathbf{k}}^*(t) + q_{\mathbf{k}}^*(t) q_{\mathbf{k}}(t) \} . \quad (3.71)$$

Now we switch to a Hamiltonian description by defining the new variables

$$Q_{\mathbf{k}} = \sqrt{\rho_0} (q_{\mathbf{k}} + q_{\mathbf{k}}^*) \quad \text{und} \quad P_{\mathbf{k}} = \frac{d}{dt} Q_{\mathbf{k}} = -i\omega_{\mathbf{k}} \sqrt{\rho_0} (q_{\mathbf{k}} - q_{\mathbf{k}}^*) \quad (3.72)$$

in terms of which the energy is given by

$$E = \frac{1}{2} \sum_{\mathbf{k}} \{ P_{\mathbf{k}}^2 + \omega_{\mathbf{k}}^2 Q_{\mathbf{k}}^2 \} \quad (3.73)$$

Thus, the system is equivalent to an ensemble of independent harmonic oscillators, one for each normal mode \mathbf{k} . Consequently, the system may be quantized by defining the operators $P_{\mathbf{k}} \rightarrow \hat{P}_{\mathbf{k}}$ and $Q_{\mathbf{k}} \rightarrow \hat{Q}_{\mathbf{k}}$ which obey the commutation relations

$$[\hat{Q}_{\mathbf{k}}, \hat{P}_{\mathbf{k}'}] = i\hbar \delta_{\mathbf{k}, \mathbf{k}'} . \quad (3.74)$$

As usually, it is more convenient to use the raising and lowering operators

$$\hat{b}_{\mathbf{k}} = \frac{1}{\sqrt{2\hbar\omega_{\mathbf{k}}}} \left(\omega_{\mathbf{k}} \hat{Q}_{\mathbf{k}} + i\hat{P}_{\mathbf{k}} \right) \quad \text{und} \quad \hat{b}_{\mathbf{k}}^\dagger = \frac{1}{\sqrt{2\hbar\omega_{\mathbf{k}}}} \left(\omega_{\mathbf{k}} \hat{Q}_{\mathbf{k}} - i\hat{P}_{\mathbf{k}} \right) , \quad (3.75)$$

with the commutation relations

$$[\hat{b}_{\mathbf{k}}, \hat{b}_{\mathbf{k}'}^\dagger] = \delta_{\mathbf{k}, \mathbf{k}'} , \quad [\hat{b}_{\mathbf{k}}, \hat{b}_{\mathbf{k}'}] = [\hat{b}_{\mathbf{k}}^\dagger, \hat{b}_{\mathbf{k}'}^\dagger] = 0 \quad (3.76)$$

⁶Note that the most general form of the elastic energy of an isotropic medium takes the form

$$E_{el} = \int d^3r \sum_{\alpha, \beta = x, y, z} \left[\frac{\lambda}{2} (\partial_\alpha u_\alpha) (\partial_\beta u_\beta) + \mu (\partial_\alpha u_\beta) (\partial_\alpha u_\beta) \right] \quad (3.67)$$

where $\partial_\alpha = \partial/\partial r_\alpha$. The *Lamé coefficients* λ and μ characterize the elastic properties. λ describes density fluctuations which lead to longitudinal elastic waves, whereas μ corresponds to shear deformations and leads to transversely polarized elastic waves. Note that transverse elastic waves are not important for the coupling of electrons and lattice vibrations.

According to the correspondence principle, the quantum mechanical Hamiltonian corresponding to the energy (3.73) is

$$\mathcal{H} = \sum_{\mathbf{k}} \hbar\omega_{\mathbf{k}} \left\{ \widehat{b}_{\mathbf{k}}^\dagger \widehat{b}_{\mathbf{k}} + \frac{1}{2} \right\} \quad (3.77)$$

In analogy to the treatment of the electrons in second quantization we say that the operators $\widehat{b}_{\mathbf{k}}^\dagger$ ($\widehat{b}_{\mathbf{k}}$) create (annihilate) a *phonon*, a quasiparticle with well-defined energy-momentum relation, $\omega_{\mathbf{k}} = c_s |\mathbf{k}|$. The phonon operators obey bosonic commutation relations, so that the phonons behave as bosons. Using Eqs.(3.70,3.72,3.75) the displacement field operator $\widehat{\mathbf{u}}(\mathbf{r})$ can now be defined as

$$\widehat{\mathbf{u}}(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{k}} \mathbf{e}_{\mathbf{k}} \sqrt{\frac{\hbar}{2\rho_0\omega_{\mathbf{k}}}} \left[\widehat{b}_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} + \widehat{b}_{\mathbf{k}}^\dagger e^{-i\mathbf{k}\cdot\mathbf{r}} \right]. \quad (3.78)$$

As mentioned above, the continuum approximation is valid for long wavelengths (or small \mathbf{k}) only. For wavevectors with $\mathbf{k} \sim \pi/a$ the discreteness of the lattice appears in the form of corrections to the linear dispersion $\omega_{\mathbf{k}} \sim |\mathbf{k}|$. Since the number of degrees of freedom ($3 \times$ number of atoms) is limited there is a maximal wave vector called the Debye wavevector k_D .⁷ We define in this way the Debye frequency $\omega_D = c_s k_D$ and the Debye temperature $\Theta_D = \hbar\omega_D/k_B$. In the continuous medium approximation there are only acoustic phonons, for the inclusion of optical phonons the arrangement of the atoms within a unit cell has to be considered.

3.3.2 Phonons in metals

The consideration above is certainly valid for semiconductors, where ionic interactions are due to covalent chemical bonds and oscillations around the equilibrium position may be approximated by a harmonic potential, so that the form of the elastic energy above is well motivated. The situation is more subtle for metals, where the ions interact through the long-ranged Coulomb interaction and held together through the mobile conduction electrons. In this situation we can take another viewpoint and consider motion of the ions as a collective excitation of the system analogous to the electronic plasma excitation. We focus here on the limiting case $\mathbf{k} \rightarrow 0$, so that the excitation energy is given by the ionic plasma frequency:

$$\text{electrons: } \omega_p^2 = \frac{4\pi n_0 e^2}{m} \quad \Rightarrow \quad \text{ions: } \Omega_p^2 = \frac{4\pi n_i (Z_i e)^2}{M_i} \quad (3.79)$$

where $n_i = n_0/Z_i$ denotes the density of ions with charge number Z_i and mass M_i . Apparently the excitation energy does not vanish as $\mathbf{k} \rightarrow 0$.

The shortcoming in this discussion is that we neglected the (feedback) effects of the electrons that can follow the slow ionic motion "instantaneously" due to their much smaller mass. The finite plasma frequency is a consequence of the long-range nature of the Coulomb potential (as mentioned earlier), but as we have seen above the electrons tend to screen the potentials, in particular for small wavevectors \mathbf{k} . The 'bare' ionic plasma frequency is thus not what is observed in metals.

The presence of the electrons leads to a renormalization of the Coulomb potential by a factor $1/\varepsilon(\mathbf{k}, \omega)$. The restoring force of the lattice vibrations is caused by the Coulomb potential, so that the phonon frequency (which is proportional to the square root of the restoring force) is given by

$$\omega_{\mathbf{k}}^2 = \frac{\Omega_p^2}{\varepsilon(\mathbf{k}, 0)} = \frac{k^2 \Omega_p^2}{k^2 + k_{TF}^2} \approx (c_s k)^2, \quad (3.80)$$

the linear dispersion of a sound wave ($\omega_{\mathbf{k}} = c_s |\mathbf{k}|$), and the renormalized velocity of sound c_s is

$$c_s^2 \approx \frac{\Omega_p^2}{k_{TF}^2} = \frac{Z m \omega_p^2}{M_i k_{TF}^2} = \frac{1}{3} Z \frac{m}{M_i} v_F^2. \quad (3.81)$$

⁷See course of Statistical Physics HS09.

For the comparison of the energy scales we find,

$$\frac{\Theta_D}{T_F} = \frac{\hbar\omega_D}{E_F} = \frac{\hbar c_s k_D}{\frac{1}{2}\hbar k_F v_F} = \frac{2k_D}{k_F} \frac{c_s}{v_F} \sim \frac{c_s}{v_F} = \sqrt{\frac{1}{3}Z\frac{m}{M_i}} \ll 1 \quad (3.82)$$

Kohn anomaly: Note that phononic frequencies are much smaller than the (electronic) plasma frequency, so that the approximation

$$\omega_{\mathbf{k}}^2 = \frac{\Omega_p^2}{\varepsilon(\mathbf{k}, 0)} \quad (3.83)$$

is valid even for larger wavevectors. Employing the Lindhard form of $\varepsilon(\mathbf{k}, 0)$, we find that the phonon frequency is singular at $|\mathbf{k}| = 2k_F$ as noted earlier. More explicitly we have

$$\frac{\partial\omega_{\mathbf{k}}}{\partial\mathbf{k}} \rightarrow \infty \quad \text{for } k \rightarrow 2k_F. \quad (3.84)$$

This behavior is called the *Kohn anomaly* and results from the interaction between electrons and phonons. This effect is not contained in the previous elastic medium model that neglected ion-electron interactions.

3.3.3 Peierls instability in one dimension

The Kohn anomaly has particularly drastic effects in one-dimensional electron systems, where the electron-phonon coupling leads to an instability of the metallic state.

We consider a one-dimensional Jellium model and treat the ions as an elastic medium with a displacement field u along the extended direction (x -axis). We neglect the electron-electron interaction, so that the Hamiltonian reads,

$$\mathcal{H} = \sum_{\mathbf{k}, s} \frac{\hbar^2 k^2}{2m} c_{\mathbf{k}s}^\dagger c_{\mathbf{k}s} - n_0 \sum_s \int dx dx' V(x-x') \frac{d}{dx} u(x) \hat{\Psi}_s^\dagger(x') \hat{\Psi}_s(x') + \frac{\lambda}{2} \int dx \left(\frac{du}{dx}(x) \right)^2. \quad (3.85)$$

In the general theory of elastic media $\nabla \cdot \mathbf{u} = -\delta n/n_0$ describes density modulations, so that the second term in (3.85) models the coupling of the electrons to charge density fluctuations of the positively charged background.⁸ $V(x-x')$ is the screened Coulomb interaction mediating the charge density modulation of the ions to the electrons. We consider the ground state of N electrons (corresponding to a density $n = N/L$). For a uniform background, the Fermi wavevector of free electrons is readily determined to be

$$N = \sum_s \int_{-k_F}^{+k_F} dk 1 = 2 \frac{L}{2\pi} 2k_F \quad \Rightarrow \quad k_F = \frac{\pi}{2} n. \quad (3.86)$$

Now we consider the Kohn anomaly of this system by finding the renormalization of the elastic modulus in (3.85) perturbatively. The electron-phonon coupling in momentum space can be written as

$$\hat{V}_{ep} = i \sum_{\mathbf{k}, q, s} q \{ \tilde{V}_{-q} u_q \hat{c}_{\mathbf{k}+q, s}^\dagger \hat{c}_{\mathbf{k}, s} - \tilde{V}_q u_{-q} \hat{c}_{\mathbf{k}, s}^\dagger \hat{c}_{\mathbf{k}+q, s} \}, \quad (3.87)$$

where $\tilde{V}_q = 4\pi e^2/q^2 \varepsilon(q, 0)$ and

$$u(x) = \frac{1}{\sqrt{L}} \sum_q u_q e^{-iqx} \quad \text{and} \quad V(x) = \frac{1}{\sqrt{L}} \sum_q \tilde{V}_q e^{iqx}; \quad (3.88)$$

⁸Note that only phonon modes with a finite value of $\nabla \cdot \mathbf{u}$ couple in lowest order to the electrons. This is only possible of longitudinal modes. Transverse modes are defined by the condition $\nabla \cdot \mathbf{u} = 0$ and do not couple to electrons in lowest order.

We compute the second order correction to the ground state energy using Rayleigh-Schrödinger perturbation theory (note that the linear energy shift vanishes):

$$\begin{aligned}\Delta E^{(2)} &= - \sum_{k,q,s} q^2 |\tilde{V}_q|^2 u_q u_{-q} \sum_n \frac{|\langle \Psi_0 | \hat{c}_{k,s}^\dagger \hat{c}_{k+q,s} | n \rangle|^2 + |\langle \Psi_0 | \hat{c}_{k+q,s}^\dagger \hat{c}_{k,s} | n \rangle|^2}{E_n - E_0} \\ &= \sum_q |\tilde{V}_q|^2 q^2 u_q u_{-q} \sum_k \frac{n_{k+q} - n_k}{\epsilon_{k+q} - \epsilon_k} = \Omega \sum_q |\tilde{V}_q|^2 q^2 \chi_0(q, 0) u_q u_{-q}\end{aligned}\quad (3.89)$$

where the virtual states $|n\rangle$ are electron-hole excitations of the filled Fermi sea. This term is a correction to the elastic term in $(\Omega\lambda/2) \sum_q q^2 u_q u_{-q} = (\Omega\rho_0/2) \sum_q \omega_q^2 u_q u_{-q}$ in (3.85) shows that the elastic modulus and, thus, the phonon frequency is renormalized:

$$\omega_q^{(ren)2} \approx \omega_q^2 + \frac{|\tilde{V}_q|^2 q^2}{\rho_0} \chi_0(q, 0) = \omega_q^2 - \frac{|\tilde{V}_q|^2 q}{2\pi\rho_0} \ln \left| \frac{q + 2k_F}{q - 2k_F} \right| \quad (3.90)$$

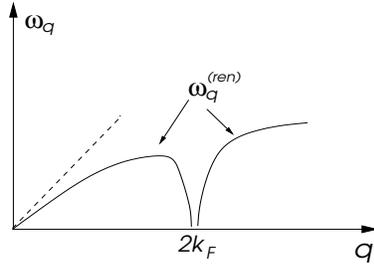


Fig. 3.7: Kohn anomaly for the one-dimensional system with electron-phonon coupling. The renormalization of the phonon frequency is divergent at $q = 2k_F$.

From the behavior for $q \rightarrow 0$ we infer that the velocity of sound is renormalized. However, a much more drastic modification occurs at $q = 2k_F$. Here the phonon spectrum is 'softened', i.e. the frequency vanishes and even becomes negative (the latter is an artifact of the perturbation theory).⁹ This hints at an instability triggered by the (Bose-Einstein) condensation of bosons with a wave vector of $q = 2k_F$. This coherent superposition of many phonons corresponds classically to a static periodic deformation of the ionic background with wave vector $2k_F$.¹⁰ The unphysical behavior of the frequency ω_q indicates that the current problem can not be dealt with using perturbation theory around the uniform state. Instead we can assume that the background shows a periodic density modulation (coherent phonon state)

$$u(x) = u_0 \cos(Qx) \quad (3.95)$$

⁹Note that the expression

$$\omega_q^2 = \frac{\Omega_p^2}{\epsilon(q, 0)} \quad (3.91)$$

in (3.83) does not yield negative energies but a zero of ω_q at $q = 2k_F$ instead.

¹⁰Coherent state: We introduce the coherent state

$$|\Phi_Q^{(coh)}\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{(\hat{b}_Q^\dagger)^n}{n!} \alpha^n |0\rangle \quad (3.92)$$

which does not have a definite phonon number for the mode of wave vector Q . On the other hand, this mode is macroscopically occupied, since

$$n_Q = \langle \Phi_Q^{(coh)} | \hat{b}_Q^\dagger \hat{b}_Q | \Phi_Q^{(coh)} \rangle = |\alpha|^2 \quad (3.93)$$

and, moreover, we find

$$\langle \Phi_Q^{(coh)} | \hat{u}(x) | \Phi_Q^{(coh)} \rangle = \frac{1}{L} \frac{\hbar}{2\rho_0\omega_Q} \left[\alpha e^{iQx} + \alpha^* e^{-iQx} \right] = u_0 \cos(Qx) \quad (3.94)$$

with $u_0 = \hbar\alpha/\rho_0 L\omega_Q$, assuming α being real.

where $Q = 2k_F$ and $u(x)$, and investigate the effect of this modulation on the electron-phonon system. To this end we show that such a modulation lowers the energy of the electrons. Assuming that u_0 is small we can evaluate the electronic energy using the approximation of nearly free electrons, treating Q as a reciprocal lattice vector. The electronic spectrum for $0 \leq k \leq Q$ is approximately determined by the secular equation

$$\det \begin{pmatrix} \frac{\hbar^2 k^2}{2m} - E & \Delta \\ \Delta^* & \frac{\hbar^2 (k-Q)^2}{2m} - E \end{pmatrix} = 0 \quad (3.96)$$

$$\Rightarrow E_{k\pm} = \frac{\hbar^2}{4m} \left[(k-Q)^2 + k^2 \pm \sqrt{\{(k-Q)^2 - k^2\}^2 + 16m^2|\Delta|^2/\hbar^4} \right]$$

Δ follows from the Fourier transform of the potential $V(x)$,

$$\Delta = -iQu_0n\tilde{V}_Q \quad \text{with} \quad \tilde{V}_Q = \int dx e^{iQx}V(x) \quad (3.97)$$

The total energy is then given by

$$E_{tot}(u_0) = 2 \sum_{0 \leq k < Q} E_{k-} + \frac{\lambda L Q^2}{4} u_0^2 \quad (3.98)$$

where all states of the lower band (E_{k-}) are occupied and all states of the upper band (E_{k+}) are empty. We minimize E_{tot} with respect to u_0 :

$$\begin{aligned} 0 = \frac{1}{L} \frac{dE_{tot}}{du_0} &= -\frac{\hbar^2}{2m} \frac{32Q^2 m^2 n^2 \tilde{V}_Q^2}{\hbar^4} u_0 \int_0^Q \frac{dk}{2\pi} \frac{1}{\sqrt{\{(k-Q)^2 - k^2\}^2 + 16m^2 Q^2 n^2 \tilde{V}_Q^2 u_0^2 / \hbar^4}} + \frac{\lambda}{2} Q^2 u_0 \\ &= -u_0 \frac{4Qmn^2 \tilde{V}_Q^2}{\hbar^2 \pi} \int_{-k_F}^{+k_F} dq \frac{1}{\sqrt{q^2 + 4m^2 n^2 \tilde{V}_Q^2 u_0^2 / \hbar^4}} + \frac{\lambda}{2} Q^2 u_0 \\ &= -u_0 \frac{8Qmn^2 \tilde{V}_Q^2}{\hbar^2 \pi} \operatorname{arsinh} \left(\frac{\hbar^2 k_F}{2mn\tilde{V}_Q u_0} \right) + \frac{\lambda}{2} Q^2 u_0. \end{aligned} \quad (3.99)$$

We solve this equation for u_0 using $\operatorname{arsinh}(x) \approx \ln(2x)$ für $x \gg 1$.

$$u_0 = \frac{\hbar^2 k_F}{mn\tilde{V}_Q} \exp \left\{ -\frac{\hbar^2 k_F \pi \lambda}{8mn^2 \tilde{V}_Q^2} \right\} = \frac{2}{k_F} \frac{\epsilon_F}{n\tilde{V}_Q} e^{-1/N(0)g} \quad (3.100)$$

where $\epsilon_F = \hbar^2 k_F^2 / 2m$ is the Fermi energy and $N(0) = 2m/\pi \hbar^2 k_F$ is the density of states at the Fermi energy. We introduce the coupling constant $g = 4n^2 \tilde{V}_Q^2 / \lambda$ that describes the phonon-induced effective electron-electron interaction. The coupling is the stronger the more polarizable (softer) ionic background, i.e. when the elastic modulus λ is small. Note that the static displacement u_0 depends exponentially on the coupling and on the density of states.

The underlying reason for this so-called *Peierls instability* to happen lies in the opening of an energy gap at $k = \pm k_F$, i.e. at the Fermi energy.

$$\Delta E = E_{k_{F+}} - E_{k_{F-}} = 2|\Delta| = 8\epsilon_F \exp \left(-\frac{1}{N(0)g} \right) \quad (3.101)$$

The gap is associated with a lowering of the energy of the electron states in the lower band in the vicinity of the Fermi energy. For this reason this kind of instability is called a Fermi surface instability. Due to the gap the metal has turned into a semiconductor with an energy gap for all electron-hole excitations.

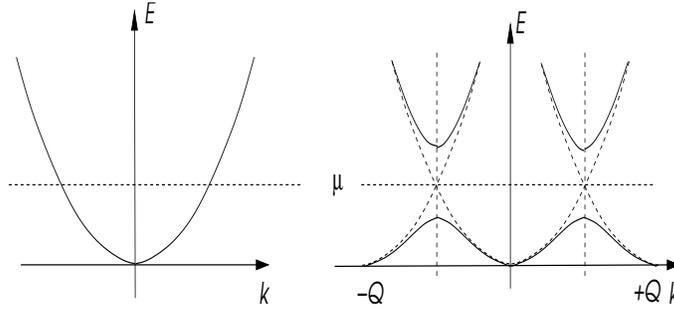


Fig. 3.8: Change of the electron spectrum. The modulation of the ionic background yields gaps at the Fermi points and the system becomes an insulator.

The modulation of the electron density follows the charge modulation due to the ionic lattice deformation, which can be seen by expressing the wave function of the electronic states,

$$\psi'_k(x) = \frac{1}{\sqrt{\Omega}} \frac{\Delta e^{ikx} + (E_k - \epsilon_k) e^{i(k-Q)x}}{\sqrt{(E_k - \epsilon_k)^2 + |\Delta|^2}} \quad (3.102)$$

which is a superposition of two plane waves with wave vectors k and $k - Q$, respectively. Hence the density is

$$\rho_k(x) = -e |\psi'_k(x)|^2 = -\frac{e}{\Omega} \left\{ 1 - \frac{2(\epsilon_k - E_k) |\Delta|}{(E_k - \epsilon_k)^2 + |\Delta|^2} \sin Qx \right\} \quad (3.103)$$

and its modulation is given by

$$\delta\rho(x) = \sum_k \rho_k(x) + en = \frac{e}{2} \int_0^{k_F} \frac{dk'}{2\pi} \frac{m|\Delta| \sin Qx}{\sqrt{\hbar^4 k_F^2 k'^2 + m^2 |\Delta|^2}} = \frac{en|\Delta|}{16\epsilon_F} \ln \left| \frac{2\epsilon_F}{|\Delta|} \right| \sin(2k_F x). \quad (3.104)$$

Such a state with a spatially modulated electronic charge density is called a charge density wave (CDW) state.

This instability is important for quasi-one-dimensional metals which are for example realized in organic conductors such as TTF·TCNQ (tetrathiafulvalene tetracyanoquinomethane). In higher dimensions the effect of the Kohn anomaly is generally less pronounced, so that in this case spontaneous deformations rarely occur. An exception that will be dealt with later are systems with a so-called nested Fermi surface that in some respects resemble one-dimensional systems. On the other hand, the electron-phonon interaction does have drastic consequences for metals that exhibit superconductivity, another kind of Fermi surface instability.

3.3.4 Phonons and the dielectric function

We have seen that an external potential V_a is screened by the polarization of the electrons. As the positively charged ionic background is polarized, too, it should also be included in the renormalization of the external potential. In general, we write

$$\varepsilon V_{ren} = V_a. \quad (3.105)$$

To proceed we define the 'bare' (unrenormalized) ionic dielectric function ε_n^{ion} and the renormalized one ε_{ren}^{ion} . The latter includes the effect of screening by the electrons onto the ionic interactions. It is illuminating to consider the renormalization from two different points of view: 1) The ionic potential is included into the external potential, so that the remaining screening is due to the electrons only:

$$\varepsilon^{el} V_{ren} = V_a + V_{ion}; \quad (3.106)$$

2) The electronic potential is included into the external potential, so that the potential is renormalized by the ions exclusively:

$$\varepsilon_n^{ion} V_{ren} = V_a + V_{el} \quad (3.107)$$

Note that in (3.107) all effects of electron polarization are included in V_{el} , so that the dielectric function results from the 'bare' ions. Adding (3.106) to (3.107) and subtracting (3.105), we obtain

$$(\varepsilon^{el} + \varepsilon_n^{ion} - \varepsilon)V_{ren} = V_a + V_{el} + V_{ion} = V_{ren} \quad \Rightarrow \quad \varepsilon = \varepsilon^{el} + \varepsilon_n^{ion} - 1 \quad (3.108)$$

In order to relate the renormalized potential to the external potential, we can make the ansatz

$$V_{ren} = \frac{1}{\varepsilon} V_a = \frac{1}{\varepsilon_{ren}^{ion}} \frac{1}{\varepsilon^{el}} V_a \quad (3.109)$$

i.e. the effective potential V_a/ε^{el} that results from polarization of the electrons is additionally screened by the ions which interact via interactions that are screened by the electrons. Using (3.108) we obtain

$$\varepsilon_{ren}^{ion} = 1 + \frac{1}{\varepsilon^{el}} (\varepsilon_n^{ion} - 1) \quad (3.110)$$

Taking into account the discussion of the plasma excitation of the bare ions in (3.47,3.79) above we approximate

$$\varepsilon_n^{ion} = 1 - \frac{\Omega_p^2}{\omega^2} \quad (3.111)$$

so that we obtain

$$\varepsilon = 1 + \frac{k_{TF}^2}{k^2} - \frac{\Omega_p^2}{\omega^2} = \left(1 + \frac{k_{TF}^2}{k^2}\right) \left(1 - \frac{\omega_{\mathbf{k}}^2}{\omega^2}\right) \quad (3.112)$$

for small wave vectors \mathbf{k} and using

$$\varepsilon^{el} = 1 + \frac{k_{TF}^2}{k^2}. \quad (3.113)$$

The Coulomb interaction between the electrons is replaced by an effective interaction:

$$V^{eff}(\mathbf{q}, \omega) = \frac{4\pi e^2}{q^2 \varepsilon(\mathbf{q}, \omega)} = \frac{4\pi e^2}{k_{TF}^2 + q^2} \left\{ 1 + \frac{\omega_{\mathbf{q}}^2}{\omega^2 - \omega_{\mathbf{q}}^2} \right\}. \quad (3.114)$$

This interaction corresponds to the matrix element for a scattering process of two electrons with momentum exchange \mathbf{q} and energy exchange ω .

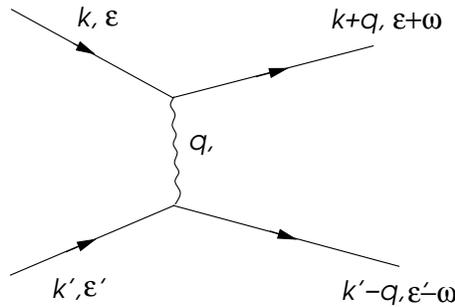


Fig. 3.9: Diagram for the electron-electron interaction involving also electron-phonon coupling.

The phonon frequency is always less than the Debye frequency ω_D . Hence the effect of the phonons is irrelevant for energy exchanges that are much larger than ω_D . The time scale for such energies would be too short for the slow ions to move and influence the interaction. Interestingly, the interaction is attractive for $\omega < \omega_D$ because of overcompensation by the ions. This aspect of the electron-phonon interaction is relevant for superconductivity.

3.4 Fermi Surfaces - the de Haas-van Alphen effect

The ground state of a metal is characterized by the existence of a discontinuity of the occupation number function $n_{\mathbf{k}}$ in momentum space - the Fermi surface. The de Haas-van Alphen experiment is one of the best methods to verify the existence and to determine the shape of a Fermi surface. It is based on the behavior of electrons at low temperatures in a strong magnetic field.

3.4.1 Landau levels

Consider a free electron gas subject to a uniform magnetic field $\mathbf{B} = (0, 0, B)$. The one-particle Hamiltonian for an electron is given by

$$\mathcal{H} = \frac{1}{2m} \left(\frac{\hbar}{i} \nabla - \frac{e}{c} \mathbf{A} \right)^2 - \frac{g\mu_B}{\hbar} \hat{S}_z B \quad (3.115)$$

where we have chosen the Landau gauge for the vector potential, $\mathbf{A} = (0, Bx, 0)$, and, as usually, $\mathbf{B} = \nabla \times \mathbf{A} = (0, 0, B)$. Hence the Hamiltonian (3.115) can be written as

$$\mathcal{H} = \frac{1}{2m} \left\{ -\hbar^2 \frac{\partial^2}{\partial x^2} + \left(\frac{\hbar}{i} \frac{\partial}{\partial y} - \frac{e}{c} Bx \right)^2 - \hbar^2 \frac{\partial^2}{\partial z^2} \right\} - \frac{g\mu_B}{\hbar} \hat{S}_z B. \quad (3.116)$$

The vector potential in this gauge acts like a confining harmonic potential along the x -axis. As translational invariance in the y - and z -directions is maintained, the eigenfunctions take the form

$$\psi(\mathbf{r}) = e^{ik_z z} e^{ik_y y} \phi(x) \xi_s \quad (3.117)$$

where ξ_s is the spin wave function. The eigenstates are then found from the solution of the harmonic oscillator problem, so that we have

$$\phi_{n,k_y}(x) = \frac{1}{\sqrt{2^n n! 2\pi \ell^2}} H_n[(x - k_y \ell^2)/\ell] e^{-(x - k_y \ell^2)^2/2\ell^2} \quad (3.118)$$

where $H_n(x)$ are the Hermite polynomials, $\ell^2 = \hbar c/|eB|$ (ℓ : magnetic length) and the energies are

$$E_{n,k_z,s} = \frac{\hbar^2 k_z^2}{2m} + \hbar \omega_c \left(n + \frac{1}{2} \right) - \frac{g\mu_B}{\hbar} B s \quad (3.119)$$

where $s = \pm \hbar/2$, $n = 0, 1, 2, \dots$ and we have introduced the cyclotron frequency $\omega_c = |eB|/mc$. Note that the energy does not depend on k_y . The apparently different spatial dependences of the wave functions for the x - and y -directions are merely a consequence of the chosen gauge.¹¹

The fact that the energy does not depend on k_y in the chosen gauge indicates a huge degeneracy of the eigenstates. To obtain the number of degenerate states we concentrate on $k_z = 0$ and neglect the electron spin. We take the electrons to be confined to a cube of volume $L \times L \times L$ with periodic boundary conditions. The wave function $\phi(x)$ is centered around $k_y \ell^2$ with the conditions $0 < k_y \ell^2 < L$ and $k_y = 2\pi n_y/L$. The degeneracy is given by the number of different values for n_y ,

$$0 < \frac{2\pi n_y \ell^2}{L} < L \quad \Rightarrow \quad N_n = \frac{L^2}{2\pi \ell^2} \quad (3.121)$$

¹¹Like the vector potential, the wave function is a gauge dependent quantity. To see this, observe that under a gauge transformation the wave function undergoes a position dependent phase shift:

$$\mathbf{A}(\mathbf{r}, t) \rightarrow \mathbf{A}'(\mathbf{r}, t) = \mathbf{A}(\mathbf{r}, t) + \nabla \chi(\mathbf{r}, t) \quad \Rightarrow \quad \psi(\mathbf{r}, t) \rightarrow \psi'(\mathbf{r}, t) = \psi(\mathbf{r}, t) e^{i\hbar c \chi(\mathbf{r}, t)/e} \quad (3.120)$$

The energies correspond to a discrete set of one-dimensional systems, so that the density of states is determined by the structure of the one-dimensional dispersion (with square root singularities at the band edges) along the z -direction:

$$\begin{aligned} N_0(E, n, s) &= \frac{N_n}{\Omega} \sum_{k_z} \delta(E - E_{n, k_z, s}) = \frac{1}{2\pi\ell^2} \int \frac{dk_z}{2\pi} \delta\left(E - \frac{\hbar^2 k_z^2}{2m} - \hbar\omega_c \left(n + \frac{1}{2}\right) + \frac{g\mu_B}{\hbar} B s\right) \\ &= \frac{(2m)^{3/2} \omega_c}{8\pi^2 \hbar^2} \frac{1}{\sqrt{E - \hbar\omega_c(n + 1/2) + g\mu_B B s/\hbar}} \end{aligned} \quad (3.122)$$

The total density of states is obtained by summing over $n = 0, 1, 2, \dots$ and $s = \pm\hbar/2$. This should be compared to the density of states without the magnetic field,

$$N_0(E) = \frac{1}{\Omega} \sum_{\mathbf{k}, s} \delta\left(E - \frac{\hbar^2 \mathbf{k}^2}{2m}\right) = \frac{(2m)^{3/2}}{2\pi^2 \hbar^3} \sqrt{E} \quad (3.123)$$

The density of states for one spin-component is shown in Fig. 3.10.

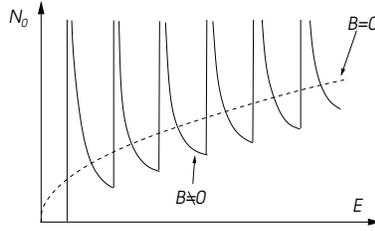


Fig. 3.10: Density of states for electrons in a magnetic field due to Landau levels. The dashed line shows the density of states in the absence of a magnetic field.

3.4.2 Oscillatory behavior of the magnetization

In the presence of a magnetic field, the smooth density of states of the three-dimensional metal is replaced by a discontinuous form dominated by square root singularities. The position of the singularities depends on the strength of the magnetic field. In order to understand the resulting effect on the magnetization, we consider the free energy

$$F = N\mu - TS = N\mu - k_B T \sum_{k_z, k_y, n, s} \ln\left(1 + e^{-(E_{n, k_z, s} - \mu)/k_B T}\right) \quad (3.124)$$

and use the general thermodynamic relation $M = -\partial F/\partial B$. For the details of the tedious calculation, we refer e.g. to J. M. Ziman, "Principles of the Theory of Solids" (German: "Prinzipien der Festkörpertheorie"), and merely present the result:

$$M = N\chi_P B \left[1 - \frac{1}{3} + \frac{\pi k_B T}{\mu_B B} \sqrt{\frac{\epsilon_F}{\mu_B B}} \sum_{\nu=1}^{\infty} \frac{1}{\sqrt{\nu}} \frac{\sin\left(\frac{\pi}{4} - \frac{\pi\nu\epsilon_F}{\mu_B B}\right)}{\sinh\left(\frac{\pi^2\nu k_B T}{\mu_B B}\right)} \right]. \quad (3.125)$$

Here χ_P is the Pauli-spin susceptibility (originating in the Zeeman-term) and the second term $-1/3\chi_P =: \chi_L$ is the diamagnetic Landau susceptibility that is due to induced orbital currents (the Landau levels).

For sufficiently low temperatures, $k_B T < \mu_B B$, the magnetization as a function of the applied field exhibits oscillatory behavior. The dominant contribution comes from the summand with $\nu = 1$. The oscillations are a consequence of the singularities in the density of states that influence the magnetic moment upon successively passing through the Fermi energy as the magnetic field is varied.

$$\frac{\pi\epsilon_F}{\mu_B} \Delta\left(\frac{1}{B}\right) = 2\pi \quad \Rightarrow \quad \Delta\left(\frac{1}{B}\right) = \frac{2\mu_B}{\epsilon_F} = \frac{2\hbar e}{2mc} \frac{2m}{\hbar^2 k_F^2} = \frac{2\pi e}{\hbar c} \frac{1}{A(k_F)} = \frac{1}{\Phi_0 A(k_F)} \quad (3.126)$$

where $A(k_F) = \pi k_F^2$ is the cross sectional area of the Fermi sphere perpendicular to the magnetic field and $\Phi_0 = hc/e$ the magnetic flux quantum.

3.4.3 Onsager equation

The behavior we have found for free electrons above generalizes to the case of arbitrary band structures. In this case, however, there are usually no exact solutions available. Instead we discuss the behavior of electrons within the quasiclassical approximation (Sect.1.7) and consider the closed orbits of a wave packet subject to a magnetic field. From the quasi-classical equations of motion for the center of mass of the wave packet (1.83) we have

$$\dot{\mathbf{r}} = \mathbf{v}_{\mathbf{k}} = \frac{1}{\hbar} \frac{\partial \epsilon_{\mathbf{k}}}{\partial \mathbf{k}} \quad \text{and} \quad \hbar \dot{\mathbf{k}} = -\frac{e}{c} \mathbf{v}_{\mathbf{k}} \times \mathbf{B}. \quad (3.127)$$

The time needed for travelling along a path between \mathbf{k}_1 and \mathbf{k}_2 is given by

$$t_2 - t_1 = \int_{\mathbf{k}_1}^{\mathbf{k}_2} dk \frac{1}{|\dot{\mathbf{k}}|} = \frac{\hbar c}{eB} \int_{\mathbf{k}_1}^{\mathbf{k}_2} \frac{dk}{|\mathbf{v}_{\mathbf{k},\perp}|} \quad (3.128)$$

where $\mathbf{v}_{\mathbf{k},\perp}$ denotes the component of the velocity that is perpendicular to \mathbf{B} . Let $\Delta \mathbf{k}$ be a vector in the plane of the motion that is both perpendicular to $\dot{\mathbf{k}}$ and \mathbf{B} , and that points from an orbit of energy ϵ to one with energy $\epsilon + \Delta \epsilon$. Then, we have

$$\Delta \epsilon = \frac{\partial \epsilon}{\partial \mathbf{k}} \cdot \Delta \mathbf{k} = \frac{\partial \epsilon}{\partial \mathbf{k}_{\perp}} \cdot \Delta \mathbf{k} = \left| \frac{\partial \epsilon}{\partial \mathbf{k}_{\perp}} \right| |\Delta \mathbf{k}| = \hbar |\mathbf{v}_{\mathbf{k},\perp}| |\Delta \mathbf{k}| \quad \Rightarrow \quad \frac{1}{|\mathbf{v}_{\mathbf{k},\perp}|} = \frac{\hbar |\Delta \mathbf{k}|}{\Delta \epsilon}, \quad (3.129)$$

because $\partial \epsilon / \partial \mathbf{k}_{\perp}$ and $\Delta \mathbf{k}$ are perpendicular to orbits of constant energy. Hence

$$t_2 - t_1 = \frac{\hbar^2 c}{eB} \frac{1}{\Delta \epsilon} \int_{\mathbf{k}_1}^{\mathbf{k}_2} |\Delta \mathbf{k}| dk = \frac{\hbar^2 c}{eB} \frac{\Delta A_{12}}{\Delta \epsilon} \quad \Rightarrow \quad t_2 - t_1 = \frac{\hbar^2 c}{eB} \frac{\partial A_{1,2}}{\partial \epsilon} \quad (3.130)$$

for infinitesimal $\Delta \mathbf{k}$.

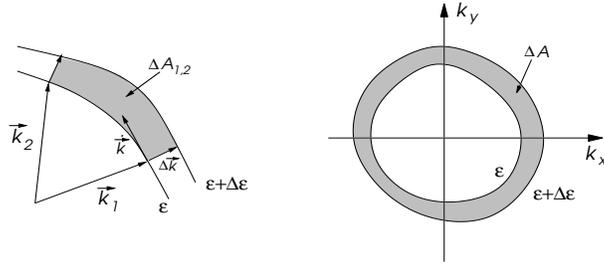


Fig. 3.11: Motion of electrons in \mathbf{k} -space. The shaded area shows the area covered by the displacement vector $\Delta \mathbf{k}$ during the motion.

Here $\Delta A_{1,2}$ is the (k -space) area swept by $\Delta \mathbf{k}$ when going from \mathbf{k}_1 to \mathbf{k}_2 . One period takes the time $T(\epsilon)$ with

$$T(\epsilon) = \frac{\hbar^2 c}{eB} \frac{\partial A(\epsilon)}{\partial \epsilon}. \quad (3.131)$$

Using now the discrete Landau levels with energies E_{n,k_z} , we can infer from Bohr's correspondence principle the following relation when the number of the Landau levels involved is large:

$$E_{n+1,k_z} - E_{n,k_z} = \frac{h}{T(E_{n,k_z}, k_z)}. \quad (3.132)$$

This result states that the difference between the energies of adjacent energy levels is given by the inverse period of classical closed orbits. As we are interested in the energy levels close to the Fermi energy ($E_{n,k_z} \sim \epsilon_F$) we have

$$n \sim \frac{\epsilon_F}{\hbar \omega_c} \gg 1 \quad (3.133)$$

Invoking (3.131) and (3.132) we can show that

$$\Delta A = A(E_{n+1, k_z}) - A(E_{n, k_z}) = \frac{2\pi e B}{\hbar c} \quad (3.134)$$

where we have used that to a good approximation

$$\frac{\partial A(\epsilon)}{\partial \epsilon} = \frac{A(E_{n+1, k_z}) - A(E_{n, k_z})}{E_{n+1, k_z} - E_{n, k_z}} = \frac{T e B}{\hbar \hbar c}. \quad (3.135)$$

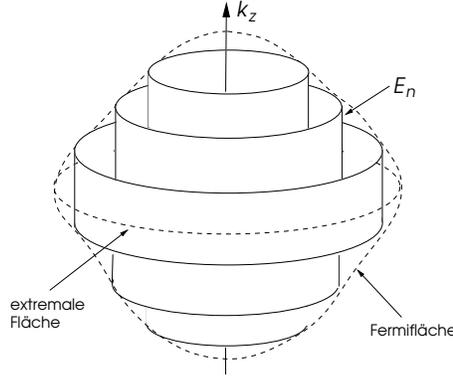


Fig. 3.12: Tubes of quantized electronic states in a magnetic field along the z-axis. A maximum of the magnetization occurs every time a tube crosses the extremal Fermi surface as the magnetic field is increased.

The area bounded by two neighboring classical orbits with quantum mechanically allowed energies is ΔA irrespective of the quantum number n . It follows that the area enclosed by one classical orbit with given quantum numbers n and k_z is

$$A(E_{n, k_z}, k_z) = (n + \gamma) \Delta A \quad (3.136)$$

where γ is an 'integration constant'. This equation is called the *Onsager equation*. The area corresponding to an extremal density of states at the Fermi surface belongs to the orbit with $E_{n, k_z} = \epsilon_F$:

$$A(\epsilon_F, k_z = 0) = \Delta A (n + \gamma) = \frac{2\pi e B}{\hbar c} (n + \gamma) \quad (3.137)$$

so that in general the period of the oscillations is given by

$$\Delta \left(\frac{1}{B} \right) = \frac{2\pi e}{\hbar c} \frac{1}{A(\epsilon_F)} \quad (3.138)$$

The oscillations in the magnetization thus allow to measure the cross sectional area of the Fermi 'sphere'. By varying the orientation of the field the topology of the Fermi surface can be mapped. As an alternative to the measurement of magnetization oscillations one can also measure resistivity oscillations known under the name Schubnikov-de Haas effect. For both methods it is crucial that the Landau levels are sufficiently clearly recognizable. Apart from low temperatures this necessitates sufficiently clean samples. In this context, sufficiently clean means that the average life-time τ (average time between two scattering events) has to be much larger than the period of the cyclotron orbits, i.e. $\omega_c \tau \gg 1$. This follows from the uncertainty relation

$$\Delta \epsilon \sim \frac{\hbar}{\tau} \ll \hbar \omega_c \quad (3.139)$$

3.5 Quantum Hall Effect

The Hall effect (discovered by Edwin Hall in 1879) is caused by the Lorentz force, a force that a magnetic field exerts on a moving charge. This force is perpendicular both to the velocity and to the magnetic field. For an electrical current, the Lorentz force leads to a transverse voltage, when a magnetic field is applied in a direction non-collinear with the current in the conductor. This so-called Hall effect can be used to investigate some properties of the charge carriers (most prominently, it can be used to distinguish between electrons and holes). Before we treat the quantum version, we briefly review the original Hall effect. To this end we consider the classical equation of motion

$$m^* \frac{d\mathbf{v}}{dt} = -e \left\{ \mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B} \right\}, \quad (3.140)$$

where m^* is the effective electron mass. For the Hall geometry shown in Fig. 3.13 with the current along the y -direction and the magnetic field in z -direction, this leads to the steady state condition $v_x = 0$, so that

$$\left\{ \mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B} \right\}_x = 0. \quad (3.141)$$

For given current and magnetic field this is a condition on E_x , the solution $E_x = -vB_z/c$ yields the Hall voltage that compensates the Lorentz force. The Hall conductivity can be defined using the definition of the current density, $\mathbf{j} = -n_0 e \mathbf{v}$:

$$\sigma_H = \frac{j_y}{E_x} = \frac{n_0 e c}{B_z} = \nu \frac{e^2}{h}, \quad (3.142)$$

where $\nu = n_0 hc/Be$. Hence knowledge of the Hall conductivity can be used to determine not only the charge density n_0 , but also the sign of the charge carriers, i.e. whether the Fermi surfaces encloses the Γ -point (electron-like, negative charge) or a point on the boundary of the Brillouin zone (hole-like, positive charge).

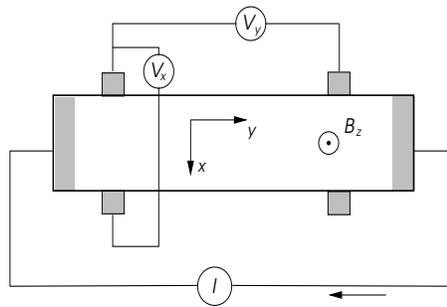


Fig. 3.13: Hall bar. The current runs along the y -direction and the magnetic field is applied along z -direction. The voltage V_y determines the conductance along the Hall bar, while V_x corresponds to the (transverse) Hall voltage.

In 1980 Klaus von Klitzing and coworkers (von Klitzing, Dorda, and Pepper, Phys. Rev. Lett. 45, 494 (1980)) made an astonishing discovery in measurements of the Hall effect in a two-dimensional electron system. This system was realized in the inversion layer of GaAs-MOSFET elements and behaves like a two-dimensional electron gas with a high mobility $b = e\tau/m^*$ (mean free path $l = 10\text{\AA}$ and low density ($n_0 \sim 10^{11}/\text{cm}^2$)). The two relevant dimensions correspond to the boundary layer, whereas the electrons are confined in the third dimension (like in a potential well, cf. section 2.5.3). For high magnetic fields of around 1 – 30T von Klitzing observed a quantization of the Hall conductivity at sufficiently low temperatures ($T < 4\text{K}$), such that the conductivity was an integer multiple of e^2/h :

$$\sigma_H = N \frac{e^2}{h} \quad \frac{e^2}{h} = \frac{1}{25812.8\Omega} \quad (3.143)$$

where $N = 1, 2, 3, \dots$. At the same time, the longitudinal conductivity σ_{yy} vanished and assumed finite values only when σ_H crossed over from one quantized value to the next.

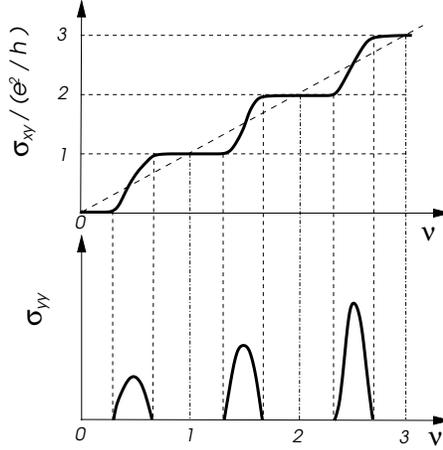


Fig. 3.14: Integer Quantum Hall effect: As a function of the filling factor ν plateaus in σ_{xy} appear at multiples of e^2/h . The longitudinal conductance σ_{yy} is only finite for fillings where σ_{xy} changes between plateaus.

In 1982 Tsui, Störmer and Gossard (Phys. Rev. Lett. 48, 1559 (1982)) discovered an additional quantization of σ_H , corresponding to certain rational multiples of e^2/h . Correspondingly, one now distinguishes between the integer quantum Hall effect (IQHE) and the fractional quantum Hall effect (FQHE). These discoveries marked the beginning of a whole new field in solid state physics that continues to produce interesting results.

3.5.1 Hall effect of the two-dimensional electron gas

The quantum mechanical treatment of the Hall conductivity takes the Hamilton operator (3.115) as its starting point. We continue to use the Landau gauge for the vector potential, $\mathbf{A} = (0, Bx, 0)$, so that we arrive again at (3.116). For the two-dimensional gas there is no motion in the z -direction (bound state in a potential well), so that the highly degenerate energy eigenvalues are given by $E_n = \hbar\omega_c(n + 1/2)$, where again $\omega_c = |eB|/m^*c$. Initially, we will concentrate on the lowest Landau level ($n = 0$) with the wave function

$$\phi_{0,k_y} = \frac{1}{\sqrt{2\pi\ell^2}} e^{-(x-k_y\ell^2)^2/2\ell^2} e^{ik_y y} . \quad (3.144)$$

where $\ell = \sqrt{\hbar c/|eB|}$ defines the magnetic length (extension of the wave function in the magnetic field). Whereas the wave function takes the form of a plane wave in y -direction the energy does not depend on k_y , leading to the huge degeneracy within the Landau levels mentioned above. In x -direction, the wave function is localized around $X = k_y\ell^2$. Note that the shape of the wave function (and thus the apparent asymmetry between x - and y -directions) is due to the particular gauge chosen and that because of the degeneracy this choice of basis is far from unique.

Now we augment the Hamiltonian by an electric field in x -direction, i.e. we add a potential $U(\mathbf{r}) = -eE_x x$. This term can be absorbed into the harmonic potential in (3.116) and leads to a shift of the center of the wave function in x -direction: $X'(k_y) = X(k_y) - eE_x/m^*\omega_c^2$. Moreover the degeneracy is lifted and the energy of the lowest Landau level takes the form (by completing the square):

$$E_{n=0}(k_y) = \frac{\hbar\omega_c}{2} - eE_x X'(k_y) + \frac{m^*}{2} \left(\frac{cE_x}{B} \right)^2 \quad (3.145)$$

for ϕ_{0k_y} . Thus the energy depends on k_y . The velocity is given by

$$v_y(k_y) = \frac{1}{\hbar} \frac{dE_0(k_y)}{dk_y} = -\frac{eE_x\ell^2}{\hbar} = -\frac{cE_x}{B\hbar} \quad (3.146)$$

From this we can determine the current density,

$$j_y = -en_0v_y(k_y) = en_0 \frac{cE_x}{B\hbar} = \frac{e\nu}{2\pi\ell^2} \frac{cE_x}{B\hbar} = \nu \frac{e^2}{h} E_x \quad (3.147)$$

where $\nu = n_0 2\pi\ell^2 \propto B^{-1}$ is the filling of the Landau level.¹² The Hall conductivity is then

$$\sigma_H = \frac{j_y}{E_x} = \nu \frac{e^2}{h}. \quad (3.148)$$

This leads to a linear relation between σ_H and $\nu \propto B^{-1}$. as derived previously based on the quasiclassical approximation.

3.5.2 Integer Quantum Hall Effect

The steps or plateaus observed by von Klitzing in the Hall conductivity σ_H of the two-dimensional electron gas as a function of the magnetic field correspond to the values $\sigma_H = Ne^2/h$, as if ν were an integer. In the plateau the longitudinal conductivity of the electron gas vanishes,

$$\sigma_{yy} = \frac{j_y}{E_y} = 0, \quad (3.149)$$

and is finite at the transition points of σ_H between two plateaus only, cf. Fig. 3.14. This fact is in contrast to the consideration above as well.

The solution to this mysterious behavior lies in the disorder that is always present in the real inversion layer. Due to the disorder the electrons move in a randomly modulated potential landscape $U(x, y)$. In two dimensions, even small amounts of disorder lead to the localization of electronic states. To investigate this new aspect we consider the lowest Landau level in the symmetric gauge $\mathbf{A} = (-y, x, 0)B/2$. The Schrödinger equation in polar coordinates is given by

$$\frac{\hbar^2}{2m^*} \left\{ -\frac{1}{r} \left(\frac{\partial}{\partial r} r \frac{\partial}{\partial r} \right) - \left(\frac{\partial}{r\partial\varphi} - i \frac{e}{2\hbar c} Br \right)^2 \right\} \psi(r, \varphi) + U(x, y)\psi(r, \varphi) = E\psi(r, \varphi) \quad (3.150)$$

Without the external potential $U(x, y)$ we find the solution

$$\psi_{n=0,m}(r, \varphi) = \frac{1}{\sqrt{2\pi\ell^2 2^m m!}} \left(\frac{r}{\ell} \right)^m e^{-im\varphi} e^{-r^2/4\ell^2} \quad (3.151)$$

where all values of $m = 0, 1, 2, 3, \dots$ correspond to the same energy $E_0 = \hbar\omega_c/2$. The wave functions are concentrated on circles of radius $r_m = \sqrt{2m\ell}$.

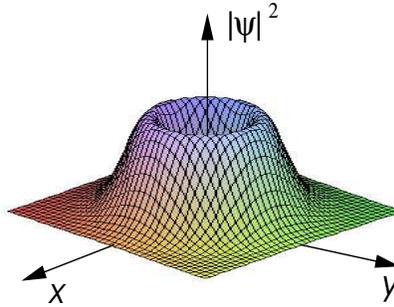


Fig. 3.15: Wavefunction of a Landau level state in the symmetric gauge.

Note that the magnetic flux threading such a circle is given by

$$\pi B r_m^2 = \pi B 2m\ell^2 = 2\pi m B \frac{\hbar c}{eB} = m \frac{\hbar c}{e} = m\Phi_0 \quad (3.152)$$

¹²Note that $\nu^{-1} = B/n_0\Phi_0$ ($\Phi_0 = hc/e$: flux quantum), i.e. ν^{-1} is the number of flux quanta Φ_0 per electron.

and that it is an integer multiple of the flux quantum $\Phi_0 = hc/e$.

Now we consider the effect of the potential. The gauge can be adjusted to the potential landscape. If we assume the potential to be rotationally invariant around the origin, the symmetric gauge is already optimal. A simple exact solution is possible for the potential

$$U(x, y) = U(r) = \frac{C_1}{r^2} + C_2 r^2 + C_3; . \quad (3.153)$$

It turns out that we obtain all eigenstates in the lowest Landau level using the ansatz

$$\tilde{\psi}_{0,m}(r, \varphi) = \frac{1}{\sqrt{2\pi\ell^{*2}2^\alpha\Gamma(\alpha+1)}} \left(\frac{r}{\ell^*}\right)^\alpha e^{-im\varphi} e^{-r^2/4\ell^{*2}} \quad (3.154)$$

with

$$\alpha^2 = m^2 + C_1^*, \quad \frac{1}{\ell^{*2}} = \frac{1}{\ell^2} \sqrt{1 + C_2^*} \quad (3.155)$$

where $C_1^* = 2m^*C_1/\hbar^2$ and $C_2^* = 8\ell^4 m^*C_2/\hbar^2$ are dimensionless parameters. The degeneracy in the Landau level is lifted:

$$E_{0,m} = \frac{\hbar\omega_c}{2} \left\{ \frac{\ell^2}{\ell^{*2}}(\alpha+1) - m \right\} + C_3 \quad (3.156)$$

The wave functions are concentrated around the radius $r_m = \sqrt{2\alpha\ell^*}$. For weak potentials $C_1^*, C_2^* \ll 1$ and $m \gg 1$ the energy is given by

$$E_{0,m} \approx \frac{\hbar\omega_c}{2} + \frac{C_1}{r_m^2} + C_2 r_m^2 + C_3 \dots , \quad (3.157)$$

i.e. the wave function adjusts itself to the potential landscape. It turns out that the same is true for arbitrarily structured potential landscapes. The wave function describes electrons on quasi-classical trajectories that trace the equipotential lines. Consequently the states described here are localized in the sense that they are attached to the structure of the potential. Application of an electric field can not set the electrons in the concentric rings in motion: The electrons are localized and do not contribute to electric transport. This holds true for any slowly varying potential.

Picture of the potential landscape: When the magnetic field is varied the filling $\nu = n_0 2\pi\ell^2$ of the Landau level is accordingly adjusted. Whereas in the translationally invariant case all states of a given level are degenerate, these states are now spread over a certain energy range. In the quasi-classical approximation, these states correspond to equipotential trajectories that are either filled or empty depending on the magnetic field, i.e. they are either below or above the chemical potential.

These considerations lead to an intuitive picture on localized and extended state trajectories. We may consider the potential landscape like a real landscape where the trajectories are contour lines. Assume that we fill now water into such a landscape. For small filling we find lakes whose shores are closed and correspond to contour lines. They are the equivalent of close trajectories and represent localized electronic states. At very high water level only the large "mountains" would reach out of the water forming islands in the sea. The coastlines again represent closed trajectories of localized electron states. At the boundary between the lakes to the islands topology as we fill water we find a water level at which the coast lines become arbitrarily long and percolate through the whole landscape. These contour lines now correspond to extended (non-localized) electron states. From this picture we conclude that when we fill gradually a Landau level in the random potential at low filling all state are localized. At some special intermediate level we fill extended states and the states at higher chemical potential would be localized again. The presence of filled extended states plays a crucial role in the following argument.

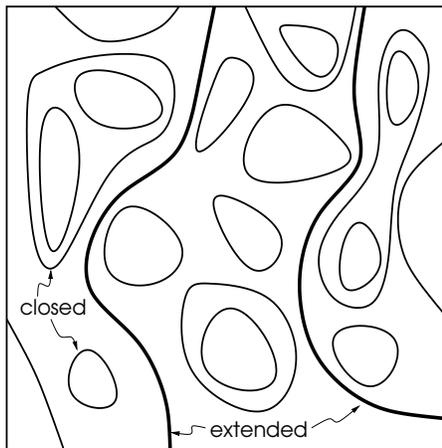


Fig. 3.16: Contour plot of potential landscape. There are closed trajectories and extended percolating trajectories.

Laughlin's gauge argument: We consider a Hall element that is deformed into a so-called Corbino geometry, i.e. a circular disc with a hole in the middle as shown in Fig. 3.17. The Hall element is threaded by a constant and uniform magnetic field B . In addition we can introduce an arbitrary flux Φ through the hole without influencing the uniform field in the disc. The flux Φ through the hole is irrelevant for all localized electron trajectories because only the extended (percolating) trajectories can wind around the hole of the disc and receive an Aharonov-Bohm phase. When the flux is increased adiabatically, the vector potential is changed by an amount

$$\delta \mathbf{A} = \nabla \chi \quad \Rightarrow \quad \delta A_\varphi = \frac{\delta \Phi}{2\pi r} \quad \Rightarrow \quad \psi \rightarrow \psi e^{ie\chi/\hbar c} = \psi e^{i\delta\Phi\varphi/\Phi_0}. \quad (3.158)$$

If the disc was translationally (clean) invariant, we could use the wave functions $\psi_{0,m}$ in (3.151), so that $B\pi r_m^2 = n\Phi_0 + \delta\Phi$. The single-valuedness of the wave function implies that m has to be adjusted, $m \rightarrow m - \delta\Phi/\Phi_0$. Increasing Φ by one flux quantum leads to a decrease of m by 1. Hence gauge invariance implies that the wave functions are shifted in their radius. The same argument may be used for the higher Landau levels.

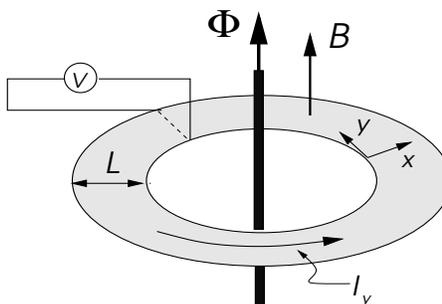


Fig. 3.17: Corbino disk for Laughlin's argument.

This argument is topological in nature and doesn't change for independent electrons when disorder is introduced. The transfer of one electron between neighboring extended states due to the change of Φ by Φ_0 leads to a net shift of one electron from the outer to the inner boundary. We assume an electric field E_x to be applied in the radial direction (x -direction), resulting in the energy change

$$\Delta\epsilon_V = -eE_x L \quad (3.159)$$

where L is the distance of the boundaries. There is a further change in electromagnetic energy because the constant current I_y in the disc when we increase the magnetic flux:

$$\Delta\epsilon_I = \frac{I_y \delta\Phi}{c}. \quad (3.160)$$

Following the Aharonov-Bohm argument that there is no change of the energy of the system when flux is change by multiples of Φ_0 , the two energies should compensate. Thus, setting $\delta\Phi = \Phi_0$ leads to

$$\Delta\epsilon_V + \Delta\epsilon_I = 0 \quad \Rightarrow \quad \sigma_H = \frac{j_y}{E_x} = \frac{I_y}{LE_x} = \frac{ec}{\Phi_0} = \frac{e^2}{h}. \quad (3.161)$$

Such a contribution to σ_H is obtained for each filled Landau level each which contributes with (percolating) extended states. Hence for n filled levels we have $\sigma_H = ne^2/h$. In summary the universality of the quantization is ensured by the topological character of the Hall conductivity.

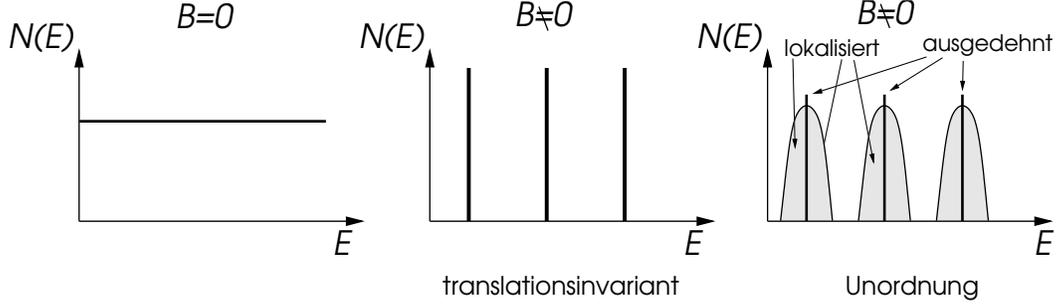


Fig. 3.18: Density of states for the two-dimensional electron gas. Left panel: without magnetic field; Middle panel: with magnetic field in a clean system, showing sharp strongly degenerate Landau levels; Right panel: with magnetic field in disordered system, showing spread Landau levels where most of the states are localized (closed trajectories) and only few states in the center percolate.

Localized and extended states: The density of states of the two-dimensional electron gas (2DEG) is given by

$$N_{\text{2DEG}}(E) = 2 \sum_{k_x, k_y} \delta \left(E - \frac{\hbar^2(k_x^2 + k_y^2)}{2m} \right) = \frac{L_x L_y m}{2\pi}, \quad (3.162)$$

whereas for the Landau levels we have

$$N_{\text{LN}}(E) = \frac{L_x L_y}{2\pi\ell^2} \sum_n \delta(E - E_n). \quad (3.163)$$

According to our previous discussion the main effect of the potential is to lift the degeneracy of the states comprising a Landau level. This remains true for random potential landscapes. Most of the states are then localized and do not contribute to electric transport. Only the few extended states can contribute to the transport when they are filled (see Fig. 3.18).

For partially filled extended states the Hall conductivity σ_H is not universal because not all states that are necessary for transferring one electron from one edge to the other, when the flux is changed by Φ_0 (in Laughlin's argument) are occupied. Thus the charge transferred does not amount to a complete $-e$. The appearance of partially filled extended states marks the transition from one plateau to the next and are accompanied by a finite longitudinal conductivity σ_{yy} . When all the extended states of a Landau level are occupied, they cannot contribute to (longitudinal) transport, i.e. in the range of a plateau σ_{yy} vanishes. Note that because of thermal occupation the plateaus shrink when the temperature is increased. This is the reason that the Quantum Hall Effect is observable only for sufficiently low temperatures ($T < 4\text{K}$).

Edge states and Büttiker's argument: The confining potential at the edge also belongs to the potential landscape. Equi-potential trajectories of states close to the edge are always extended and "percolate" along the edge. These wave functions have been discussed in section 3.5.1. From (3.145) we find that the energy is not symmetric in k_y (wave vector along the edge), i.e.

$E(k_y) \neq E(-k_y)$. This implies that the states are chiral and can move in one direction only for a given energy. The edge states on the opposing edges move in opposite directions, a fact that can be readily verified by inspection of (3.145) based on Fig.3.19.

The total current flowing along the edge for a given Landau level is

$$I = \sum_{k_y} \frac{e}{L_y} v_y, \quad (3.164)$$

i.e. we have one state per k_y that extends over the whole length of the Hall element. Thus the density is given by $1/L_y$. For the wave vector we assume quantization corresponding to periodic boundary conditions: $k_y = 2\pi n_y/L_y$ with $n_y = 0, \pm 1, \pm 2, \dots$. The velocity may be found from (3.146). In summary we have

$$I = \frac{e}{2\pi\hbar} \int_{\text{occupied}} dk_y \frac{dE_n(k_y)}{dk_y} = \frac{e}{h} \int_{\text{occupied}} dX \frac{dE}{dX} = \frac{e}{h} (\mu - E_n^{(0)}) \quad (3.165)$$

where $X = k_y \ell^2$ is the transversal position of the wave function. Sufficiently far away from the boundary E_n is independent of X and approaches the value $E_n^{(0)} = \hbar\omega_c(1/2 + n)$ of a translationally invariant electron gas. μ is the chemical potential.

The potential difference between the two opposing edges leads to a net current along the edge direction of the Hall bar,

$$\mu_A - \mu_B = eV_H = eE_x L_x = \frac{h}{e} (I_A + I_B) = \frac{h}{e} I_H \quad \Rightarrow \quad \sigma_H = \frac{I_H}{E_x L_x} = \frac{e^2}{h}. \quad (3.166)$$

where for $\mu_A = \mu_B$ we have $I_A = -I_B$. Note that $I_H = I_A + I_B$ only for the case that there are no currents in the bulk of the system, which is ensured by the localization of the states at the chemical potential.

This approach leads to the same quantization, as every Landau level contributes one edge state: $\sigma_H = ne^2/h$ (n is the number of occupied Landau levels). Note that this argument is independent of the precise shape of the confining edge potential.

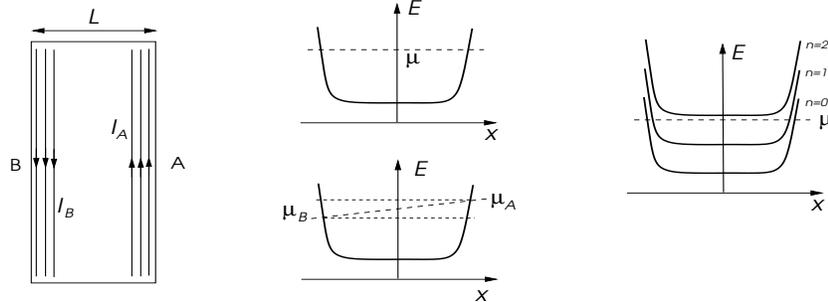


Fig. 3.19: Edge state picture: Left panel: chiral edge state exist on both edges of the Hall bar with opposite chirality; middle panel: single Landau level without and with transverse potential different, where the latter yields a finite net current due to current imbalance between left and right edge; Right panel: many Landau levels, whose number defines the Hall conductance value.

The effect on the longitudinal conductivity can be discussed from this point of view, too. However it is simpler to discuss the resistivity. Like the conductivity $\hat{\sigma}$ the resistivity $\hat{\rho}$ is a tensor:

$$\left. \begin{array}{l} \mathbf{j} = \hat{\sigma} \mathbf{E} \\ \mathbf{E} = \hat{\rho} \mathbf{j} \end{array} \right\} \Rightarrow \quad \sigma_{yy} = \frac{\rho_{yy}}{\rho_{yy}^2 + \rho_{xy}^2}, \quad \sigma_{xy} = \frac{\rho_{xy}}{\rho_{yy}^2 + \rho_{xy}^2}. \quad (3.167)$$

For a finite Hall resistivity ρ_{xy} it follows that the longitudinal resistivity $\rho_{yy} = 0$ in two dimensions.

For the edge states electrons with a given energy can move in one direction only. Thus, there is no backward scattering by obstacles as long as the edges are far apart from each other such that no scattering between the two edges can occur. Then $\rho_{yy} = 0$ and hence $\sigma_{yy} = 0$. A finite resistivity can occur only for the case that there are extended states in the bulk so that the edge states on opposite edges are no longer separated from each other.

3.5.3 Fractional Quantum Hall Effect

Only two years after the discovery of the Integer Quantum Hall Effect Störmer and Tsui observed more series of plateaus of the Hall resistivity in a 2DEG realized with very high quality MOSFET inversion layers at low temperatures. The most pronounced of these plateaus is observed at a filling of $\nu = 1/3$ ($\rho_{xy} = h/\nu e^2$ or $\sigma_{xy} = \nu e^2/h$). Afterwards an entire hierarchy of plateaus at fractional values of ν , $\nu = p/m$ (p and m are integers).

$$\nu = \frac{1}{3}, \frac{2}{3}, \frac{2}{5}, \frac{3}{5}, \frac{3}{7}, \dots \quad (3.168)$$

The emergence of these new plateaus is called the Fractional Quantum Hall Effect.

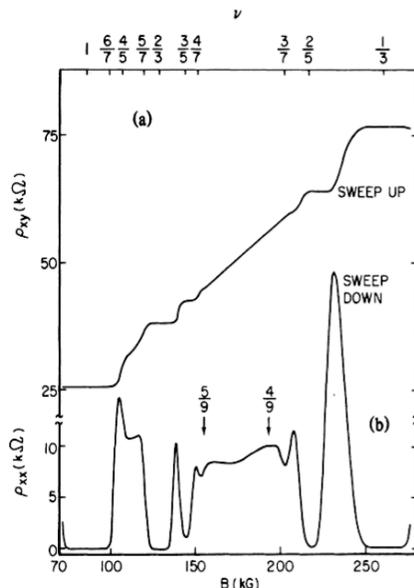


Fig. 3.20: Fractional Quantum Hall effect

Again it was Laughlin who found the key concept to explain the FQHE. Unlike the IQHE, this new quantization feature cannot be understood from a single-electron picture, and it is based on the Coulomb repulsion between the electrons and the accompanying correlation. Laughlin investigated the case $\nu = 1/3$ and made the following ansatz for the wave function:

$$\Psi_{1/m}(z_1, \dots, z_N) \propto \prod_{i < j} (z_i - z_j)^m \exp\left(-\sum_i \frac{|z_i|^2}{4\ell^2}\right) \quad (3.169)$$

where $z = x - iy$ is a complex number representing the coordinates of the two-dimensional system. We have limited ourselves to a consideration of the lowest Landau level. For the plateau at $\nu = 1/3$ the exponent is $m = 3$. This state gives a stable plateau with $\sigma_H = e^2/3h$.

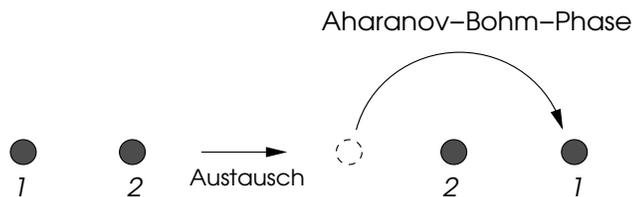


Fig. 3.21: Exchange of two particles in two dimensions involves the motion of the particles around each other. There are two topologically distinct paths.

Later Jain came up with a very intuitive interpretation of the Laughlin state based on the concept of so-called composite fermions. In fact, Laughlin's state can be written as

$$\Psi_{1/m} = \prod_{i < j} (z_i - z_j)^{m-1} \Psi_1 \quad (3.170)$$

where Ψ_1 is the Slater determinant describing the completely filled lowest Landau level.¹³ Thus $\prod_{i < j} (z_i - z_j)^{m-1}$ acts as a so-called Jastro factor that introduces correlation effects into the wave function, as Ψ_1 contains only the correlations due to the Pauli exclusion principle, but not those due to the Coulomb repulsion. The Jastro factor leads to an additional suppression of the wave function when two electrons approach each other. In the form introduced above, it leads to an additional phase factor for the electrons when they encircle each other. In particular, exchanging two electrons leads to a phase

$$\exp(i(m-1)\pi) = \exp\left(i \frac{e}{\hbar c} \frac{m-1}{2} \Phi_0\right) = 1 \quad (3.174)$$

as shown in Fig. 3.21, which holds for m odd only. Thus, the sign of the wave function still changes when two electrons are exchanged.

For the case $m = 3$ there are three flux quanta Φ_0 per electron. Now we define a composite fermion as an electron that has two negative flux quanta attached to it. These objects may be considered as independent fermions because the attached flux quanta compensate the Jastro factor through factors of the type $(z_i - z_j)^{-2}$. The exchange of two such objects in two dimensions leads to an additional Aharonov-Bohm phase that is just the opposite of the phase in (3.174). The presence of the flux $-2\Phi_0$ per electron has the consequence that the composite fermions are subject to an effective field composed of the external field and the attached flux quanta:

$$B_{eff} = \frac{1}{3}B + \left\{ \frac{2}{3}B - \sum_i 2\Phi_0(z_i) \right\} \quad (3.175)$$

For an external field of $B = 3n_0\Phi_0$ the expression in the brackets vanishes and we have an effective field of $B_{eff} = n_0\Phi_0$ (Fig. 3.22). Thus we can treat the composite fermions as independent electrons that form an Integer Quantum Hall state with $\nu = 1$ for $B = 3n_0\Phi_0$, that we have discussed above.

¹³ The Slater determinant of the lowest Landau level is obtained from the states of the independent electrons. We use the symmetric gauge, so that states are labelled by the quantum number m' . Apart from the normalization the states are given by

$$\phi_{m'}(z) = z^{m'} e^{|z|^2/4\ell^2} \quad \text{where } m' = 0, 1, 2, \dots \quad (3.171)$$

as in (3.151) ($z = x - iy$). The Slater determinant for N independent electrons is

$$\begin{aligned} \Psi_1(z_1, \dots, z_N) &= \frac{1}{\sqrt{N!}} \text{Det} \begin{bmatrix} \phi_0(z_1) & \cdots & \phi_N(z_1) \\ \vdots & & \vdots \\ \phi_0(z_N) & \cdots & \phi_N(z_N) \end{bmatrix} \\ &= \frac{1}{\sqrt{N!}} \text{Det} \begin{bmatrix} 1 & z_1 & z_1^2 & \cdots & z_1^N \\ 1 & z_2 & z_2^2 & \cdots & z_2^N \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & z_N & z_N^2 & \cdots & z_N^N \end{bmatrix} \exp\left(-\sum_i \frac{|z_i|^2}{4\ell^2}\right). \end{aligned} \quad (3.172)$$

The last determinant is a so-called Vandermonde determinant that can be written in the form of a product so that

$$\Psi_1 = \prod_{i < j} (z_i - z_j) \exp\left(-\sum_i \frac{|z_i|^2}{4\ell^2}\right) \quad (3.173)$$

where the prefactor is a homogenous polynomial with roots $z_i = z_j$ (Pauli principle). We also see that the state has a definite total angular momentum $L_z = N\hbar$.

This point of view can be applied to other Fractional Quantum Hall states, too, e.g.

$$\nu = \frac{n}{2nk + 1} \quad (3.176)$$

where we assume n filled Landau levels and consider composite fermions with attached flux of $-2k\Phi_0$ each:

$$B_{\text{eff}} = n_0 \frac{\Phi_0}{\nu} - 2kn_0\Phi_0 = n_0 \frac{\Phi_0}{n} \Rightarrow \frac{1}{n} + 2k = \frac{1}{\nu}. \quad (3.177)$$

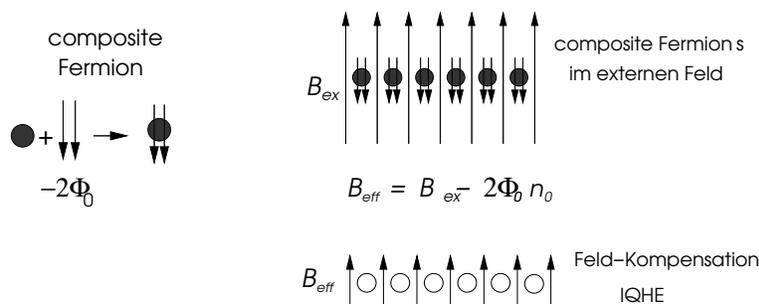


Fig. 3.22: Composite Fermions: electrons with attached magnetic flux lines, here for the state of $\nu = 1/3$.

Despite the apparent simplicity of the treatment in terms of independent composite fermions one should keep in mind that in fact one is dealing with a strongly correlated electron state. The structure of the composite fermions is a manifestation of the fact that the fermions are not independent electrons. There are no composite fermions in the vacuum, they can exist within a certain many-body state only. The Fractional Quantum Hall state also supports unconventional excitations with fractional charges. For example for the case $\nu = 1/3$ there are excitations with charge $e^* = e/3$. These are so-called 'topological' excitations, that can exist only in correlated systems. The Fractional Quantum Hall system is a very peculiar 'ordered' state of a two-dimensional electron system that has many interesting and complex properties.¹⁴

¹⁴Additional literature on the QHE:

Integer QHE: *K. von Klitzing et al.*, Physik Journal 4 (6), 37 (2005);

Fractional QHE: *R. Morf*, Physik in unserer Zeit 33, 21 (2002); *J.K. Jain*, Advances in Physics 41, 105 (1992).

Chapter 4

Landau's Theory of Fermi Liquids

Up to now, we have considered electrons as more or less independent particles, the effect of their interactions entering in terms of the renormalization of potentials and collective excitations (plasma resonance) only. The underlying assumption of our discussion was that electrons in the presence of interactions may still be described as particles with a well-defined energy-momentum-relation, and that their groundstate is a Fermi sea with a sharp Fermi surface. While there is no guarantee that this assumption holds in general (and in fact it doesn't hold always), we will show in this chapter that in metals the description of electrons as quasiparticles can be justified. This quasiparticle picture will lead us to Landau's phenomenological theory of Fermi liquids.

4.1 Life-Time of quasiparticles

We first consider the life-time of a state consisting of a filled Fermi sea to which one electron of given momentum and energy is added. Let $\hbar\mathbf{k}$ be its momentum and $\epsilon_{\mathbf{k}} = \hbar^2\mathbf{k}^2/2m$ with $|\mathbf{k}| > k_F$ and $\epsilon_{\mathbf{k}} > \epsilon_F$ its energy. The interaction between the electrons can induce this state to decay into a many-body state. In momentum space the interaction has the form

$$\mathcal{H}_{ee} = \sum_{\mathbf{k}, \mathbf{k}', \mathbf{q}} \sum_{s, s'} \tilde{V}(\mathbf{q}) \hat{c}_{\mathbf{k}-\mathbf{q}, s}^\dagger \hat{c}_{\mathbf{k}', s'}^\dagger \hat{c}_{\mathbf{k}', s'} \hat{c}_{\mathbf{k}, s} , \quad (4.1)$$

where $\tilde{V}(\mathbf{q})$ represents the interaction in momentum space (\mathbf{q} is the momentum transfer in the scattering process of two electrons). Below we will use the screened Coulomb potential,

$$\tilde{V}(\mathbf{q}) = \frac{4\pi e^2}{q^2 \epsilon(\mathbf{q}, 0)} \quad (4.2)$$

which is short-ranged. As we are only interested in very small energy transfers $\omega (\ll \epsilon_F)$ the static approximation is admissible. In a perturbative treatment, to lowest order the effect of the interaction is the creation of a particle-hole excitation in addition to the single electron above the Fermi energy.

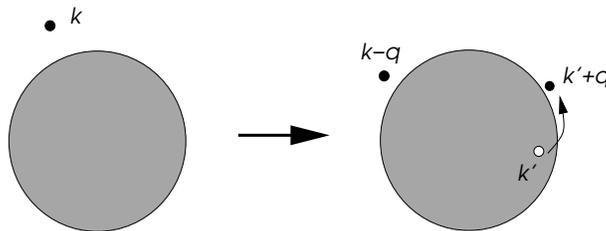


Fig. 4.1: The decay of an electron state above the Fermi energy happens through scattering by creating particle-hole excitations.

The transition is allowed whenever energy and momentum are conserved:

$$\mathbf{k} = (\mathbf{k} - \mathbf{q}) - \mathbf{k}' + (\mathbf{k}' + \mathbf{q}) \quad \text{und} \quad \epsilon_{\mathbf{k}} = \epsilon_{\mathbf{k}-\mathbf{q}} - \epsilon_{\mathbf{k}'} + \epsilon_{\mathbf{k}'+\mathbf{q}} \quad (4.3)$$

We calculate the life-time ($\tau_{\mathbf{k}}$) of the initial state with momentum \mathbf{k} using Fermi's golden rule which yields the transition rate from the initial state of a filled Fermi sea and one particle with momentum \mathbf{k} to a state with two electrons above the Fermi sea, with momenta $\mathbf{k} - \mathbf{q}$ and $\mathbf{k}' + \mathbf{q}$, and a hole with \mathbf{k}' , as shown in fig. 4.1:

$$\frac{1}{\tau_{\mathbf{k}}} = \frac{2\pi}{\hbar} \frac{1}{\Omega^2} \sum_{\mathbf{k}', \mathbf{q}} \sum_{s'} \left| \tilde{V}(\mathbf{q}) \right|^2 n_{0, \mathbf{k}'} (1 - n_{0, \mathbf{k}-\mathbf{q}}) (1 - n_{0, \mathbf{k}'+\mathbf{q}}) \delta(\epsilon_{\mathbf{k}-\mathbf{q}} - \epsilon_{\mathbf{k}} - (\epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}'+\mathbf{q}})) \quad (4.4)$$

Note that the terms $n_{0, \mathbf{k}'} (1 - n_{0, \mathbf{k}-\mathbf{q}}) (1 - n_{0, \mathbf{k}'+\mathbf{q}})$ take care of the Pauli principle, in the sense, that final state after scattering exists, i.e. the hole state (\mathbf{k}') lies inside and the two particle states ($\mathbf{k} - \mathbf{q}, \mathbf{k}' + \mathbf{q}$) lies outside the Fermi sea.

First we perform the integral running over \mathbf{k}' under the condition that the energy $\epsilon_{\mathbf{k}'+\mathbf{q}} - \epsilon_{\mathbf{k}'}$ is small. Then we can reduce the integral to the following for:

$$\begin{aligned} S(\omega_{\mathbf{q}, \mathbf{k}}, q) &= \frac{1}{\Omega} \sum_{\mathbf{k}'} n_{0, \mathbf{k}'} (1 - n_{0, \mathbf{k}'+\mathbf{q}}) \delta(\epsilon_{\mathbf{k}-\mathbf{q}} - \epsilon_{\mathbf{k}} - (\epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}'+\mathbf{q}})) \\ &= \frac{1}{(2\pi)^3} \int d^3 k' n_{0, \mathbf{k}'} (1 - n_{0, \mathbf{k}'+\mathbf{q}}) \delta(\epsilon_{\mathbf{k}'+\mathbf{q}} - \epsilon_{\mathbf{k}'} - \hbar\omega_{\mathbf{q}, \mathbf{k}}) = \frac{N(\epsilon_F)}{4} \frac{\omega_{\mathbf{q}, \mathbf{k}}}{qv_F} \end{aligned} \quad (4.5)$$

where $N(\epsilon_F) = mk_F/\pi^2\hbar^2$ is the density of states of the electrons at the Fermi surface and $\omega_{\mathbf{q}, \mathbf{k}} = \hbar(2\mathbf{k} \cdot \mathbf{q} - q^2)/2m > 0$.¹

To compute the remaining integral over \mathbf{q} we assume that the matrix element $|4\pi e^2/q^2 \varepsilon(\mathbf{q}, 0)|^2$ depends only weakly on \mathbf{q} , especially for small \mathbf{q} , i.e. the interaction is short-ranged. In spherical

¹For small ω (justified, because $\hbar\omega \leq (2k_F q - q^2)/2m$ for most allowed ω) the integral may be computed using cylindrical coordinates, where \mathbf{q} points along the axis of the cylinder:

$$S(\mathbf{q}, \omega) = \frac{1}{(2\pi)^2} \int_{k_2}^{k_1} dk'_{\perp} k'_{\perp} \int_0^{k_F} dk'_{\parallel} \delta\left(\frac{\hbar^2 q^2}{2m} + \frac{\hbar^2 q k'_{\parallel}}{m} - \hbar\omega\right) = \frac{m}{4\pi^2 \hbar^2 q} (k_1^2 - k_2^2) \quad (4.6)$$

with $k_1^2 = k_F^2 - k_{\parallel,0}^2$ und $k_2^2 = k_F^2 - (k_{\parallel,0} + q)^2$, where $k_{\parallel,0} = (2m\omega - \hbar q^2)/2\hbar q$ is enforced by the delta function.

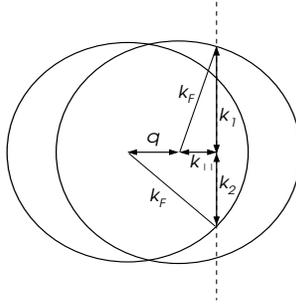


Fig. 4.2: Geometry of the conservation laws.

The wave vectors k_2 and k_1 are the upper and lower limits of integration determined from the condition $n_{0, \mathbf{k}'} (1 - n_{0, \mathbf{k}'+\mathbf{q}}) > 0$ and can be obtained by simple geometric considerations. Eq. (4.5) follows immediately.

coordinates, the integral reads

$$\begin{aligned}
\frac{1}{\tau_{\vec{k}}} &= \frac{2\pi}{\hbar} \cdot \frac{N(\epsilon_F)}{4v_F\Omega} \sum_{\vec{q}, s'} \left| \tilde{V}(\mathbf{q}) \right|^2 \frac{\omega_{\mathbf{q}, \mathbf{k}}}{q} \\
&= \frac{N(\epsilon_F)}{(2\pi)^2 2\hbar v_F} \int d^3\vec{q} \left| \tilde{V}(\mathbf{q}) \right|^2 \frac{\omega_{\mathbf{q}, \mathbf{k}}}{q} \\
&= \frac{N(\epsilon_F)}{(2\pi)4mv_F} \int dq \left| \tilde{V}(\mathbf{q}) \right|^2 q^2 \int_{\theta_1}^{\theta_2} d\theta \sin\theta (2k \cos\theta - q) \\
&= \frac{N(\epsilon_F)}{(2\pi)4mv_F} \int dq \left| \tilde{V}(\mathbf{q}) \right|^2 q^2 \left[-\frac{1}{4k} (2k \cos\theta - q)^2 \right]_{\theta_1}^{\theta_2}.
\end{aligned} \tag{4.7}$$

The restriction of the domain of integration of θ follows from the two conditions $k^2 \geq (\vec{k} - \vec{q})^2 \geq k_F^2$ and $(\vec{k} - \vec{q})^2 = k^2 - 2kq \cos\theta + q^2$. From the first condition, $\cos\theta_2 = q/2k$, and from the second, $\cos\theta_1 = (k^2 - k_F^2 + q^2)/2kq$. Thus,

$$\begin{aligned}
\frac{1}{\tau_{\vec{k}}} &= \frac{N(\epsilon_F)}{(2\pi)4mv_F} \int dq \left| \tilde{V}(\mathbf{q}) \right|^2 \frac{1}{4k} (k^2 - k_F^2)^2 \\
&\approx \frac{N(\epsilon_F)}{(2\pi)4v_F} \frac{m}{k_F} \frac{1}{\hbar^4} (\epsilon_{\vec{k}} - \epsilon_F)^2 \int dq \left| \tilde{V}(\mathbf{q}) \right|^2 \\
&= \frac{1}{8\pi\hbar^3} \frac{N(\epsilon_F)}{v_F^2} (\epsilon_{\vec{k}} - \epsilon_F)^2 \int dq \left| \tilde{V}(\mathbf{q}) \right|^2.
\end{aligned} \tag{4.8}$$

Note that convergence of the last integral over q requires that the integrand does not diverge stronger than q^α ($\alpha < 1$) for $q \rightarrow 0$. The dielectric constant obtained in the previous chapter certainly fulfills this condition.

Essentially, the result states that

$$\frac{1}{\tau_{\mathbf{k}}} \propto (\epsilon_{\mathbf{k}} - \epsilon_F)^2 \tag{4.9}$$

for \mathbf{k} slightly above the Fermi surface. This implies that the state $|\mathbf{k}s\rangle$ occurs as a resonance with decay width $\hbar/\tau_{\mathbf{k}}$. Nevertheless, we find that

$$\frac{\hbar/\tau_{\mathbf{k}}}{\epsilon_{\mathbf{k}} - \epsilon_F} \xrightarrow{k \rightarrow k_F} 0 \tag{4.10}$$

i.e. the resonance becomes arbitrarily sharp as the Fermi surface is approached, so that the quasiparticle concept is asymptotically valid in this limit. Consequently, the momentum of an electron is a good quantum number in the vicinity of the Fermi surface. Underlying this result is the Pauli exclusion principle, which restricts the phase space for decay processes of single particle states close to the Fermi surface. In addition, the assumption of short ranged interactions is crucial, and long ranged interactions can change the behavior drastically due to the larger number of decay channels.

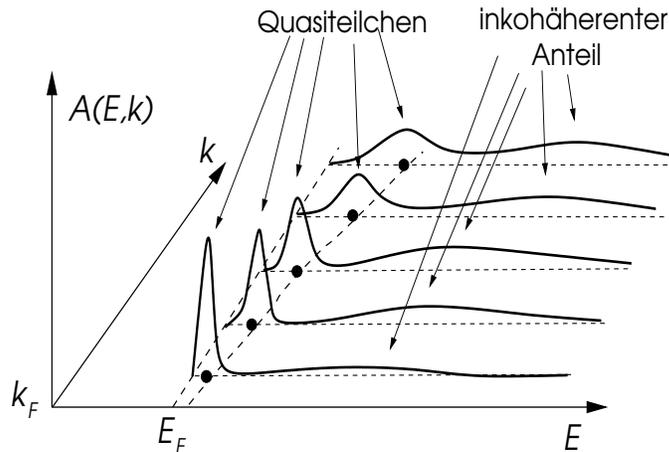


Fig. 4.3: Quasiparticle spectrum: Quasiparticle peaks appear the sharper the closer the energy lies to the Fermi energy. The area under the "sharp" quasiparticle peak corresponds to the quasiparticle weight. The missing quasiparticle weight is transferred to higher energies (incoherent part).

4.2 Phenomenological Theory of Fermi Liquids

The existence of well-defined fermionic quasiparticles in spite of the underlying complex many-body physics led Landau to the following phenomenological theory. Like the states of independent electrons, quasiparticle states can be characterized by their momentum (\mathbf{k} and spin σ). In fact, there is a one-to-one mapping of the free electrons to the quasiparticles. Consequently, the number of quasiparticles and the number of electrons coincide. We define the momentum distribution function of quasiparticles, $n_\sigma(\mathbf{k})$, subject to the condition:

$$N = \sum_{\mathbf{k},\sigma} n_\sigma(\mathbf{k}) . \quad (4.11)$$

In analogy with the Fermi-Dirac distribution of free electrons, in the ground state this distribution function is a simple step function,

$$n_\sigma^{(0)}(\mathbf{k}) = \Theta(k_F - |\mathbf{k}|) \quad (4.12)$$

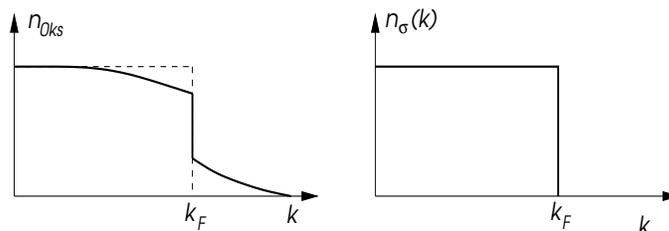


Fig. 4.4: Schematic picture of the distribution function: Left panel: modified distribution function of the original electron states; right panel: distribution function of quasiparticle states making a simple step function.

For a spherically symmetric electron system, the quasiparticle Fermi surface is a sphere with the same radius as the one for free electrons of the same density. For a general point group symmetry the Fermi surface may be deformed by the interactions without changing the underlying symmetry. The volume enclosed by the Fermi surface is always conserved despite the deformation (This is the content of the Luttinger theorem, J.M. Luttinger, Phys. Rev. 119, 1153 (1960)). Note that the distributions $n_\sigma^{(0)}(\mathbf{k})$ and $n_{0\mathbf{k}s} = \langle \hat{c}_{\mathbf{k}\sigma}^\dagger \hat{c}_{\mathbf{k}\sigma} \rangle$ are not identical (Fig. 4.3). Interestingly, $n_{0\mathbf{k}s}$ is still discontinuous at the Fermi surface, but the height of the jump

is smaller than 1, in general. The latter is a measure for the quasiparticle 'weight' at the Fermi surface, i.e. the amplitude of the corresponding free electron state in the quasiparticle state. The modification of the electron distribution function indicates the involvement of electron-hole excitations in the renormalization of the electronic properties, which deplete the Fermi sea and populate the states above the Fermi level.

Small deviations from the ground state distribution are described by

$$\delta n_\sigma(\mathbf{k}) = n_\sigma(\mathbf{k}) - n_\sigma^{(0)}(\mathbf{k}) \quad (4.13)$$

which contains the essential information on the low-energy physics of the system. Consequently, $\delta n_\sigma(\mathbf{k})$ is concentrated on momenta \mathbf{k} very close to the Fermi energy only, where the quasiparticle concept is valid. This distribution function enters a phenomenological energy functional of the following form:

$$E = E_0 + \sum_{\mathbf{k}, \sigma} \epsilon_\sigma(\mathbf{k}) \delta n_\sigma(\mathbf{k}) + \frac{1}{2\Omega} \sum_{\mathbf{k}, \mathbf{k}'} \sum_{\sigma, \sigma'} f_{\sigma\sigma'}(\mathbf{k}, \mathbf{k}') \delta n_\sigma(\mathbf{k}) \delta n_{\sigma'}(\mathbf{k}') + O(\delta n^3) \quad (4.14)$$

where E_0 denotes the energy of the ground state. Here $\epsilon_\sigma(\mathbf{k})$ and $f_{\sigma\sigma'}(\mathbf{k}, \mathbf{k}')$ are phenomenological parameters which have to be determined by experiments. The variational derivative

$$\tilde{\epsilon}_\sigma(\mathbf{k}) = \frac{\delta E}{\delta n_\sigma(\mathbf{k})} = \epsilon_\sigma(\mathbf{k}) + \frac{1}{\Omega} \sum_{\mathbf{k}', \sigma'} f_{\sigma\sigma'}(\mathbf{k}, \mathbf{k}') \delta n_{\sigma'}(\mathbf{k}') + \dots \quad (4.15)$$

yields an effective energy-momentum relation $\tilde{\epsilon}_\sigma(\mathbf{k})$, which depends on the distribution of all quasiparticles. A quasiparticle moves in the "mean-field" of all other quasiparticles, so that changes in the distribution affect $\tilde{\epsilon}_\sigma(\mathbf{k})$. The second variational derivative defines a coupling between the quasiparticles

$$\frac{\delta^2 E}{\delta n_\sigma(\mathbf{k}) \delta n_{\sigma'}(\mathbf{k}')} = \frac{1}{\Omega} f_{\sigma\sigma'}(\mathbf{k}, \mathbf{k}') . \quad (4.16)$$

We introduce a parametrization for these couplings by assuming that the dependence on \mathbf{k}, \mathbf{k}' can be reduced to the relative angle $\theta_{\hat{k}, \hat{k}'}$ (spherical symmetry). The radial dependence is ignored as we consider quasiparticles in the vicinity of the Fermi surface only.

$$f_{\sigma\sigma'}(\mathbf{k}, \mathbf{k}') = f^s(\hat{k}, \hat{k}') + \sigma\sigma' f^a(\hat{k}, \hat{k}') \quad \Rightarrow \quad f^{s,a}(\hat{k}, \hat{k}') = \sum_{l=0}^{\infty} f_l^{s,a} P_l(\cos \theta_{\hat{k}, \hat{k}'}) \quad (4.17)$$

with $\hat{k} = \mathbf{k}/|\mathbf{k}|$ and $P_l(z)$ are Legendre-polynomials (s, a stands for symmetric bzw. antisymmetric). We introduce the following parameters:

$$N(\epsilon_F) f_l^s = F_l^s \quad \text{und} \quad N(\epsilon_F) f_l^a = F_l^a , \quad (4.18)$$

where $N(\epsilon_F) = m^* k_F / \hbar^2 \pi^2$ is the density of states. The notation $F_l^s = F_l$ and $F_l^a = Z_l$ is also used frequently in the literature.

4.2.1 Specific heat and density of states

For the Fermi-Dirac distribution at low temperatures

$$\delta n_\sigma(\mathbf{k}) = n_\sigma^{(0)}(T, \mathbf{k}) - n_\sigma^{(0)}(0, \mathbf{k}) \quad \Rightarrow \quad \frac{1}{\Omega} \sum_{\mathbf{k}} \delta n_\sigma(\mathbf{k}) \propto T^2 + O(T^4). \quad (4.19)$$

Thus, to leading order one can use

$$n_\sigma(\mathbf{k}) = \frac{1}{e^{(\epsilon(\mathbf{k}) - \mu)/k_B T} + 1} \quad (4.20)$$

with $\epsilon_\sigma(\mathbf{k})$ in place of the renormalized $\tilde{\epsilon}_\sigma(\mathbf{k})$, as $\tilde{\epsilon}$ differs from ϵ at order T^2 only. When focussing on leading terms (usually T^0 and T^1) these corrections may be neglected. In order to discuss the specific heat, we employ the expression for the entropy of a fermion gas. For each quasiparticles with a given spin there is one 'orbital' labelled by \mathbf{k} . The entropy density may be computed from the distribution function.

$$S = -\frac{k_B}{\Omega} \sum_{\mathbf{k}, \sigma} \{n_\sigma(\mathbf{k}) \ln[n_\sigma(\mathbf{k})] + [1 - n_\sigma(\mathbf{k})] \ln[1 - n_\sigma(\mathbf{k})]\} \quad (4.21)$$

Taking a derivative with respect to T we obtain the specific heat,

$$\frac{C}{T} = \frac{\partial S}{\partial T} = -\frac{k_B}{\Omega} \sum_{\mathbf{k}, \sigma} \frac{e^{\xi(\mathbf{k})/k_B T}}{(e^{\xi(\mathbf{k})/k_B T} + 1)^2} \frac{\xi(\mathbf{k})}{k_B T^2} \underbrace{\ln \frac{n_\sigma(\mathbf{k})}{1 - n_\sigma(\mathbf{k})}}_{=\xi(\mathbf{k})/k_B T} \quad (4.22)$$

$= (2 \cosh(\xi(\mathbf{k})/2k_B T))^{-2}$

where we have introduced $\xi(\mathbf{k}) = \epsilon(\mathbf{k}) - \mu$. In the limit $T \rightarrow 0$

$$\frac{C}{T} \approx \frac{N(\epsilon_F)}{4k_B T^3} \int d\xi \frac{\xi^2}{\cosh^2(\xi/2k_B T)} \approx \frac{k_B^2 N(\epsilon_F)}{4} \int_{-\infty}^{+\infty} dy \frac{y^2}{\cosh^2(y/2)} = \frac{\pi^2 k_B^2 N(\epsilon_F)}{3} = \gamma, \quad (4.23)$$

which is the usual behavior $C = \gamma T$. The density of states at the Fermi surface follows from

$$\nabla_{\mathbf{k}} \epsilon(\mathbf{k})|_{\mathbf{k}_F} = \mathbf{v}_F = \frac{\hbar \mathbf{k}_F}{m^*} \Rightarrow N(\epsilon_F) = \frac{2}{\Omega} \sum_{\mathbf{k}} \delta(\epsilon(\mathbf{k}) - \epsilon_F) = \frac{k_F^2}{\pi^2 \hbar v_F} = \frac{m^* k_F}{\pi^2 \hbar^2} = \frac{3}{2} \frac{n}{\epsilon_F} \quad (4.24)$$

which defines the effective mass m^* .

4.2.2 Compressibility

A Fermi gas has a finite compressibility because each fermion occupies a certain amount of space due to the Pauli principle. The compressibility κ is defined by

$$\kappa = -\frac{1}{\Omega} \left(\frac{\partial \Omega}{\partial p} \right)_{T, N} \quad (4.25)$$

where p is the uniform (hydrostatic) pressure, which - at zero temperature - is determined by the change of the ground state energy upon changing the volume:

$$p = -\frac{\partial E}{\partial \Omega} = -\frac{\partial \Omega \hat{\epsilon}(n)}{\partial \Omega} = -\hat{\epsilon} + \frac{N}{\Omega} \frac{\partial \hat{\epsilon}}{\partial n} \Rightarrow \frac{1}{\kappa} = \Omega \frac{\partial^2 E}{\partial \Omega^2} = n^2 \frac{\partial^2 \hat{\epsilon}}{\partial n^2} \quad (4.26)$$

where $E = \Omega \hat{\epsilon}(n)$ and n is the quasiparticle density ($n = N/\Omega$). The chemical potential follows from

$$\mu = \left(\frac{\partial E}{\partial N} \right)_{T, \Omega} = \left(\frac{\partial \hat{\epsilon}}{\partial n} \right)_{T, \Omega} \Rightarrow \frac{1}{\kappa} = n^2 \left(\frac{\partial \mu}{\partial n} \right)_{T, \Omega}. \quad (4.27)$$

A change of the volume via applying pressure yields a change of the particle density n for fixed particle number N . On the other hand, we may keep the volume constant and change the particle number by altering the chemical potential, which then also leads to a change of n . In momentum space increasing (decreasing) n means enlarging (decreasing) the volume of the Fermi sphere. The relation between n and k_F is given by,

$$n = \frac{N}{\Omega} = \frac{k_F^3}{3\pi^2} \Rightarrow \delta n = \frac{k_F^2}{\pi^2} \delta k_F. \quad (4.28)$$

This yields a deviation of the distribution function, $\delta n_\sigma(\mathbf{k})$ which is isotropic (see Fig. 4.5 , left panel).

The chemical potential also depends on k_F and the distribution function, $\mu = \tilde{\epsilon}(\mathbf{k}_F, n_\sigma(\mathbf{k}))$.

$$\delta\mu = v_F \hbar \delta k_F + \frac{\delta \tilde{\epsilon}(\mathbf{k}_F)}{\delta n_{\sigma'}(\mathbf{k}')} \frac{\delta n_{\sigma'}(\mathbf{k}')}{\delta k_F} \delta k_F = v_F \hbar \delta k_F + \frac{1}{\Omega} \sum_{\mathbf{k}', \sigma'} f_{\sigma, \sigma'}(\mathbf{k}_F, \mathbf{k}') \delta(k' - k_F) \delta k_F \quad (4.29)$$

so that

$$\frac{\partial \mu}{\partial k_F} = \hbar v_F + \frac{4\pi k_F^2}{(2\pi)^3} \int \frac{d\Omega_{\hat{k}'}}{4\pi} \sum_{\sigma'} f_{\sigma \sigma'}(\hat{k}, \hat{k}') = \frac{\hbar^2 k_F}{m^*} \left\{ 1 + N(\epsilon_F) \int \frac{d\Omega_{\hat{k}'}}{4\pi} f^s(\hat{k}, \hat{k}') \right\} \quad (4.30)$$

Finally, using (4.18) and (4.27), we find

$$\frac{1}{\kappa} = n^2 \frac{\partial \mu}{\partial k_F} \frac{\partial k_F}{\partial n} = \frac{n^2}{N(\epsilon_F)} \{1 + F_0^s\} = \frac{2}{3} n \epsilon_F \{1 + F_0^s\}. \quad (4.31)$$

The Landau parameter F_0^s as well as the effective mass m^* included in $N(\epsilon_F)$ describe the renormalization of the compressibility.

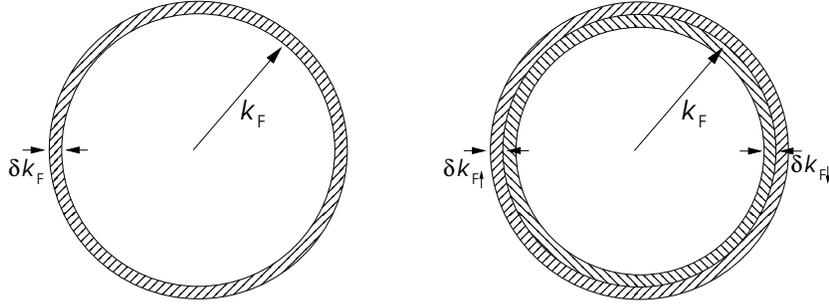


Fig. 4.5: Deviations of the distribution functions: Left panel: isotropic increase of the Fermi surface as used for the uniform compressibility; right panel: spin dependent change of size of the Fermi surface as used for the uniform spin susceptibility.

4.2.3 Spin susceptibility

If we apply a magnetic field, the Zeeman coupling yields a shift of the quasiparticle energies depending on their spin orientation. For an isotropic system, without loss of generality, we can assume the magnetic field to be directed into the z -direction. The energy shift is

$$\delta \tilde{\epsilon}_\sigma(\mathbf{k}) = -g\mu_B H \frac{\sigma}{2} + \frac{1}{\Omega} \sum_{\mathbf{k}', \sigma'} f_{\sigma, \sigma'}(\mathbf{k}, \mathbf{k}') \delta n_{\sigma'}(\mathbf{k}') = -\tilde{g}\mu_B H \frac{\sigma}{2}. \quad (4.32)$$

The gyromagnetic factor \tilde{g} differs from the value of $g = 2$ for free electrons due to interactions,

$$\begin{aligned} \frac{1}{\Omega} \sum_{\mathbf{k}', \sigma'} f_{\sigma \sigma'}(\mathbf{k}, \mathbf{k}') \delta n_{\sigma'}(\mathbf{k}') &= \frac{1}{\Omega} \sum_{\mathbf{k}', \sigma'} f_{\sigma \sigma'}(\mathbf{k}, \mathbf{k}') \frac{\partial n_{\sigma'}(\mathbf{k}')}{\partial \tilde{\epsilon}_{\sigma'}(\mathbf{k}')} \delta \tilde{\epsilon}_{\sigma'}(\mathbf{k}') \\ &= \frac{1}{\Omega} \sum_{\mathbf{k}', \sigma'} f_{\sigma \sigma'}(\mathbf{k}, \mathbf{k}') \delta(\tilde{\epsilon}_{\sigma'}(\mathbf{k}') - \epsilon_F) \tilde{g}\mu_B H \frac{\sigma'}{2} \end{aligned} \quad (4.33)$$

From this and (4.32) we derive

$$\tilde{g} = g - \tilde{g} N(\epsilon_F) \int \frac{d\Omega_{\hat{k}'}}{4\pi} f^a(\hat{k}, \hat{k}') = g - \tilde{g} F_0^a \quad \Rightarrow \quad \tilde{g} = \frac{g}{1 + F_0^a}. \quad (4.34)$$

The magnetization can be computed from the distribution function,

$$M = g\mu_B \sum_{\mathbf{k}, \sigma} \frac{\sigma}{2} \delta n_\sigma(\mathbf{k}) = g\mu_B \sum_{\mathbf{k}, \sigma} \frac{\sigma}{2} \frac{\partial n_\sigma(\mathbf{k})}{\partial \tilde{\epsilon}_\sigma(\mathbf{k})} \delta \tilde{\epsilon}_\sigma(\mathbf{k}) = g\mu_B \sum_{\mathbf{k}, \sigma} \frac{\sigma}{2} \delta(\tilde{\epsilon}_\sigma(\mathbf{k}) - \epsilon_F) \tilde{g}\mu_B H \frac{\sigma}{2} \quad (4.35)$$

and the susceptibility is given by

$$\chi = \frac{M}{H\Omega} = \frac{\mu_B^2 N(\epsilon_F)}{1 + F_0^a} \quad (4.36)$$

The changes in the distribution function induced by the magnetic field feed back into the susceptibility, so that the latter may be either weakened or enhanced.

4.2.4 Effective mass and Galilei invariance

We have initially introduced the effective mass of quasiparticles in $\epsilon_\sigma(\mathbf{k})$ by hand. In this section we would like to show that overall consistency requires a relation of the effective mass with a certain Landau parameter. The effective mass is the result of the interactions among the electrons.

This consistency is connected with the Galilean invariance when we shift the momenta of all particles by $\hbar\mathbf{q}$ which shall be very small compared to the Fermi momentum in order to be compatible with the assumption of the Fermi liquid theory, i.e. the distribution function given by

$$\delta n_\sigma(\mathbf{k}) = n_\sigma^{(0)}(\mathbf{k} + \mathbf{q}) - n_\sigma^{(0)}(\mathbf{k}) \approx \mathbf{q} \cdot \nabla n_\sigma^{(0)}(\mathbf{k}) . \quad (4.37)$$

is concentrated around the Fermi energy (see Fig.4.6).

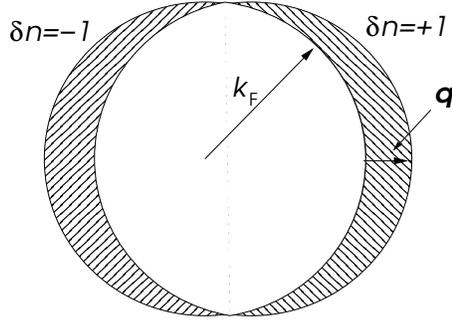


Fig. 4.6: Distribution function due to a Fermi surface shift (Galilei transformation).

We calculate now the current density using the distribution function $n_\sigma(\mathbf{k}) = n_\sigma^{(0)}(\mathbf{k}) + \delta n_\sigma(\mathbf{k})$. Within the Fermi liquid theory this yields,

$$\mathbf{j}_\mathbf{q} = \frac{1}{\Omega} \sum_{\mathbf{k}, \sigma} \mathbf{v}(\mathbf{k}) n_\sigma(\mathbf{k}) \quad (4.38)$$

with

$$\mathbf{v}(\mathbf{k}) = \frac{1}{\hbar} \nabla_{\mathbf{k}} \tilde{\epsilon}_\sigma(\mathbf{k}) = \frac{1}{\hbar} \left\{ \nabla_{\mathbf{k}} \epsilon_\sigma(\mathbf{k}) + \frac{1}{\Omega} \sum_{\mathbf{k}', \sigma'} \nabla_{\mathbf{k}} f_{\sigma\sigma'}(\mathbf{k}, \mathbf{k}') \delta n_\sigma(\mathbf{k}') \right\} . \quad (4.39)$$

Thus we obtain for the current density,

$$\begin{aligned} \mathbf{j}_\mathbf{q} &= \frac{1}{\Omega} \sum_{\mathbf{k}, \sigma} \frac{\hbar\mathbf{k}}{m^*} n_\sigma(\mathbf{k}) + \frac{1}{\Omega^2} \sum_{\mathbf{k}, \sigma} \sum_{\mathbf{k}', \sigma'} \{n_\sigma^{(0)}(\mathbf{k}) + \delta n_\sigma(\mathbf{k})\} \frac{1}{\hbar} \nabla_{\mathbf{k}} f_{\sigma\sigma'}(\mathbf{k}, \mathbf{k}') \delta n_\sigma(\mathbf{k}') \\ &= \frac{1}{\Omega} \sum_{\mathbf{k}, \sigma} \frac{\hbar\mathbf{k}}{m^*} \delta n_\sigma(\mathbf{k}) - \frac{1}{\Omega^2} \sum_{\mathbf{k}, \sigma} \sum_{\mathbf{k}', \sigma'} \frac{1}{\hbar} \{ \nabla_{\mathbf{k}} n_\sigma^{(0)}(\mathbf{k}) \} f_{\sigma\sigma'}(\mathbf{k}, \mathbf{k}') \delta n_\sigma(\mathbf{k}') + O(q^2) \\ &= \frac{1}{\Omega} \sum_{\mathbf{k}, \sigma} \frac{\hbar\mathbf{k}}{m^*} \delta n_\sigma(\mathbf{k}) + \frac{1}{\Omega^2} \sum_{\mathbf{k}, \sigma} \sum_{\mathbf{k}', \sigma'} f_{\sigma\sigma'}(\mathbf{k}, \mathbf{k}') \frac{\hbar\mathbf{k}'}{m^*} \delta n_\sigma(\mathbf{k}') + O(q^2) = \mathbf{j}_1 + \mathbf{j}_2 . \end{aligned} \quad (4.40)$$

where, for the second line, we performed an integration by parts and neglect terms quadratic in δn and, in the third line, use $f_{\sigma\sigma'}(\mathbf{k}, \mathbf{k}') = f_{\sigma'\sigma}(\mathbf{k}', \mathbf{k})$ and

$$\nabla_{\mathbf{k}} n_{\sigma}^{(0)}(\mathbf{k}) = \frac{\partial n_{\sigma}^{(0)}(\mathbf{k})}{\partial \epsilon_{\sigma}(\mathbf{k})} \nabla_{\mathbf{k}} \epsilon_{\sigma}(\mathbf{k}) = -\delta(\epsilon_{\sigma}(\mathbf{k}) - \epsilon_F) \nabla_{\mathbf{k}} \epsilon_{\sigma}(\mathbf{k}) = -\delta(\epsilon_{\sigma}(\mathbf{k}) - \epsilon_F) \frac{\hbar \mathbf{k}}{m^*}. \quad (4.41)$$

The first term of Eq.(4.40) denotes quasiparticle current, \mathbf{j}_1 , while the second term can be interpreted as a drag current, the induced motion of the other particles due to interaction (backflow), \mathbf{j}_2 .

Now we turn to a different viewpoint, looking at the system as being in the inertial frame with a velocity $\hbar \mathbf{q}/m$, as all particles received the same momentum. The current density then is given by

$$\mathbf{j}_{\mathbf{q}} = \frac{N \hbar \mathbf{q}}{\Omega m} = \frac{1}{\Omega} \sum_{\mathbf{k}, \sigma} \frac{\hbar \mathbf{k}}{m} n_{\sigma}(\mathbf{k}) = \frac{1}{\Omega} \sum_{\mathbf{k}, \sigma} \frac{\hbar \mathbf{k}}{m} \delta n_{\sigma}(\mathbf{k}). \quad (4.42)$$

Since these two viewpoints are equivalent the resulting currents should be the same. Thus, we compare Eq.(4.40) and (4.42) and obtain the equation,

$$\frac{\hbar \mathbf{k}}{m} = \frac{\hbar \mathbf{k}}{m^*} + \frac{1}{\Omega} \sum_{\mathbf{k}', \sigma'} f_{\sigma\sigma'}(\mathbf{k}, \mathbf{k}') \delta(\epsilon_{\sigma}(\mathbf{k}') - \epsilon_F) \frac{\hbar \mathbf{k}'}{m^*} \quad (4.43)$$

which then leads to

$$\frac{1}{m} = \frac{1}{m^*} + N(\epsilon_F) \int \frac{d\Omega_{\hat{k}'}}{4\pi} f^s(\hat{k}, \hat{k}') \frac{\hat{k} \cdot \hat{k}'}{m^*} \quad (4.44)$$

$$\Rightarrow \frac{m^*}{m} = 1 + \frac{1}{3} F_1^s.$$

The factor 1/3 originates in the term $1/(2l+1)$ for $l=1$ as

$$\int_{-1}^{+1} dz P_l(z) P_{l'}(z) = \frac{2\delta_{ll'}}{2l+1} \quad (4.45)$$

Thus, for consistency the relation of Eq.(4.44) between m^* and F_1^s must hold. Generally, we find that $F_1^s > 0$ so that quasiparticles in a Fermi liquid are "heavier" than bare electrons.

4.2.5 Stability of the Fermi liquid

Upon inspection of the renormalization of the quantities

$$\frac{\kappa}{\kappa_0} = \frac{m^*}{m} \frac{1}{1+F_0^s}, \quad \frac{\chi}{\chi_0} = \frac{m^*}{m} \frac{1}{1+F_0^a} \quad \text{mit} \quad \frac{m^*}{m} = 1 + \frac{1}{3} F_1^s \quad (4.46)$$

one notes that for example the compressibility κ and the susceptibility χ diverge for $F_0^s \rightarrow -1$ or $F_0^a \rightarrow -1$, indicating an instability of the system. A diverging spin susceptibility leads to a ferromagnetic state with a split Fermi surface, one for each spin direction. A diverging compressibility leads to a spontaneous contraction of the system.

More generally, the deformation of the quasiparticle distribution function may vary over the Fermi surface, so that arbitrary deviations of the simple Fermi liquid state may be classified by this deformation,

$$\delta n_{\sigma}(\hat{k}) = \sum_{l=0}^{\infty} \delta n_{\sigma,l} P_l(\cos \theta_{\hat{k}}) \quad (4.47)$$

For charge density deformations $\delta n_{+l}(\hat{k}) = \delta n_{-l}(\hat{k})$ holds, whereas spin density deformations are described by $\delta n_{+l}(\hat{k}) = -\delta n_{-l}(\hat{k})$. Stability of the Fermi liquid against any of these deformations requires

$$1 + \frac{F_l^{s,a}}{2l+1} > 0. \quad (4.48)$$

In general, the renormalization of the Fermi liquid leads to a change in the Wilson ratio

$$\frac{R}{R_0} = \frac{\chi}{\chi_0} \frac{\gamma_0}{\gamma} = \frac{1}{1 + F_0^a} \quad (4.49)$$

where $R_0 = \chi_0/\gamma_0 = 6\mu_B^2/\pi^2 k_B^2$. Note that the Wilson ratio does not depend on the effective mass.

A remarkable feature of the Fermi liquid theory is that even very strongly interacting Fermions remain Fermi liquids, notably the quantum liquid ^3He and so-called heavy Fermion systems, which are compounds of transition metals and rare earths. Both are strongly renormalized Fermi liquids. For ^3He we give a some of the parameters here for zero pressure and for pressures just below the critical pressure at which He solidifies ($p_c \approx 2.5 \text{ MPa} = 25 \text{ bar}$):

pressure	m^*/m	F_0^s	F_0^a	F_1^s	κ/κ_0	χ/χ_0
0	3.0	10.1	-0.52	6.0	0.27	6.3
$< p_c$	6.2	94	-0.74	15.7	0.065	24

The trends are obvious here. The higher the pressure the denser the liquid and the stronger the interaction are. Approaching the solidification the compressibility is reduced, the quasiparticles become heavier (slower) and the magnetic response increases drastically.

Finally the heavy Fermion systems are characterized by the extraordinary enhancements of the effective mass which for many of these compounds lie between 100 and 1000 times the electron mass (e.g. CeAl_3 , UBe_{13} etc.). This large mass leads also the notion of an "almost localized Fermi liquid", as the large effective mass is generated by the hybridization of itinerant conduction electrons with strongly interacting (localized) electron states in partially filled $4f$ - or $5f$ -orbitals of Lanthanide and Actinide atoms, respectively.

4.3 Microscopic considerations

A rigorous derivation of the Landau Fermi-liquid theory requires methods of quantum field theory and is beyond the scope of these lectures. However, plain Rayleigh-Schrödinger theory allows to gain some insights into the microscopic fundament of Landau's Fermi liquid theory. In the following, we consider a model of fermions with contact interaction described by the Hamiltonian,

$$\begin{aligned} \mathcal{H} &= \sum_{\mathbf{k},s} \epsilon_{\mathbf{k}} \hat{c}_{\mathbf{k}s}^\dagger \hat{c}_{\mathbf{k}s} + \int d^3r d^3r' \hat{\Psi}_\uparrow(\mathbf{r})^\dagger \hat{\Psi}_\downarrow(\mathbf{r}')^\dagger U \delta(\mathbf{r} - \mathbf{r}') \hat{\Psi}_\downarrow(\mathbf{r}') \hat{\Psi}_\uparrow(\mathbf{r}) \\ &= \sum_{\mathbf{k},s} \epsilon_{\mathbf{k}} \hat{c}_{\mathbf{k}s}^\dagger \hat{c}_{\mathbf{k}s} + \frac{U}{\Omega} \sum_{\mathbf{k},\mathbf{k}',\mathbf{q}} \hat{c}_{\mathbf{k}+\mathbf{q}\uparrow}^\dagger \hat{c}_{\mathbf{k}'-\mathbf{q}\downarrow}^\dagger \hat{c}_{\mathbf{k}'\downarrow} \hat{c}_{\mathbf{k}\uparrow} \end{aligned} \quad (4.50)$$

wobei $\epsilon_{\mathbf{k}} = \hbar^2 \mathbf{k}^2 / 2m$. Note that we had previously found that in order to find well-defined quasiparticles the interaction between the Fermions has to be short ranged.

4.3.1 Landau parameters

For a given momentum distribution $n_{\mathbf{k}s} = \langle c_{\mathbf{k}s}^\dagger c_{\mathbf{k}s} \rangle = n_{\mathbf{k}s}^{(0)} + \delta n_{\mathbf{k}s}$ we can expand the corresponding energy following the Rayleigh-Schrödinger perturbation method,

$$E = E^{(0)} + E^{(1)} + E^{(2)} + \dots \quad \text{with}$$

$$E^{(0)} = \sum_{\mathbf{k},s} \epsilon_{\mathbf{k}} n_{\mathbf{k}s}, \quad E^{(1)} = \frac{U}{\Omega} \sum_{\mathbf{k},\mathbf{k}'} n_{\mathbf{k}\uparrow} n_{\mathbf{k}'\downarrow}, \quad (4.51)$$

$$E^{(2)} = \frac{U^2}{\Omega^2} \sum_{\mathbf{k},\mathbf{k}',\mathbf{q}} \frac{n_{\mathbf{k}\uparrow} n_{\mathbf{k}'\downarrow} (1 - n_{\mathbf{k}+\mathbf{q}\uparrow}) (1 - n_{\mathbf{k}'-\mathbf{q}\downarrow})}{\epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}'-\mathbf{q}}}$$

The second order term describes virtual processes corresponding to a pair of particle-hole excitations. The numerator of the term $E^{(2)}$ can be split into four different contributions. We first consider the term quadratic in $n_{\mathbf{k}}$ and combine it with the first order term which has the same structure,

$$\tilde{E}^{(1)} = E^{(1)} + \frac{U^2}{\Omega^2} \sum_{\mathbf{k},\mathbf{k}',\mathbf{q}} \frac{n_{\mathbf{k}\uparrow} n_{\mathbf{k}'\downarrow}}{\epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}'-\mathbf{q}}} \approx \frac{\tilde{U}}{\Omega} \sum_{\mathbf{k},\mathbf{k}'} n_{\mathbf{k}\uparrow} n_{\mathbf{k}'\downarrow}. \quad (4.52)$$

Here we have defined a renormalized interaction $U \rightarrow \tilde{U}$ through,

$$\tilde{U} = U + \frac{U^2}{\Omega} \sum_{\mathbf{q}} \frac{1}{\epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}'-\mathbf{q}}} \quad (4.53)$$

In principle, \tilde{U} depends on the wave vectors \mathbf{k} and \mathbf{k}' . When restricted to the Fermi surface ($|\mathbf{k}| = |\mathbf{k}'| = k_F$), however, this dependency may be neglected, if the range of the interaction ℓ is small compared to the mean electron spacing, $k_F \ell \ll 1$.²

The next higher term is cubic in $n_{\mathbf{k}}$:

$$\tilde{E}^{(2)} = -\frac{\tilde{U}^2}{\Omega^2} \sum_{\mathbf{k},\mathbf{k}',\mathbf{q}} \frac{n_{\mathbf{k}\uparrow} n_{\mathbf{k}'\downarrow} (n_{\mathbf{k}+\mathbf{q}\uparrow} + n_{\mathbf{k}'-\mathbf{q}\downarrow})}{\epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}'-\mathbf{q}}} \quad (4.55)$$

We have replaced U^2 by \tilde{U}^2 , which is admissible at this order. The term quartic in $n_{\mathbf{k}}$ vanishes due to symmetry.

We can vary the energy E with respect to $\delta n_{\mathbf{k}\uparrow}$ and obtain

$$\tilde{\epsilon}_\uparrow(\mathbf{k}) = \epsilon_{\mathbf{k}} + \frac{\tilde{U}}{\Omega} \sum_{\mathbf{k}'} n_{\mathbf{k}'\downarrow} - \frac{\tilde{U}^2}{\Omega^2} \sum_{\mathbf{k}',\mathbf{q}} \frac{n_{\mathbf{k}'\downarrow} (n_{\mathbf{k}+\mathbf{q}\uparrow} + n_{\mathbf{k}'-\mathbf{q}\downarrow}) - n_{\mathbf{k}+\mathbf{q}\uparrow} n_{\mathbf{k}'-\mathbf{q}\downarrow}}{\epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}'-\mathbf{q}}} \quad (4.56)$$

and analogously for $\epsilon_\downarrow(\mathbf{k})$. The coupling parameters may be determined using the definition (4.15). Beginning with $f_{\uparrow\uparrow}(\mathbf{k}_F, \mathbf{k}'_F)$ with wavevectors on the Fermi surface ($\mathbf{k} \rightarrow \mathbf{k}_F$), the terms

²Care has to be taken for our contact interaction which would lead to a divergence in the large- q range. A cutoff for q of order $Q_c \sim \ell^{-1}$ would regularize the integral which is dominated by the large- q part. Thus we may use the following expansion,

$$I = \frac{1}{\Omega} \sum_{\mathbf{q}} \frac{1}{\epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}'-\mathbf{q}}} = \frac{1}{(2\pi)^3} \int_0^{Q_c} dq q^2 \int d\Omega_q \frac{m}{(\mathbf{k}' - \mathbf{k}) \cdot \mathbf{q} - q^2} = \frac{m}{(2\pi)^2} \int dq q \int_{-1}^{+1} \frac{d \cos \theta}{K \cos \theta - q}$$

$$= \frac{m}{(2\pi)^2} \int_0^{Q_c} dq q \ln \left| \frac{q - K}{q + K} \right| = -\frac{m}{(2\pi)^2} \left\{ Q_c + \frac{K^2 - Q_c^2}{2K} \ln \left| \frac{Q_c - K}{Q_c + K} \right| \right\} \approx -\frac{2mQ_c}{(2\pi)^2} \left\{ 1 - \frac{K^2}{Q_c^2} + O\left(\frac{K^4}{Q_c^4}\right) \right\} \quad (4.54)$$

where we use $K = |\mathbf{k}' - \mathbf{k}| \leq 2k_F \ll Q_c$. Thus, the momentum dependence of \tilde{U} is indeed weak.

contributing to the coupling can be written as

$$\begin{aligned} & \frac{\tilde{U}^2}{\Omega^2} \sum_{\mathbf{k}', \mathbf{q}} n_{\mathbf{k}+\mathbf{q}\uparrow} \frac{n_{\mathbf{k}'-\mathbf{q}\downarrow} - n_{\mathbf{k}'\downarrow}}{\epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}'-\mathbf{q}}} \\ & \xrightarrow{\mathbf{k}+\mathbf{q} \rightarrow \mathbf{k}'_F} \frac{1}{\Omega} \sum_{\mathbf{k}'_F} n_{\mathbf{k}'_F\uparrow} \frac{\tilde{U}^2}{\Omega} \sum_{\mathbf{k}'} \frac{n_{\mathbf{k}'-\mathbf{q}\downarrow}^{(0)} - n_{\mathbf{k}'\downarrow}^{(0)}}{\epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}'-\mathbf{q}}} \Bigg|_{\mathbf{q}=\mathbf{k}'_F-\mathbf{k}_F} = -\frac{1}{\Omega} \sum_{\mathbf{k}'_F} n_{\mathbf{k}'_F\uparrow} \frac{\tilde{U}^2}{2} \chi_0(\mathbf{k}'_F - \mathbf{k}_F) \end{aligned} \quad (4.57)$$

where we consider $n_{\mathbf{k}'_F\uparrow} = n_{\mathbf{k}'_F\uparrow}^{(0)} + \delta n_{\mathbf{k}'_F\uparrow}$. Note that the part of $n_{\mathbf{k}'_F\uparrow}^{(0)}$ in this term will contribute the ground state energy in Landau's energy functional. Here, $\chi_0(\mathbf{q})$ is the static susceptibility in (3.34) (Lindhard function). It follows that

$$f_{\uparrow\uparrow}(\mathbf{k}_F, \mathbf{k}'_F) = f_{\downarrow\downarrow}(\mathbf{k}_F, \mathbf{k}'_F) = \frac{\tilde{U}^2}{2} \chi_0(\mathbf{k}_F - \mathbf{k}'_F) \quad (4.58)$$

The other couplings are obtained in a similar manner,

$$f_{\uparrow\downarrow}(\mathbf{k}_F, \mathbf{k}'_F) = f_{\downarrow\uparrow}(\mathbf{k}_F, \mathbf{k}'_F) = \tilde{U} - \frac{\tilde{U}^2}{2} \{2\tilde{\chi}_0(\mathbf{k}_F + \mathbf{k}'_F) - \chi_0(\mathbf{k}_F - \mathbf{k}'_F)\} \quad (4.59)$$

where the function $\tilde{\chi}_0(\vec{q})$ is defined as

$$\tilde{\chi}_0(\mathbf{q}) = \frac{1}{\Omega} \sum_{\mathbf{k}'} \frac{n_{\mathbf{k}'+\mathbf{q}\uparrow}^{(0)} + n_{\mathbf{k}'\downarrow}^{(0)}}{2\epsilon_F - \epsilon_{\mathbf{k}'+\mathbf{q}} - \epsilon_{\mathbf{k}'}} \quad (4.60)$$

The couplings may be parametrized by the angle θ between \mathbf{k}_F and \mathbf{k}'_F .

$$\begin{aligned} f_{\sigma\sigma'}(\theta) = \frac{\tilde{U}}{2} \left[\left\{ 1 + \frac{\tilde{U}N(\epsilon_F)}{2} \left(2 + \frac{\cos\theta}{2\sin(\theta/2)} \ln \frac{1+\sin(\theta/2)}{1-\sin(\theta/2)} \right) \right\} \delta_{\sigma\sigma'} \right. \\ \left. - \left\{ 1 + \frac{\tilde{U}N(\epsilon_F)}{2} \left(1 - \frac{\sin(\theta/2)}{2} \ln \frac{1+\sin(\theta/2)}{1-\sin(\theta/2)} \right) \right\} \sigma\sigma' \right] \end{aligned} \quad (4.61)$$

Now we are in a position to determine the most important Landau parameters ($\tilde{u} = \tilde{U}N(\epsilon_F) > 0$):

$$\begin{aligned} F_0^s &= \tilde{u} \left\{ 1 + \tilde{u} \left(1 + \frac{1}{6}(2 + \ln 2) \right) \right\} = \tilde{u} + 1.449 \tilde{u}^2 > 0 \\ F_0^a &= -\tilde{u} \left\{ 1 + \tilde{u} \left(1 - \frac{2}{3}(1 - \ln 2) \right) \right\} = -\tilde{u} - 0.895 \tilde{u}^2 < 0 \\ F_1^s &= \tilde{u}^2 \frac{2}{15} (7\ln 2 - 1) = 0.514 \tilde{u}^2 > 0 \end{aligned} \quad (4.62)$$

In any case the effective mass m^* is enhanced as compared to the bare mass m , as the interaction between the particles enforces the motion of many particles whenever one is moved. The behavior of the susceptibility and the compressibility depends on the sign of the interaction. If the interaction is repulsive ($\tilde{u} > 0$), the compressibility decreases, implying that it is harder to compress the Fermi liquid. The susceptibility is enhanced in this case, so that it is easier to polarize the electrons' spins. Conversely, for attractive interactions the compressibility is enhanced whereas the susceptibility is suppressed. The attractive case, however, is more subtle because the Fermi liquid becomes unstable at low temperatures, turning into a superfluid or superconductor, by forming so-called Cooper pairs.

4.3.2 Distribution function

Finally, we examine the effect of interactions on the ground state properties again using Rayleigh-Schrödinger perturbation theory. The calculation of the corrections to the ground state $|\Psi_0\rangle$, the filled Fermi sea, is relatively simple:

$$|\Psi\rangle = |\Psi^{(0)}\rangle + |\Psi^{(1)}\rangle + \dots$$

$$|\Psi^{(0)}\rangle = |\Psi_0\rangle \quad \text{und} \quad |\Psi^{(1)}\rangle = \frac{U}{\Omega} \sum_{\mathbf{k}, \mathbf{k}', \mathbf{q}} \sum_{s, s'} \frac{\hat{c}_{\mathbf{k}+\mathbf{q}, s}^\dagger \hat{c}_{\mathbf{k}'-\mathbf{q}, s'}^\dagger \hat{c}_{\mathbf{k}', s'} \hat{c}_{\mathbf{k}, s}}{\epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}'-\mathbf{q}}} |\Psi_0\rangle \quad (4.63)$$

The lowest order correction involves again particle-hole excitations, depleting the Fermi sea by lifting particles virtually above the Fermi energy. This will affect the distribution function as we will discuss now. The momentum distribution $\langle \hat{n}_{\mathbf{k}s} \rangle = \langle \hat{c}_{\mathbf{k}s}^\dagger \hat{c}_{\mathbf{k}s} \rangle$ is obtain as the expectation value,

$$\langle \hat{n}_{\mathbf{k}s} \rangle = \frac{\langle \Psi | \hat{c}_{\mathbf{k}s}^\dagger \hat{c}_{\mathbf{k}s} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \langle \hat{n}_{\mathbf{k}s} \rangle^{(0)} + \langle \hat{n}_{\mathbf{k}s} \rangle^{(2)} + \dots \quad (4.64)$$

where $\langle \hat{n}_{\mathbf{k}s} \rangle^{(0)}$ is the unperturbed (step-function) distribution $\Theta(k_F - |\mathbf{k}|)$, and

$$\langle \hat{n}_{\mathbf{k}s} \rangle^{(2)} = \begin{cases} -\frac{U^2}{\Omega^2} \sum_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3} \frac{(1-n_{\mathbf{k}_1})(1-n_{\mathbf{k}_2})n_{\mathbf{k}_3}}{(\epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}_3} - \epsilon_{\mathbf{k}_1} - \epsilon_{\mathbf{k}_2})^2} \delta_{\mathbf{k}+\mathbf{k}_3, \mathbf{k}_1+\mathbf{k}_2} & |\mathbf{k}| < k_F \\ \frac{U^2}{\Omega^2} \sum_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3} \frac{n_{\mathbf{k}_1}n_{\mathbf{k}_2}(1-n_{\mathbf{k}_3})}{(\epsilon_{\mathbf{k}_1} + \epsilon_{\mathbf{k}_2} - \epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}_3})^2} \delta_{\mathbf{k}+\mathbf{k}_3, \mathbf{k}_1+\mathbf{k}_2} & |\mathbf{k}| > k_F \end{cases} \quad (4.65)$$

This yields the modification of the distribution functions as depicted in Fig.4.7. It allows us also to determine the size of the discontinuity of the distribution function at the Fermi surface:

$$\langle \hat{n}_{\mathbf{k}_{F-}} \rangle - \langle \hat{n}_{\mathbf{k}_{F+}} \rangle = 1 - \left(\frac{UN(\epsilon_F)}{2} \right)^2 \ln 2 \quad (4.66)$$

The jump of $\langle \hat{n}_{\mathbf{k}} \rangle$ is reduced independently of the sign of the interaction. This jump is also a measure for the weight of the quasiparticle state at the Fermi surface.

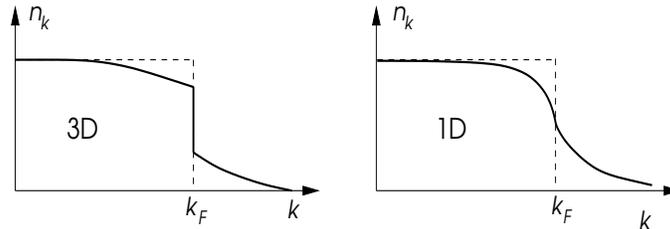


Fig. 4.7: Momentum distribution functions of electrons for a three-dimensional (left panel) and one-dimensional (right panel) Fermion system.

It is interesting to extend our considerations for an ordinary three dimensional Fermion liquid to the case of a one-dimensional system. The result is obtained straightforwardly,

$$\langle \hat{n}_{\mathbf{k}s} \rangle^{(2)} \approx \begin{cases} \frac{1}{8\pi^2} \frac{U^2}{\hbar^2 v_F^2} \ln \frac{k_+}{k - k_F} & k > k_F \\ -\frac{1}{8\pi^2} \frac{U^2}{\hbar^2 v_F^2} \ln \frac{k_-}{k_F - k} & k < k_F \end{cases} \quad (4.67)$$

where k_+ and k_- are cutoff parameters of the order of the Fermi wave vector k_F .

Apparently the quality of the perturbative calculation deteriorates as $k \rightarrow k_{F\pm}$, since we encounter a logarithmic divergence from both sides. Indeed, a more sophisticated approach shows that the distribution function is continuous at $k = k_F$ in one dimension, without any jump. Correspondingly, the quasiparticle weight vanishes and the elementary excitations cannot be described by Fermionic quasiparticles but rather by collective modes. This kind of behavior can be well described by the so-called Bosonization of Fermions in one dimension, a topic that is beyond the scope of these lectures. However, one of its surprising results is the fact that in one dimension, the fermionic excitations decay into independent charge and spin excitations (spin-charge separation). This behavior can be visualized by considering a half-filled lattice with predominantly antiferromagnetic spin correlations. In this case both charge and spin excitations are like different kinds of domain walls, moving at different velocities.

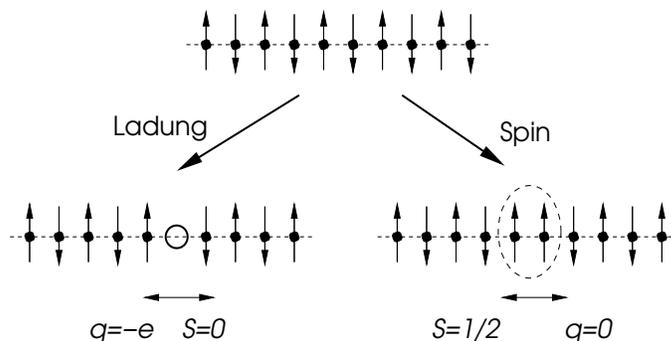


Fig. 4.8: Spin-charge separation: The dominant spin correlation is staggered (up-down-up-down-). A charge excitation is a vacancy which can move. A spin excitation may be considered as domain wall. Both excitations exist and move independently. A quasiparticle in a 3D Fermi liquid has always spin and charge together.

Chapter 5

Transport properties of metals

The ability to transport electrical current is one of the most remarkable and characteristic properties of metals. There are various influences determining their electrical conductivity: At zero temperature, a perfect metal is a perfect electrical conductor, i.e., the resistivity is zero. However, a *residual resistivity* due to defects in the material remains. At finite temperature, electron-electron and electron-phonon scattering can lead to a temperature-dependent resistivity. Furthermore, a magnetic field may influence the resistivity, a phenomenon called magnetoresistance, and lead to the previously mentioned Hall effect. We will, however, not consider the effects of a magnetic field in this chapter. Finally, we discuss heat transport which is also mostly mediated by electrons in metals. In this context other transport phenomena such as thermoelectricity (Seebeck and Peltier effect) will be of interest.

5.1 Electrical conductivity

In a normal metal electrical current density $\mathbf{j}(\mathbf{q}, \omega)$ is the result of an applied electrical field $\mathbf{E}(\mathbf{q}, \omega)$. We define the electrical conductivity $\sigma(\mathbf{q}, \omega)$ within a linear response through

$$\mathbf{j}(\mathbf{q}, \omega) = \sigma(\mathbf{q}, \omega) \mathbf{E}(\mathbf{q}, \omega) \quad (5.1)$$

for a homogeneous metal. It is interesting to see that a relation between the conductivity and the dynamical dielectric susceptibility $\chi_0(\mathbf{q}, \omega)$ follows from the continuity equation

$$\frac{\partial}{\partial t} \rho(\mathbf{r}, t) + \nabla \cdot \mathbf{j}(\mathbf{r}, t) = 0, \quad (5.2)$$

with ρ as the charge density, or, in Fourier space,

$$\omega \rho(\mathbf{q}, \omega) - \mathbf{q} \cdot \mathbf{j}(\mathbf{q}, \omega) = 0. \quad (5.3)$$

From this, we can calculate

$$\chi_0(\mathbf{q}, \omega) = -\frac{\rho(\mathbf{q}, \omega)}{eV_a(\mathbf{q}, \omega)} = -\frac{\mathbf{q} \cdot \mathbf{j}(\mathbf{q}, \omega)}{e\omega V_a(\mathbf{q}, \omega)} = -\frac{\sigma(\mathbf{q}, \omega) \mathbf{q} \cdot \mathbf{E}(\mathbf{q}, \omega)}{\omega e V_a(\mathbf{q}, \omega)} = \frac{\sigma(\mathbf{q}, \omega)}{\omega} \frac{\{-i\mathbf{q}^2 V_a(\mathbf{q}, \omega)\}}{e^2 V_a(\mathbf{q}, \omega)}, \quad (5.4)$$

and thus

$$\varepsilon(\mathbf{q}, \omega) = 1 - \frac{4\pi e^2}{q^2} \chi_0(\mathbf{q}, \omega) = 1 + \frac{4\pi i}{\omega} \sigma(\mathbf{q}, \omega). \quad (5.5)$$

In the limit $q \ll k_F$ of large wavelengths that $\varepsilon(0, \omega) = 1 - \omega_p^2/\omega^2$ as discussed previously¹, and thus

$$\sigma(\omega) = \frac{i\omega_p^2}{4\pi\omega}. \quad (5.6)$$

¹In the small- q limit we approximate $\chi_0(\mathbf{q}, \omega) \approx nq^2/m\omega^2$ from Eq.(3.45).

This might be interpreted in a way that the conductivity is purely imaginary. However, this conclusion is wrong, as we can find the real part of $\sigma(\omega)$ from the Kramers-Kronig relations. With $\sigma = \sigma_1 + i\sigma_2$, they state that

$$\sigma_1(\omega) = -\frac{1}{\pi} \int_{-\infty}^{+\infty} d\omega' \mathcal{P} \frac{1}{\omega - \omega'} \sigma_2(\omega') \quad \text{and} \quad \sigma_2(\omega) = \frac{1}{\pi} \int_{-\infty}^{+\infty} d\omega' \mathcal{P} \frac{1}{\omega - \omega'} \sigma_1(\omega'). \quad (5.7)$$

A simple calculation then yields

$$\sigma_1(\omega) = \frac{\omega_p^2}{4} \delta(\omega), \quad \sigma_2(\omega) = \frac{\omega_p^2}{4\pi\omega}. \quad (5.8)$$

Obviously this metal is a perfect conductor with $\sigma \rightarrow \infty$ for $\omega \rightarrow 0$. This comes about as we have only considered systems without dissipation.

An additional important property is the existence of the so-called f -sum rule,

$$\int_0^{\infty} d\omega' \sigma_1(\omega') = \frac{1}{2} \int_{-\infty}^{+\infty} d\omega' \sigma_1(\omega') = \frac{\omega_p^2}{8} = \frac{\pi e^2 n}{2m}. \quad (5.9)$$

It is valid for all electronic systems, in which charges are moved by external fields (even for semiconductors).

5.2 Transport equations and relaxation time

5.2.1 The Boltzmann equation

In order to tackle the problem of a finite conductivity, it is practicable to use a formalism analogous to Fermi liquid theory, on the basis of a distribution function of quasi particles. In transport theory, the distribution function describes the deviation from an equilibrium. If the system is isolated from the exterior, equilibrium is reached by relaxation after some time.

Let us introduce the distribution function $f(\mathbf{k}, \sigma; \mathbf{r}, t)$, where

$$f(\mathbf{k}, \sigma; \mathbf{r}, t) \frac{d^3 k}{(2\pi)^3} d^3 r \quad (5.10)$$

is the number of particles in the infinitesimal phase space volume $d^3 r d^3 k / (2\pi)^3$ at position (\mathbf{k}, \mathbf{r}) , time t , and with spin σ . Such a description is applicable only if temporal and spacial variations occur at long wavelength and small frequency, respectively, i.e., if typically $q \ll k_F$ and $\hbar\omega \ll \epsilon_F$. In the following, we neglect spin. Consequently, the total number of particles N is given by

$$N = \int \frac{d^3 k}{(2\pi)^3} d^3 r f(\mathbf{k}, \mathbf{r}, t). \quad (5.11)$$

We know, that the equilibrium distribution f_0 for the quasi particles is given by the Fermi-Dirac distribution, i.e.,

$$f_0(\mathbf{k}, \mathbf{r}, t) = \frac{1}{e^{(\epsilon_{\mathbf{k}} - \mu)/k_B T} + 1}, \quad (5.12)$$

and is independent of \mathbf{r} and t . Usually, we study processes close to equilibrium, where the deviation $f(\mathbf{k}, \mathbf{r}, t) - f_0(\mathbf{k}, \mathbf{r}, t)$ is small. The distribution function obeys the *Boltzmann equation*

$$\frac{D}{Dt} f(\mathbf{k}, \mathbf{r}, t) = \left\{ \frac{\partial}{\partial t} + \dot{\mathbf{r}} \cdot \nabla_{\mathbf{r}} + \dot{\mathbf{k}} \cdot \nabla_{\mathbf{k}} \right\} f(\mathbf{k}, \mathbf{r}, t) = \left(\frac{\partial f}{\partial t} \right)_{\text{coll}}, \quad (5.13)$$

where we have defined the substantial derivative in phase space D/Dt (the total temporal derivative in a frame moving with the phase-space volume). The right side is called collision

integral and describes the rate of change in f by collisions. Let us consider the temporal derivatives of \mathbf{r} and \mathbf{k} from a quasi-classical viewpoint. Then,

$$\dot{\mathbf{r}} = \mathbf{v} = \frac{\hbar\mathbf{k}}{m}, \quad (5.14)$$

$$\hbar\dot{\mathbf{k}} = -e(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \quad (5.15)$$

i.e., the force, which is our central interest, originates from the electric and magnetic fields. We will, however, only consider the case $\mathbf{B} = 0$.

In the collision integral, the probability $W(\mathbf{k}, \mathbf{k}')$ to scatter a quasi particle with wave vector \mathbf{k} to \mathbf{k}' is important. For simple scattering on static potentials, the collision integral is given by

$$\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = - \int \frac{d^3k'}{(2\pi)^3} [W(\mathbf{k}, \mathbf{k}')f(\mathbf{k}, \mathbf{r}, t)\{1 - f(\mathbf{k}', \mathbf{r}, t)\} - W(\mathbf{k}', \mathbf{k})f(\mathbf{k}', \mathbf{r}, t)\{1 - f(\mathbf{k}, \mathbf{r}, t)\}]. \quad (5.16)$$

The first term, describing the scattering from \mathbf{k} to \mathbf{k}' , requires a quasi particle at \mathbf{k} [hence the factor $f(\mathbf{k}, \mathbf{r}, t)$] and the absence of a particle at \mathbf{k}' [therefore the factor $1 - f(\mathbf{k}', \mathbf{r}, t)$]; note that spin is conserved here. This process describes the scattering out of the phase space volume under consideration, i.e., reduces the number of particles in it. Therefore, it enters the collision integral with negative sign. The second term describes the opposite process and thus enters with positive sign. For a system with time inversion symmetry, $W(\mathbf{k}, \mathbf{k}') = W(\mathbf{k}', \mathbf{k})$, i.e., the opposite scattering process happens with the same probability. Then, we can collect both terms and end up with

$$\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = \int \frac{d^3k'}{(2\pi)^3} W(\mathbf{k}, \mathbf{k}') \{f(\mathbf{k}', \mathbf{r}, t) - f(\mathbf{k}, \mathbf{r}, t)\}. \quad (5.17)$$

The Boltzmann equation is a relatively complicated integro-differential equation and suitable approximations are required. At small deviations from equilibrium, we can approximate the collision integral by the so-called *relaxation-time approximation*. For simplicity, we assume that the system is isotropic (quasiparticle dispersion $\epsilon_{\mathbf{k}}$ only depends on $|\mathbf{k}|$) and, furthermore, that the scattering probabilities depend only on the angle between \mathbf{k} and \mathbf{k}' and are elastic. Then, we make the Ansatz

$$\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = - \frac{f(\mathbf{k}, \mathbf{r}, t) - f_0(\mathbf{k}, \mathbf{r}, t)}{\tau(\epsilon_{\mathbf{k}})}. \quad (5.18)$$

Here, as in (5.12), $f_0(\mathbf{k}, \mathbf{r}, t)$ is a local equilibrium distribution with temperature $T = T(\mathbf{r}, t)$ and chemical potential $\mu = \mu(\mathbf{r}, t)$ generally depending weakly/slowly on \mathbf{r} and t . $\tau(\epsilon_{\mathbf{k}})$ is called *relaxation time* and is the characteristic time within which the system relaxes to equilibrium.

Consider a (small) uniform electric field $\mathbf{E}(t)$ and constant temperature. With $f(\mathbf{k}, \mathbf{r}, t) = f_0(\mathbf{k}, \mathbf{r}, t) + \delta f(\mathbf{k}, \mathbf{r}, t)$, we write the Fourier-transformed Boltzmann equation (5.13) in relaxation-time approximation and find, after linearizing in small δf ,

$$-i\omega\delta f(\mathbf{k}, \omega) + \frac{e\mathbf{E}(\omega)}{\hbar} \cdot \frac{\partial f_0(\mathbf{k})}{\partial \mathbf{k}} = - \frac{\delta f(\mathbf{k}, \omega)}{\tau(\epsilon_{\mathbf{k}})}, \quad (5.19)$$

where we have used that $f(\mathbf{k}, \mathbf{r}, t) = f(\mathbf{k}, t)$ for $E = E(t)$. In linearizing Eq. (5.19) in δf , we have actually assumed that $\delta f \propto |\mathbf{E}|$. This was indeed consistent, since Eq. (5.19) can easily be transformed into

$$\delta f(\mathbf{k}, \omega) = - \frac{e\tau\mathbf{E}(\omega)}{\hbar(1 - i\omega\tau)} \cdot \frac{\partial f_0(\mathbf{k})}{\partial \mathbf{k}} = - \frac{e\tau\mathbf{E}(\omega)}{\hbar(1 - i\omega\tau)} \cdot \frac{\partial f_0(\epsilon)}{\partial \epsilon} \frac{\partial \epsilon_{\mathbf{k}}}{\partial \mathbf{k}}. \quad (5.20)$$

From this, we can calculate the quasi-particle current $\mathbf{j}(\omega)$,

$$\mathbf{j}(\omega) = 2e \int \frac{d^3k}{(2\pi)^3} \mathbf{v}_{\mathbf{k}} f(\mathbf{k}, \omega) = - \frac{e^2}{4\pi^3} \int d^3k \frac{\tau(\epsilon_{\mathbf{k}})[\mathbf{E}(\omega) \cdot \mathbf{v}]\mathbf{v}}{1 - i\omega\tau(\epsilon_{\mathbf{k}})} \frac{\partial f_0(\epsilon_{\mathbf{k}})}{\partial \epsilon_{\mathbf{k}}}, \quad (5.21)$$

i.e.,

$$j_\alpha(\omega) = \sum_\beta \sigma_{\alpha\beta}(\omega) E_\beta(\omega) \quad (5.22)$$

with the conductivity tensor

$$\sigma_{\alpha\beta} = -\frac{e^2}{4\pi^3} \int d\epsilon \frac{\partial f_0(\epsilon)}{\partial \epsilon} \frac{\tau(\epsilon)}{1 - i\omega\tau(\epsilon)} \int d\Omega_{\mathbf{k}} k^2 \frac{v_{\alpha\mathbf{k}} v_{\beta\mathbf{k}}}{\hbar |\mathbf{v}_{\mathbf{k}}|}. \quad (5.23)$$

This corresponds to the Ohmic law. Note that $\sigma_{\alpha\beta}$ is diagonal in isotropic systems.

5.2.2 The Drude form

For $\omega\tau \gg 1$ Eq. (5.23) is independent on the relaxation time. In an isotropic system ($\sigma_{\alpha\beta} = \sigma\delta_{\alpha\beta}$) with $T \ll T_F$, this leads to

$$\sigma(\omega) \approx i \frac{e^2 m^2 v_F}{4\pi^3 \hbar^3 \omega} \int d\Omega_{\mathbf{k}} v_{Fz}^2 = i \frac{e^2 n}{m\omega} = i \frac{\omega_p^2}{4\pi\omega}, \quad (5.24)$$

which we have found already in Eq. (5.6). This does, however, not mean that we are dealing with a perfect conductor here.

We are actually interested in the static limit, i.e., in the dc conductivity ($\omega = 0$),

$$\sigma = -\frac{e^2 n}{m} \int d\epsilon \frac{\partial f_0}{\partial \epsilon} \tau(\epsilon) = \frac{e^2 n \bar{\tau}}{m} = \frac{\omega_p^2 \bar{\tau}}{4\pi}. \quad (5.25)$$

This form is the well-known Drude form of the conductivity.² If τ depends only weakly on energy, we can simply calculate the optical conductivity at finite frequency,

$$\sigma(\omega) = \frac{\omega_p^2}{4\pi} \frac{\bar{\tau}}{1 - i\omega\bar{\tau}} = \frac{\omega_p^2}{4\pi} \left\{ \frac{\bar{\tau}}{1 + \omega^2 \bar{\tau}^2} + \frac{i\bar{\tau}^2 \omega}{1 + \omega^2 \bar{\tau}^2} \right\} = \sigma_1 + i\sigma_2. \quad (5.26)$$

Note that the real part satisfies the f -sum rule,

$$\int_0^\infty d\omega \sigma_1(\omega) = \int_0^\infty d\omega \frac{\omega_p^2}{4\pi} \frac{\bar{\tau}}{1 + \omega^2 \bar{\tau}^2} = \frac{\omega_p^2}{8} \quad (5.27)$$

and $\sigma(\omega)$ in the limit $\bar{\tau} \rightarrow 0$ recovers the behavior of Eq.(5.8). This form of the conductivity yields the dielectric function

$$\varepsilon(\omega) = 1 - \frac{\omega_p^2 \bar{\tau}}{\omega(i + \omega\bar{\tau})} = 1 - \frac{\omega_p^2 \bar{\tau}^2}{1 + \omega^2 \bar{\tau}^2} + \frac{i}{\omega} \frac{1}{1 + \omega^2 \bar{\tau}^2}, \quad (5.28)$$

which can be used to discuss the optical properties of metals. The complex index of refraction N is given through $N^2 = (n + ik)^2 = \varepsilon$. We discuss three important regimes of frequency:

Relaxation-free regime ($\omega\bar{\tau} \ll 1 \ll \omega_p\bar{\tau}$): Here,

$$\varepsilon_1(\omega) \approx -\omega_p^2 \bar{\tau}^2, \quad (5.29)$$

$$\varepsilon_2(\omega) \approx \frac{\omega_p^2 \bar{\tau}}{\omega}. \quad (5.30)$$

The real part ε_1 is constant and negative, whereas the imaginary part ε_2 becomes singular for $\omega \rightarrow 0$. Thus, the refractive index turns out to be dominated by ε_2 :

$$n(\omega) \approx k(\omega) \approx \sqrt{\frac{\varepsilon_2(\omega)}{2}} \approx \sqrt{\frac{\omega_p^2 \bar{\tau}}{2\omega}}, \quad (5.31)$$

² Note, that the (phenomenological) Drude theory of electron transport can be deduced from purely classical considerations.

i.e., $n(\omega) \gg 1$. As a result, the reflectivity R is practically 100%, since

$$R = \frac{(n-1)^2 + k^2}{(n+1)^2 + k^2}. \quad (5.32)$$

The absorption index $k(\omega)$ determines the penetration depth δ through

$$\delta(\omega) = \frac{c}{\omega k(\omega)} \approx \frac{c}{\omega_p} \sqrt{\frac{2}{\omega \bar{\tau}}}. \quad (5.33)$$

This is the skin depth of a metal with the famous relation $\delta(\omega) \propto \omega^{-1/2}$. This length is much larger than the Debye length c/ω_p ($\sim 10^{-6} \text{cm} \sim 100 \text{\AA}$ für $\hbar\omega_p = 10 \text{ eV}$). δ is in the cm range for frequencies of $\sim 10 - 100 \text{Hz}$ (cf. Fig. 5.1 on a logarithmic scale to emphasize the behavior at small frequencies).

Relaxation regime ($1 \ll \omega \bar{\tau} \ll \omega_p \bar{\tau}$): We expand in $(\omega \bar{\tau})^{-1}$, yielding

$$\varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2} + i \frac{\omega_p^2}{\omega^3 \bar{\tau}}. \quad (5.34)$$

The real part $\varepsilon_1 \approx -\omega_p^2/\omega^2$ is still negative and ε_2 is still dominant. For the optical properties, we obtain

$$k(\omega) \approx \frac{\omega_p}{\omega} \quad (5.35)$$

$$n(\omega) \approx \frac{\omega_p}{2\omega^2 \bar{\tau}}. \quad (5.36)$$

$k(\omega) \gg n(\omega) \gg 1$ implies a large reflectivity of metals in this frequency range as well. Note that visible frequencies are part of this regime (see Fig. 5.2, 5.3). The frequency dependence of the penetration depth becomes weak, and its magnitude is approximately the Debye length, $\delta \sim c/\omega_p$.

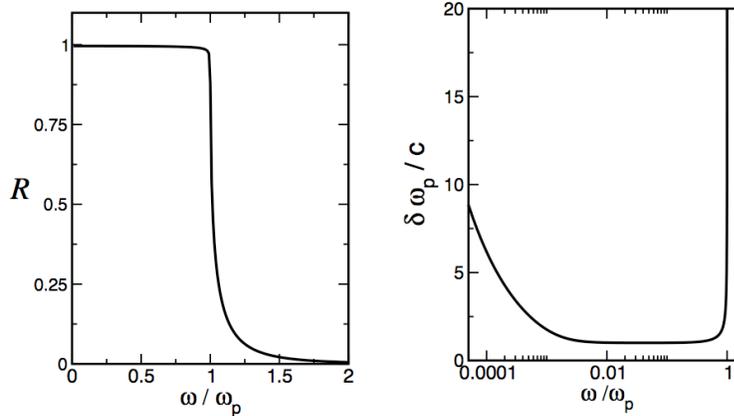


Figure 5.1: The frequency dependent reflectivity and penetration depth for $\omega_p \bar{\tau} = 500$.

Ultraviolet regime ($\omega \approx \omega_p$ and $\omega > \omega_p$): In this regime, the imaginary part of ε is approximately zero and the real part has the well known form

$$\varepsilon_1(\omega) = 1 - \frac{\omega_p^2}{\omega^2}, \quad (5.37)$$

such that the reflectivity changes drastically, from ~ 1 to ~ 0 (cf. Fig. 5.1). Metals are nearly transparent for $\omega > \omega_p$. In Fig. 5.1, one also notices the rapid increase in the penetration depth δ , showing the transparency of the metal.

In these considerations, we have neglected the contributions to the dielectric function due to the ion cores (core electrons and nuclei). They do, however, influence the reflecting properties of metals; particularly, the value of ω_p is lowered to $\omega'_p = \omega_p/\sqrt{\epsilon_\infty}$, where ϵ_∞ is the frequency-independent part of the dielectric function due to the ions. With this, the reflectivity for frequencies above ω'_p approaches $R = (\epsilon_\infty - 1)^2/(\epsilon_\infty + 1)^2$, and $0 < R < 1$ (see Fig. 5.2, 5.3).

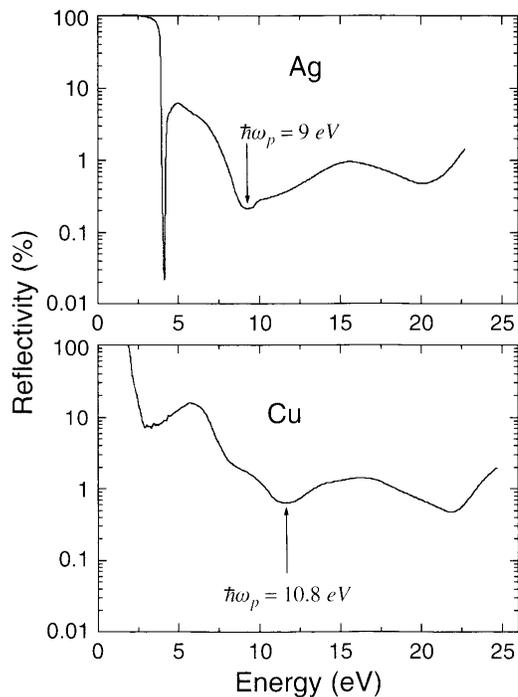


Figure 5.2: Reflectance spectra for silver and copper: In both cases the drop of reflectance is due to optical transitional between the completely filled d -band and the partially filled s -band. Note the logarithmic scale for the reflectivity. (Source: *An introduction to the optical spectroscopy of inorganic solids*, J. García Solé, L.E. Bausá and D. Jaque, Wiley (2005))

Color of metals: The practically 100% of reflectance for frequencies below ω_p is a typical feature of metals. Since for most metals the plasma frequency lies well above the range of visible light ($\hbar\omega = 1.5 - 3.5\text{eV}$), they appear shiny to our eye. While most polished metal surfaces appear shiny white, like silver, there are some metals with a color, like gold which is yellow and copper which is reddish. White shininess results from reflectance on the whole visible frequency range, while for colored metals there is a certain threshold above which the reflectance drops and frequencies towards blue are not or much weaker reflected. In most cases this drop is not connected with the plasma frequency, but with light absorption due to interband transitions. Note that the single band metal which was used for the Drude theory does not allow for optical absorption apart from the plasma excitation. Interband transition play, in particular, an important role in the case of the noble metals, Cu, Ag and Au. For these metals the reflectance drop is caused by the transition from the completely occupied d -band to the partially filled s -band (in case of Cu: $3d \rightarrow 4s$). For Cu this drop appear below 2.5 eV so that predominantly red light is reflected (see Fig.5.2). For Au this threshold frequency is slightly higher, but still in the visible, while for Ag it lies beyond the visible range (see Fig.5.2). For these cases the plasma frequency is not so easily recognizable in the reflectance. Note that the single band metal which was used for the

Drude theory does not allow for optical absorption apart from the plasma excitation. On the other hand, Al shows a reflectance rather close to the expected behavior (see. Fig.5.3). Also in this case there is some reduction of the reflection due to interband absorption. However, this effect is weak and the strong drop occurs at the plasma frequency of $\hbar\omega_p = 15.8\text{eV}$. Like Ag also polished Al is white shiny.

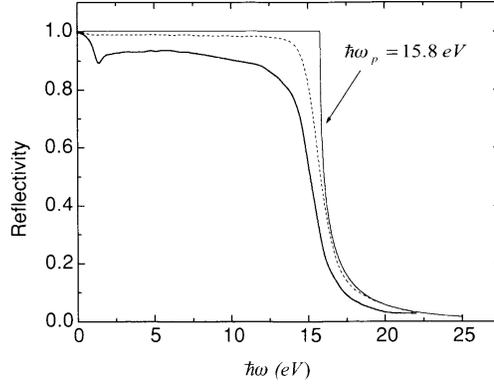


Figure 5.3: Reflectance spectrum of aluminium: The slight reduction of reflectivity below ω_p is due to interband transitions. The thin line is the theoretical behavior for $\tau = 0$ and the dashed line for finite τ . (Source: *An introduction to the optical spectroscopy of inorganic solids*, J. García Solé, L.E. Bausá and D. Jaque, Wiley (2005))

5.2.3 The relaxation time

With the help of the approximation Eq. (5.20) to the Boltzmann equation, we study the connection between the scattering rate $W(\mathbf{k}, \mathbf{k}')$ and the relaxation time τ . The relation

$$\frac{f(\mathbf{k}) - f_0(\mathbf{k})}{\tau(\epsilon_{\mathbf{k}})} = \int \frac{d^3k'}{(2\pi)^3} W(\mathbf{k}, \mathbf{k}') \{f(\mathbf{k}) - f(\mathbf{k}')\} \quad (5.38)$$

shall be valid, and we consider an isotropic system and assume elastic scattering. The solution of Eq. (5.19) is of the form

$$f(\mathbf{k}) = f_0(\mathbf{k}) + A(k)\mathbf{k} \cdot \mathbf{E} \quad \Rightarrow \quad f(\mathbf{k}) - f(\mathbf{k}') = A(k)(\mathbf{k} - \mathbf{k}') \cdot \mathbf{E}. \quad (5.39)$$

We introduce a parametrization of the angles according to $(\mathbf{k} \parallel z)$

$$\begin{aligned} \mathbf{k} \cdot \mathbf{E} &= kE \cos \theta, \\ \mathbf{k} \cdot \mathbf{k}' &= kk' \cos \theta', \\ \mathbf{k}' \cdot \mathbf{E} &= k'E(\cos \theta \cos \theta' + \sin \theta \sin \theta' \cos \phi'). \end{aligned} \quad (5.40)$$

With $k = k'$ (elastic scattering) we obtain

$$f(\mathbf{k}) - f(\mathbf{k}') = A(k)kE[\cos \theta(1 - \cos \theta') - \sin \theta \sin \theta' \cos \phi']. \quad (5.41)$$

After insertion in the right side of Eq. (5.38), the ϕ' -dependent part of the integration vanishes for an isotropic system, and

$$\begin{aligned} \int d\Omega_{\mathbf{k}'} [f(\mathbf{k}) - f(\mathbf{k}')] W(\mathbf{k}, \mathbf{k}') &= A(k)kE \cos \theta \int d\Omega_{\mathbf{k}'} (1 - \cos \theta') W(\mathbf{k}, \mathbf{k}') \\ &= [f(\mathbf{k}) - f_0(\mathbf{k})] \int d\Omega_{\mathbf{k}'} (1 - \cos \theta') W(\mathbf{k}, \mathbf{k}'), \end{aligned} \quad (5.42)$$

i.e., we can drop the factor $f(\mathbf{k}) - f_0(\mathbf{k})$ in both sides of Eq. (5.38) and obtain

$$\frac{1}{\tau(\epsilon_{\mathbf{k}})} = \int \frac{d^3 k'}{(2\pi)^3} W(\mathbf{k}, \mathbf{k}') (1 - \cos \theta'), \quad (5.43)$$

where one should remember that, for elastic scattering, $\epsilon_{\mathbf{k}} = \epsilon_{\mathbf{k}'}$.

5.3 Impurity scattering

5.3.1 Potential scattering

Every deviation from the perfect periodicity of the ionic lattice is a source of quasiparticle scattering, leading to the loss of their original momentum. Without translational invariance, the conservation of momentum is violated (energy, however, is still conserved). Possible static scatterers are, e.g., vacancies, dislocations, and impurity atoms. The scattering rate $W(\mathbf{k}, \mathbf{k}')$ for a potential \widehat{V} can be determined by the Golden Rule,³

$$W(\mathbf{k}, \mathbf{k}') = \frac{2\pi}{\hbar} n_{\text{imp}} |\langle \mathbf{k}' | \widehat{V} | \mathbf{k} \rangle|^2 \delta(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}). \quad (5.44)$$

By n_{imp} we denote the density of impurities, where we have assumed identical impurities. We neglect interference effects between different impurities, which is a reasonable approximation for small densities.

According to Eq. (5.43), the lifetime of a quasi particle with momentum $\hbar \mathbf{k}$ is given by

$$\begin{aligned} \frac{1}{\tau(\epsilon_{\mathbf{k}})} &= \frac{2\pi}{\hbar} n_{\text{imp}} \int \frac{d^3 k'}{(2\pi)^3} |\langle \mathbf{k}' | \widehat{V} | \mathbf{k} \rangle|^2 (1 - \hat{k} \cdot \hat{k}') \delta(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}) \\ &= n_{\text{imp}} (\hat{k} \cdot \mathbf{v}_{\mathbf{k}}) \int \frac{d\sigma}{d\Omega}(\mathbf{k}, \mathbf{k}') (1 - \hat{k} \cdot \hat{k}') \frac{d\Omega_{\mathbf{k}'}}{4\pi}, \end{aligned} \quad (5.45)$$

with the differential scattering cross section $d\sigma/d\Omega$ and $\hat{k} = \mathbf{k}/|\mathbf{k}|$. Here, we have used the connection between the Golden Rule and the Born approximation.⁴

Note the difference between the relaxation time τ and the lifetime $\tilde{\tau}$, where the Golden rule yields for the latter

$$\frac{1}{\tilde{\tau}} = \int \frac{d^3 k}{(2\pi)^3} W(\mathbf{k}, \mathbf{k}'). \quad (5.48)$$

The factor $1 - \cos \theta'$ in Eq. (5.43) emphasizes backscattering (scattering angle $\theta' \approx \pi$) in contrast to forward scattering (scattering angle $\theta' \approx 0$), as the latter has less influence on transport. Therefore, τ is called *transport lifetime*.

Assuming the defects to be point charges with charge Ze , whose potential is screened in a metal, leads to

$$\langle \mathbf{k}' | \widehat{V} | \mathbf{k} \rangle = \frac{4\pi Z e^2}{|\mathbf{k} - \mathbf{k}'|^2 + k_{\text{TF}}^2}. \quad (5.49)$$

³This corresponds to the first Born approximation in scattering theory. Note, that this approximation is insufficient to describe resonant scattering.

⁴The scattering of particles with momentum $\hbar \mathbf{k}$ into the solid angle $d\Omega_{\mathbf{k}'}$ around \mathbf{k}' yields

$$\begin{aligned} W(\mathbf{k}, \mathbf{k}') d\Omega_{\mathbf{k}'} &= \frac{2\pi}{\hbar} \sum_{\mathbf{k}' \in d\Omega_{\mathbf{k}'}} |\langle \mathbf{k}' | \widehat{V} | \mathbf{k} \rangle|^2 \delta(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}) \\ &= \frac{2\pi}{\hbar} d\Omega_{\mathbf{k}'} \int_{\mathbf{k}' \in d\Omega_{\mathbf{k}'}} \frac{d^3 k'}{(2\pi)^3} |\langle \mathbf{k}' | \widehat{V} | \mathbf{k} \rangle|^2 \delta(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}) = \frac{2\pi}{\hbar} d\Omega_{\mathbf{k}'} N(\epsilon) |\langle \mathbf{k}' | \widehat{V} | \mathbf{k} \rangle|^2. \end{aligned} \quad (5.46)$$

The scattering per incoming particle current $j_{\text{in}} d\sigma(\mathbf{k}, \mathbf{k}') = W(\mathbf{k}, \mathbf{k}') d\Omega_{\mathbf{k}'}$ determines the differential cross section

$$\hat{k} \cdot \mathbf{v}_{\mathbf{k}} \frac{d\sigma}{d\Omega}(\mathbf{k}, \mathbf{k}') = \frac{2\pi}{\hbar} \frac{N(\epsilon)}{4\pi} |\langle \mathbf{k}' | \widehat{V} | \mathbf{k} \rangle|^2. \quad (5.47)$$

leading to Eq. (5.45).

In the case of $k_{\text{TF}} \gg k_F$ (very strong screening), the differential cross section becomes independent of the deviation—in the context of partial wave expansion one speaks of s-wave scattering, i.e., $\delta_{l>0} \rightarrow 0$ —the transport and the usual lifetime become equal, $\tau = \tilde{\tau}$, where

$$\frac{1}{\tau} \approx \frac{\pi}{\hbar} N(\epsilon_F) n_{\text{imp}} \left(\frac{4\pi Z e^2}{k_{\text{TF}}^2} \right)^2. \quad (5.50)$$

We determine the conductivity for scattering on defects, assuming s-wave scattering only. Then, $\tau(\epsilon)$ depends weakly on energy and Eq. (5.25) yields

$$\sigma = \frac{e^2 n \tau(\epsilon_F)}{m}, \quad (5.51)$$

or, equivalently,

$$\rho = \frac{1}{\sigma} = \frac{m}{e^2 n \tau(\epsilon_F)}, \quad (5.52)$$

with the specific resistivity ρ . Both σ and ρ are independent of temperature. This contribution is called the *residual resistivity* of a metal, which approaches zero for a perfect material. The temperature dependence of the resistivity originates in other scattering processes like electron-phonon scattering and electron-electron scattering, which consider below. The so-called RRR (residual resistance ratio) is an often used figure to benchmark the quality of a material. It is defined as $= R(T = 300\text{K})/R(T = 0)$, i.e., as the ratio between the resistance at room temperature and the resistance at zero temperature. The bigger the RRR, the better the quality of the material.

5.3.2 Resonant scattering – Kondo effect

There are impurity atoms inducing so-called resonant scattering. If the resonance is close to the Fermi energy, the scattering rate is strongly energy dependent, inducing a more pronounced temperature dependence of the resistivity. An important example is the scattering off magnetic impurities with a spin degree of freedom, yielding a dramatic energy dependence of the scattering rate. This problem was first studied by Kondo in 1964, who used this to explain the peculiar minima in resistivity in some materials.

The coupling between the local spins \mathbf{S}_i at \mathbf{R}_i and the quasi-particle spin \mathbf{s} has the exchange form

$$\begin{aligned} \hat{V}_K &= \sum_i \hat{V}_{Ki} = J \sum_i \hat{\mathbf{S}}_i \cdot \hat{\mathbf{s}}(\mathbf{r}) \delta(\mathbf{r} - \mathbf{R}_i) \\ &= J \sum_i \left\{ \hat{S}_i^z \hat{s}^z(\mathbf{r}) + \frac{1}{2} \hat{S}_i^+ \hat{s}^-(\mathbf{r}) + \frac{1}{2} \hat{S}_i^- \hat{s}^+(\mathbf{r}) \right\} \delta(\mathbf{r} - \mathbf{R}_i) \\ &= \frac{J\hbar}{2\Omega} \sum_{\mathbf{k}, \mathbf{k}', i} \left[\hat{S}_i^z (\hat{c}_{\mathbf{k}\uparrow}^\dagger \hat{c}_{\mathbf{k}'\uparrow} - \hat{c}_{\mathbf{k}\downarrow}^\dagger \hat{c}_{\mathbf{k}'\downarrow}) + S_i^+ \hat{c}_{\mathbf{k}\downarrow}^\dagger \hat{c}_{\mathbf{k}'\uparrow} + S_i^- \hat{c}_{\mathbf{k}\uparrow}^\dagger \hat{c}_{\mathbf{k}'\downarrow} \right] e^{-i(\mathbf{k}-\mathbf{k}') \cdot \mathbf{R}_i}. \end{aligned} \quad (5.53)$$

It is important that spin flip processes, that change the spin state of the impurity and the scattered electron, are enabled.

We present here the resulting scattering rate without derivation,

$$W(\mathbf{k}, \mathbf{k}') \approx J^2 S(S+1) \left\{ 1 + 2JN(\epsilon_F) \ln \frac{D}{|\epsilon_{\mathbf{k}} - \epsilon_F|} + \dots \right\}, \quad (5.54)$$

where we have assumed that $JN(\epsilon_F) \ll 1$ (bandwidth D), and the relaxation time turns out to be

$$\frac{1}{\tau(\epsilon_{\mathbf{k}})} \approx \frac{J^2 S(S+1)}{\hbar} N(\epsilon) \left\{ 1 + 2JN(\epsilon_F) \ln \frac{D}{|\epsilon_{\mathbf{k}} - \epsilon_F|} \right\}. \quad (5.55)$$

Note that $W(\mathbf{k}, \mathbf{k}')$ does not depend on angle (s -wave scattering). The energy dependence is singular at the Fermi energy, indicating that we are not dealing with simple resonant potential scattering, but with a much more subtle many-particle effect involving the electrons on the Fermi surface. The fact that the local spin can be flipped makes the scattering center dynamical, i.e., the scatterer is constantly changing. The scattering process of an electron is influenced by previous scattering events, leading to the singularity at ϵ_F . This cannot be described within the first Born approximation, but requires at least the second approximation or the full solution. We refer to J.M. Ziman, *Principles of the Theory of Solids*, and A.C. Hewson, *The Kondo Problem to Heavy Fermions* for more details.

As indicated before, the resonant behavior induces a strong temperature dependence of the conductivity. Indeed,

$$\begin{aligned}\sigma(T) &= \frac{e^2 k_F^3}{6\pi^2 m} \int d\epsilon \frac{1}{4k_B T \cosh^2(\epsilon - \epsilon_F)/2k_B T} \tau(\epsilon) \\ &\approx \frac{e^2 n}{8m k_B T} \int d\tilde{\epsilon} \frac{J^2 S(S+1) \{1 - 2JN(\epsilon_F) \ln(D/\tilde{\epsilon})\}}{\cosh^2(\tilde{\epsilon}/2k_B T)}.\end{aligned}\quad (5.56)$$

A simple substitution in the integral leads to

$$\sigma(T) \approx \frac{e^2 n}{2m} J^2 S(S+1) \left\{ 1 - 2JN(\epsilon_F) \ln \frac{D}{k_B T} \right\}.\quad (5.57)$$

Usual contributions to the resistance (like electron-phonon scattering) typically decrease with temperature. The contribution due to Eq. (5.57) is strongly increasing, inducing a minimum in the resistance. At even lower temperatures, the conductivity seems to be suppressed (within our approximation it even becomes negative). This is, however, an artifact of the approximation. In reality, the conductivity saturates at a finite value, the crossover occurring at a characteristic Kondo temperature T_K ,

$$k_B T_K = D e^{-1/JN(\epsilon_F)},\quad (5.58)$$

a characteristic energy scale of this system. The real behavior at temperatures $T \ll T_K$ is not accessible by simple perturbation theory. It is known as the Kondo problem and constitutes one of the most interesting correlation effects of many-particle physics.

5.4 Electron-phonon interaction

Even in perfect metals, the conductivity remains finite at finite temperature. The thermally induced distortions of the lattice acts as fluctuating scattering centers. In the language of electron-phonon interaction, electrons are scattered by the absorption and emission of phonons, which induce local fluctuations in volume (cf. Chapter 3). The corresponding coupling term has the form

$$\begin{aligned}\widehat{V}_{\text{ep}} &= i \sum_{\mathbf{k}, \mathbf{q}, s} \tilde{V}_{\mathbf{q}} \mathbf{q} \cdot \left\{ \widehat{\mathbf{u}}_{\mathbf{q}} \widehat{c}_{\mathbf{k}+\mathbf{q}, s}^\dagger \widehat{c}_{\mathbf{k}, s} - \widehat{\mathbf{u}}_{-\mathbf{q}}^\dagger \widehat{c}_{\mathbf{k}, s}^\dagger \widehat{c}_{\mathbf{k}+\mathbf{q}, s} \right\} \\ &= 2i \sum_{\mathbf{k}, \mathbf{q}, s} \tilde{V}_{\mathbf{q}} \sqrt{\frac{\hbar}{2\rho_0 \omega_{\mathbf{q}}}} |\mathbf{q}| (\widehat{b}_{\mathbf{q}} - \widehat{b}_{-\mathbf{q}}^\dagger) \widehat{c}_{\mathbf{k}+\mathbf{q}, s}^\dagger \widehat{c}_{\mathbf{k}, s}.\end{aligned}\quad (5.59)$$

The interaction is similar to the interaction between electrons and electromagnetic radiation (photons). The most important processes consist of single-phonon processes, i.e., the absorption or emission of one phonon. Energy and momentum are conserved; for the scattering of an electron with momentum \mathbf{k} to \mathbf{k}' due to the emission of a phonon with momentum \mathbf{q} we have thus

$$\mathbf{k} = \mathbf{k}' + \mathbf{q} + \mathbf{G},\quad (5.60)$$

$$\epsilon_{\mathbf{k}} = \hbar \omega_{\mathbf{q}} + \epsilon_{\mathbf{k}'},\quad (5.61)$$

where \mathbf{G} is a reciprocal lattice vector. By this, the phase space available for scattering is strongly reduced, especially near the Fermi energy. Note that $\hbar\omega_{\mathbf{q}} \leq \hbar\omega_D \ll \epsilon_F$. In order to calculate the scattering rates, we require the matrix element of the available processes,⁵

$$\begin{aligned} & \langle \mathbf{k} + \mathbf{q}; N_{\mathbf{q}'} | (\hat{b}_{\mathbf{q}} - \hat{b}_{-\mathbf{q}}^\dagger) \hat{c}_{\mathbf{k}+\mathbf{q},s}^\dagger \hat{c}_{\mathbf{k},s} | \mathbf{k}; N_{\mathbf{q}'} \rangle \\ &= \langle \mathbf{k} + \mathbf{q} | \hat{c}_{\mathbf{k}+\mathbf{q},s}^\dagger \hat{c}_{\mathbf{k},s} | \mathbf{k} \rangle \left\{ \sqrt{N_{\mathbf{q}'}} \delta_{N_{\mathbf{q}'}, N_{\mathbf{q}'}-1} \delta_{\mathbf{q}, \mathbf{q}'} - \sqrt{N_{\mathbf{q}'} + 1} \delta_{N_{\mathbf{q}'}, N_{\mathbf{q}'}+1} \delta_{\bar{\mathbf{q}}, -\mathbf{q}'} \right\}. \end{aligned} \quad (5.62)$$

From the Golden Rule we then obtain

$$\begin{aligned} \left(\frac{\partial f}{\partial t} \right)_{\text{coll}} &= -\frac{2\pi}{\hbar} \sum_{\mathbf{q}} |g(\mathbf{q})|^2 \left[\{ f(\mathbf{k}) (1 - f(\mathbf{k} + \mathbf{q})) (1 + N_{-\mathbf{q}}) \right. \\ &\quad - f(\mathbf{k} + \mathbf{q}) (1 - f(\mathbf{k})) N_{-\mathbf{q}} \} \delta(\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}} + \hbar\omega_{-\mathbf{q}}) \\ &\quad - \{ f(\mathbf{k} + \mathbf{q}) (1 - f(\mathbf{k})) (1 + N_{\mathbf{q}}) \\ &\quad \left. - f(\mathbf{k}) (1 - f(\mathbf{k} + \mathbf{q})) N_{\mathbf{q}} \} \delta(\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}} - \hbar\omega_{\mathbf{q}}) \right], \end{aligned} \quad (5.63)$$

where $g(\mathbf{q}) = \tilde{V}_{\mathbf{q}} |\mathbf{q}| \sqrt{2\hbar/\rho_0\omega_{\mathbf{q}}}$. The four terms describe the four processes depicted in Fig. 5.4.

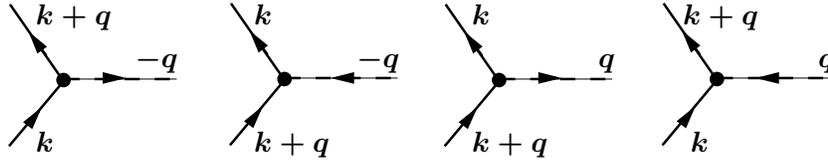


Figure 5.4: The four single-phonon electron-phonon scattering processes.

The collision integral leads to a complicated integro-differential equation, whose solution is very tedious. Instead of a rigorous calculation, we will explain the behavior in various temperature regimes by qualitative considerations.

The characteristic temperature of phonons, the Debye temperature $\Theta_D \ll T_F$, is much smaller than the Fermi temperature. Hence, the phonon energy is virtually unimportant for the energy conservation, $\epsilon_{\mathbf{k}+\mathbf{q}} \approx \epsilon_{\mathbf{k}}$. We can thus concentrate on momentum conservation and consider the lattice distortion as being essentially static, i.e., in the sense of an adiabatic approximation according to Born-Oppenheimer. The approximate collision integral then reads

$$\begin{aligned} \left(\frac{\partial f}{\partial t} \right)_{\text{coll}} &= \frac{2\pi}{\hbar} \sum_{\mathbf{q}} |g(\mathbf{q})|^2 2N(\omega_{\mathbf{q}}) \\ &\quad \times \underbrace{\{ f(\mathbf{k} + \mathbf{q}) [1 - f(\mathbf{k})] - f(\mathbf{k}) [1 - f(\mathbf{k} + \mathbf{q})] \}}_{f(\mathbf{k}+\mathbf{q}) - f(\mathbf{k})} \delta(\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}}), \end{aligned} \quad (5.64)$$

where we assume the occupation of phonon states according to the equilibrium distribution for bosons,

$$N(\omega_{\mathbf{q}}) = \frac{1}{e^{\hbar\omega_{\mathbf{q}}/k_B T} - 1}. \quad (5.65)$$

In analogy to previous approaches, we make the relaxation-time Ansatz and obtain

$$\frac{1}{\tau(\epsilon_{\mathbf{k}})} = \frac{2\pi}{\hbar} \frac{\lambda}{N(\epsilon_F)} \int \frac{d^3q}{(2\pi)^3} \hbar\omega_{\mathbf{q}} N(\omega_{\mathbf{q}}) (1 - \cos\theta) \delta(\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}}), \quad (5.66)$$

⁵In analogy to the discussion on electromagnetic radiation, the phenomenon of *spontaneous phonon emission* due to zero-point fluctuations exists. It is formally visible in the additional “+1” in the factors $(1 + N_{\pm\mathbf{q}})$.

where $|\mathbf{k}| = |\mathbf{k} + \mathbf{q}| = k_F$, i.e., the electrons close to the Fermi surface are relevant. Furthermore, we have parametrized $g(\mathbf{q})$ according to

$$|g(\mathbf{q})|^2 = \frac{\lambda}{2N(\epsilon_F)\Omega} \hbar\omega_{\mathbf{q}}, \quad (5.67)$$

where λ is a dimensionless electron-phonon coupling constant. In usual metals $\lambda < 1$. As in the case of defect scattering, the relaxation time depends only weakly on the electron energy. But, unlike previously, the direct temperature dependence of phonon occupation enters the game.

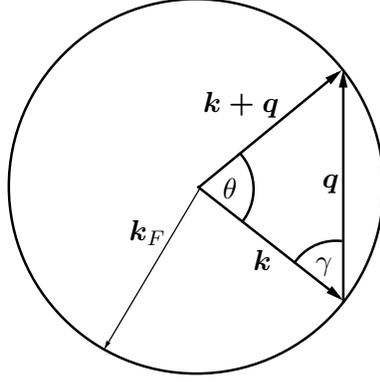


Figure 5.5: The geometry of electron-phonon scattering.

In order to execute the integral in Eq. (5.66), we have to analyze $\delta(\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}})$ and write

$$\delta(\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}}) = \delta\left(\frac{\hbar^2}{2m}\{q^2 - 2k_F q \cos \gamma\}\right) = \frac{m}{\hbar^2 k_F q} \delta\left(\frac{q}{2k_F} - \cos \gamma\right), \quad (5.68)$$

cf. Fig. 5.5. From there, we also see that $2\gamma + \theta = \pi$, and thus

$$1 - \cos \theta = 1 + \cos(2\gamma) = 2 \cos^2 \gamma. \quad (5.69)$$

Obviously, we have to integrate over $0 \leq q \leq 2k_F$, and we can reformulate the integrand according to

$$\begin{aligned} \frac{1}{\tau(T)} &= \frac{\lambda}{N(\epsilon_F)} \frac{m}{\hbar^2 \pi k_F} \int_0^{2k_F} dq q \omega_{\mathbf{q}} N(\omega_{\mathbf{q}}) \int d \cos \gamma \cos^2 \gamma \delta\left(\frac{q}{2k_F} - \cos \gamma\right) \\ &= \frac{\lambda}{4N(\epsilon_F)} \frac{m c_s}{\hbar^2 \pi k_F^3} \int_0^{2k_F} \frac{q^4 dq}{e^{\hbar c_s q / k_B T} - 1} = \frac{\lambda}{4N(\epsilon_F)} \frac{m c_s k_F^2}{\hbar^2 \pi} \left(\frac{T}{\Theta_D}\right)^5 \int_0^{2\Theta_D/T} \frac{y^4 dy}{e^y - 1}, \end{aligned} \quad (5.70)$$

where we have approximated the Debye temperature by $k_B \Theta_D \approx 2\hbar c_s k_F$. We notice two distinct characteristic temperature regimes,

$$\frac{1}{\tau} = \begin{cases} 6\zeta(5)\lambda\pi \frac{k_B \Theta_D}{\hbar} \left(\frac{T}{\Theta_D}\right)^5, & T \ll \Theta_D, \\ \lambda\pi \frac{k_B \Theta_D}{\hbar} \left(\frac{T}{\Theta_D}\right), & T \gg \Theta_D. \end{cases} \quad (5.71)$$

The prefactors depend on the details of the approximation. The qualitative temperature dependence, however, does not. We finally obtain the conductivity and resistivity from Eq. (5.23),

$$\sigma = \frac{e^2 n}{m} \tau(T), \quad (5.72)$$

$$\rho = \frac{m}{e^2 n \tau(T)}, \quad (5.73)$$

where we have used the weak energy dependence of τ ($\epsilon \approx \epsilon_F$). With this, we obtain the well-known Bloch-Grüneisen form

$$\rho(T) \propto \begin{cases} T^5, & T \ll \Theta_D, \\ T, & T \gg \Theta_D. \end{cases} \quad (5.74)$$

At high temperatures, ρ is determined by the occupation of phonon states, whereas the constraint of the scattering phase space is decisive at low temperatures.

5.5 Electron-electron scattering

In Chapter 4 we have learned, that, taking a short-ranged electron-electron interaction into account, the lifetime of a quasi particle strongly increases close to the Fermi surface. The basic reason was the constraint of the scattering phase space due to the Pauli principle. The lifetime, which we identify with the relaxation time here, has the form

$$\frac{1}{\tau(\epsilon)} = C(\epsilon - \epsilon_F)^2. \quad (5.75)$$

This allows determining the resistivity from Eq. (5.23), and we find

$$\sigma(T) = \frac{e^2 n}{m C (k_B T)^2}, \quad (5.76)$$

i.e., the resistivity $\rho \propto T^2$. This is a key property of a Fermi liquid and is often considered an identifying criterion.

However, an important point requires some explanation. One could, in principle, argue, that the momentum of the Fermi liquid is conserved upon the collision of two electrons. It is thus not quite clear what causes a finite resistance. But this argument ignores the existence of the lattice: The kinematics is also satisfied for electrons being scattered from the Fermi surface of one Brillouin zone to the one of another Brillouin zone, while incorporating a reciprocal lattice vector. By this, the momentum is transferred to the lattice. Such processes are called *Umklapp* processes and play an important role in electron-phonon scattering as well.

5.6 Matthiessen's rule and the Ioffe-Regel limit

Matthiessen's rule states, that the scattering rates of different scattering processes can simply be added, i.e.,

$$W(\mathbf{k}, \mathbf{k}') = W_1(\mathbf{k}, \mathbf{k}') + W_2(\mathbf{k}, \mathbf{k}'), \quad (5.77)$$

or, equivalently, in the relaxation time approximation,

$$\frac{1}{\tau} = \frac{1}{\tau_1} + \frac{1}{\tau_2}, \quad (5.78)$$

and

$$\rho = \frac{m}{ne^2\tau} = \frac{m}{ne^2} \left(\frac{1}{\tau_1} + \frac{1}{\tau_2} \right) = \rho_1 + \rho_2. \quad (5.79)$$

This is a rule and no theorem and corresponds effectively to a serial coupling of resistors. It is only applicable if the different scattering processes are independent. Actually, already the linear dependence of the impurity scattering rate on the impurity density n_{imp} is a consequence of Matthiessen's rule. Mutual influences of impurities, e.g., through interference effects due to the coherent scattering of a particle on different impurities, would invalidate this simplification. An

example where this happens is the reduction to one spacial dimension, where a single scatterer induces a finite resistance R . Two serial scatterers then lead to a resistance of

$$R = R_1 + R_2 + \frac{2e^2}{h} R_1 R_2 \geq R_1 + R_2. \quad (5.80)$$

The reason is that, in one-dimensional systems, the interference of backscattered waves is unavoidable and no impurity can be treated as isolated. Furthermore, every particle traversing the whole system has to pass all scatterers. A more general Matthiessen's rule,

$$\rho \geq \rho_1 + \rho_2, \quad (5.81)$$

is still valid.

For the analysis of resistance data of simple metals, we often assume the validity of Matthiessen's rule. A typical example is the resistance minimum explained by Kondo, where

$$\rho(T) = \rho_0 + \rho_{e-p}(T) + \rho_K(T) + \rho_{e-e}(T) = \rho_0 + CT^5 + C'(1 + 2JN(\epsilon_F)\ln(D/k_B T)) + AT^2. \quad (5.82)$$

Upon decreasing temperature, the Kondo term is increasing, whereas the electron-phonon and electron-electron contributions decrease. Consequently, there is a minimum.

Difficulties with Matthiessen's rule also arise, if the relaxation time depends on \mathbf{k} , since then the averaging is not the same for all scattering processes. The electron-phonon coupling can be modified by the scattering on impurities, most importantly in the presence of anisotropic Fermi surfaces.

We now turn to the discussion of resistivity in the high-temperature limit. Were we to believe the previous considerations entirely, the electrical resistivity would grow indefinitely with temperature. In most cases, however, the resistivity will saturate at a finite limiting value. We can see this from simple considerations regarding the mean free path $\ell = v_F \tau(\epsilon_F)$, the mean distance an electron travels freely between two collisions. The lattice constant a is a natural lower boundary to ℓ in the crystal lattice. Furthermore, we have so far assumed the scattering between two states with sharp momentum ($\mathbf{k} \rightarrow \mathbf{k}'$). If the de Broglie wavelength becomes comparable to the mean free path, this picture (i.e., this basis) becomes unsuitable, and k_F^{-1} is a boundary. In most systems a and k_F^{-1} are comparable lengths.

Empirically, the resistivity is described via the formula

$$\frac{1}{\rho(T)} = \frac{1}{\rho_{BT}(T)} + \frac{1}{\rho_{\max}}, \quad (5.83)$$

i.e., the parallel addition of two resistivities: $\rho_{BT}(T)$, which we have investigated using the Boltzmann transport theory, and the limiting value ρ_{\max} . This is in clear contrast to Matthiessen's rule, which is to be expected since $k_F \ell \sim 1$ will definitely lead to complex interference effects. We can, however, estimate ρ_{\max} from the Jellium model to be

$$\begin{aligned} \rho_{\max} &= \frac{m}{e^2 n \tau(\epsilon_F)} = \frac{3\pi^2 m}{e^2 k_F^3 \tau(\epsilon_F)} = \frac{h}{e^2} \frac{3\pi}{2k_F^2 \ell} \\ &\sim \frac{h}{e^2} \frac{3\pi}{2k_F} \sim (25\text{k}\Omega) \frac{3\pi}{2 \times 10^8 \text{cm}^{-1}} \approx 1\text{m}\Omega \times \text{cm}, \end{aligned} \quad (5.84)$$

where we have used $\ell^{-1} \sim k_F \sim 10^8 \text{cm}^{-1}$. This is called the Ioffe-Regel limit.⁶ Estimating ρ_{\max} for a given material is often difficult. There are even materials whose resistivity surpasses the Ioffe-Regel limit.

⁶ $\rho_{\max} \sim 1\text{m}\Omega\text{cm}$ should be compared to the room-temperature resistivity of good conductors, which are

metal	Cu	Au	Ag	Pt	Al	Sn	Na	Fe	Ni	Pb
$\rho[\mu\Omega\text{cm}]$	1.7	2.2	1.6	10.5	2.7	11	4.6	9.8	7	21

i.e., these values are well below ρ_{\max} .

5.7 General transport coefficients

Besides charge, electrons transport energy, i.e., heat and entropy. Naturally, charge and heat transport are thus interconnected. In the following, we generalize the transport theory set up above to include this.

5.7.1 Generalized Boltzmann equation and the Wiedemann-Franz law

We consider a metal with weakly space-dependent temperature $T(\mathbf{r})$ and chemical potential $\mu(\mathbf{r})$. Note that in this section we will work with the electrochemical potential $\eta(\mathbf{r}) = e\phi(\mathbf{r}) + \mu(\mathbf{r})$ where $\phi(\mathbf{r})$ denotes the electrostatic potential. The distribution function then reads

$$\delta f(\mathbf{k}; \mathbf{r}) = f(\mathbf{k}; \mathbf{r}) - f_0[\mathbf{k}, T(\mathbf{r}), \mu(\mathbf{r})], \quad (5.85)$$

where

$$f_0[\mathbf{k}, T(\mathbf{r}), \mu(\mathbf{r})] = \frac{1}{\exp\{[\epsilon_{\mathbf{k}} - \mu(\mathbf{r})]/k_B T(\mathbf{r})\} + 1}. \quad (5.86)$$

Additionally, we require the charge density to remain constant in space, i.e.,

$$\int d^3k \delta f(\mathbf{k}; \mathbf{r}) = 0 \quad (5.87)$$

for all \mathbf{r} . With this, we find the Boltzmann equation for the stationary situation,

$$\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = \mathbf{v}_{\mathbf{k}} \cdot \frac{\partial f}{\partial \mathbf{r}} + \dot{\mathbf{k}} \cdot \frac{\partial f}{\partial \mathbf{k}} = -\frac{\partial f}{\partial \epsilon_{\mathbf{k}}} \mathbf{v}_{\mathbf{k}} \cdot \left\{ \frac{\partial T}{\partial \mathbf{r}} \frac{\epsilon_{\mathbf{k}} - \mu}{T} - \mathcal{E} \right\}, \quad (5.88)$$

with $\mathcal{E} = -\nabla(e\phi + \mu)$, where $e\phi + \mu$ is called *electrochemical potential* ($\mathbf{E} = -\nabla\phi$).

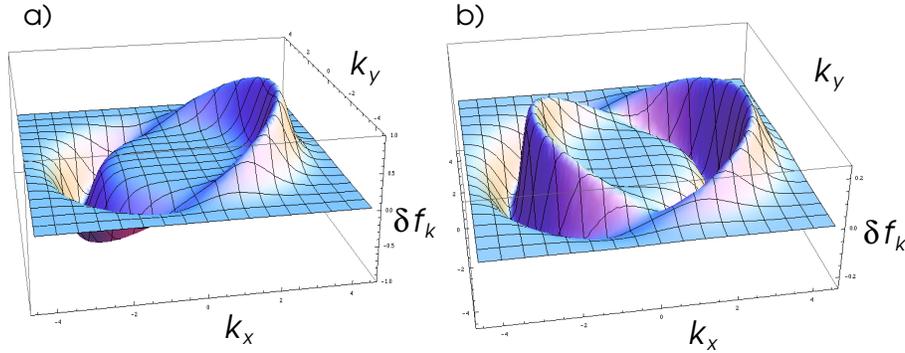


Figure 5.6: Schematic distribution functions $\delta f(\mathbf{k})$ on a cut through the k -space ($k_z = 0$) with a circular Fermi surface for two situations: a) for an applied electric field negative x -direction; b) for a temperature gradient in x -direction.

In the relaxation time approximation for the collision integral we obtain the solution

$$\delta f(\mathbf{k}) = -\frac{\partial f_0}{\partial \epsilon_{\mathbf{k}}} \tau(\epsilon_{\mathbf{k}}) \mathbf{v}_{\mathbf{k}} \cdot \left\{ \mathcal{E} - \frac{\epsilon_{\mathbf{k}} - \mu}{T} \frac{\partial T}{\partial \mathbf{r}} \right\}, \quad (5.89)$$

from which we find the charge and heat currents,

$$\mathbf{J}_e = 2 \int \frac{d^3k}{(2\pi)^3} e \mathbf{v}_{\mathbf{k}} \delta f_{\mathbf{k}}, \quad (5.90)$$

$$\mathbf{J}_q = 2 \int \frac{d^3k}{(2\pi)^3} (\epsilon_{\mathbf{k}} - \mu) \mathbf{v}_{\mathbf{k}} \delta f_{\mathbf{k}}, \quad (5.91)$$

respectively. Inserting the solution yields

$$\mathbf{J}_e = e^2 \hat{K}^{(0)} \mathbf{E} + \frac{e}{T} \hat{K}^{(1)} (-\nabla T), \quad (5.92)$$

$$\mathbf{J}_q = e \hat{K}^{(1)} \mathbf{E} + \frac{1}{T} \hat{K}^{(2)} (-\nabla T), \quad (5.93)$$

where

$$K_{\alpha\beta}^{(n)} = -\frac{1}{4\pi^3} \int d\epsilon \frac{\partial f_0}{\partial \epsilon} \tau(\epsilon) (\epsilon - \mu)^n \int d\Omega_{\mathbf{k}} \frac{v_{F\alpha} v_{F\beta}}{\hbar |\mathbf{v}_F|}. \quad (5.94)$$

In the case $T \ll T_F$ we can calculate the coefficients,⁷

$$K_{\alpha\beta}^{(0)} = \frac{1}{4\pi^3 \hbar} \tau(\epsilon_F) \int d\Omega_{\mathbf{k}} \frac{v_{F\alpha} v_{F\beta}}{|\mathbf{v}_F|}, \quad (5.97)$$

$$\hat{K}^{(1)} = \frac{\pi^2}{3} (k_B T)^2 \left. \frac{\partial}{\partial \epsilon} \hat{K}^{(0)}(\epsilon) \right|_{\epsilon=\epsilon_F}, \quad (5.98)$$

$$\hat{K}^{(2)} = \frac{\pi^2}{3} (k_B T)^2 \hat{K}^{(0)}(\epsilon_F). \quad (5.99)$$

We measure the electrical resistivity assuming thermal equilibrium, i.e., $\nabla T = 0$ for all \mathbf{r} . With this, as before,

$$\sigma_{\alpha\beta} = e^2 K_{\alpha\beta}^{(0)}. \quad (5.100)$$

To determine the thermal conductivity κ , we set $\mathbf{J}_e = 0$ (open circuit). Then, Eqs. (5.92) and (5.93) reveal the appearance of an electrochemical field according to

$$\boldsymbol{\mathcal{E}} = \frac{1}{T} \hat{K}^{(0)-1} \hat{K}^{(1)} \nabla T. \quad (5.101)$$

Thus, the heat current is given by

$$\mathbf{J}_q = -\frac{1}{T} \left(\hat{K}^{(2)} - \hat{K}^{(1)} \hat{K}^{(0)-1} \hat{K}^{(1)} \right) \nabla T = -\hat{\kappa} \nabla T. \quad (5.102)$$

In simple metals, the second term is often negligible and one obtains

$$\hat{\kappa} = \frac{1}{T} \hat{K}^{(2)} = \frac{\pi^2 k_B^2}{3} T \hat{K}^{(0)} = \frac{\pi^2 k_B^2}{3 e^2} T \hat{\sigma}, \quad (5.103)$$

which is the well-known *Wiedemann-Franz law*. Note, that we can write the thermal conductivity in the form

$$\hat{\kappa} = \frac{C}{e^2 N(\epsilon_F)} \hat{\sigma}, \quad (5.104)$$

with the electronic specific heat $C = \pi^2 k_B^2 T/3$.

⁷ If $g(\epsilon)$ depends only weakly on ϵ , we can use the Taylor expansion to derive a general approximation for $T \rightarrow 0$ according to

$$-\int d\epsilon g(\epsilon) \frac{\partial f_0}{\partial \epsilon} = g(\epsilon_F) + \frac{\pi^2}{6} (k_B T)^2 \left. \frac{\partial^2 g(\epsilon)}{\partial \epsilon^2} \right|_{\epsilon=\epsilon_F} + \dots \quad (5.95)$$

and

$$-\int d\epsilon g(\epsilon) (\epsilon - \epsilon_F) \frac{\partial f_0}{\partial \epsilon} = \frac{\pi^2}{3} (k_B T)^2 \left. \frac{\partial g(\epsilon)}{\partial \epsilon} \right|_{\epsilon=\epsilon_F}, \quad (5.96)$$

where we have used $\mu \rightarrow \epsilon_F$ for $T \rightarrow 0$.

5.7.2 Thermoelectric effect

In Eq. (5.101) we have seen that a temperature gradient induces an electric field. For simplicity, we assume an isotropic system. Then,

$$\mathcal{E} = Q \nabla T \quad (5.105)$$

with the Seebeck coefficient

$$Q = - \left. \frac{\pi^2 k_B^2 T}{3} \frac{\sigma'(\epsilon)}{e \sigma(\epsilon)} \right|_{\epsilon=\epsilon_F}. \quad (5.106)$$

Using $\sigma(\epsilon) = n(\epsilon)e^2\tau(\epsilon)/m$, we investigate $\sigma'(\epsilon)$,

$$\sigma'(\epsilon) = \frac{\tau'(\epsilon)}{\tau(\epsilon)}\sigma(\epsilon) + \frac{n'(\epsilon)}{n(\epsilon)}\sigma(\epsilon) = \frac{\tau'(\epsilon)}{\tau(\epsilon)}\sigma(\epsilon) + \frac{N(\epsilon)}{n(\epsilon)}\sigma(\epsilon), \quad (5.107)$$

i.e., we obtain a contribution if the relaxation time depends strongly on energy. This is most prominent, if there is resonant scattering involved (e.g., the Kondo effect). If the first term is irrelevant, we find

$$Q = - \frac{\pi^2 k_B^2 T}{3} \frac{N(\epsilon_F)}{e n(\epsilon_F)} = - \frac{S}{ne}, \quad (5.108)$$

which corresponds to the entropy per electron.

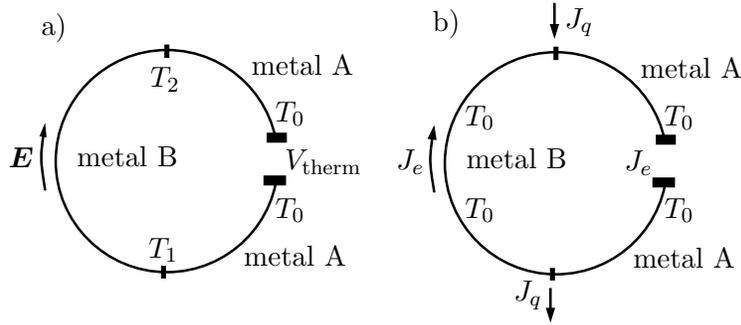


Figure 5.7: Schematics of thermoelectric effects: a) Seebeck effect, b) Peltier effect.

We consider two types of thermoelectric effects. The first is the *Seebeck effect*, where a thermoelectric voltage appears in a bi-metallic system (cf. Fig. 5.7). With Eq. (5.105), a temperature gradient across metal B induces an electromotoric force

$$\begin{aligned} \text{EMF} &= \int d\mathbf{s} \cdot \mathbf{E} = \int d\mathbf{s} \cdot Q \nabla T \\ &= \int_{T_0}^{T_1} Q_A d\vec{s} \cdot \vec{\nabla} T + \int_{T_1}^{T_2} Q_B d\vec{s} \cdot \vec{\nabla} T + \int_{T_2}^{T_0} Q_A d\vec{s} \cdot \vec{\nabla} T = (Q_B - Q_A)(T_2 - T_1). \end{aligned} \quad (5.109)$$

The resulting voltage appears in a second metal A, whose contacts are kept at the same temperature T_0 . Here, a bi-metallic configuration was chosen to reveal voltage differences across the contacts which are absent in a single metal.

The second phenomenon, the so-called *Peltier effect*, emerges in a system kept at the same temperature everywhere. Here, an electric current induces a heat current in a bi-metallic system (see Fig. 5.7), such that heat is transferred from one reservoir to another. This follows from Eqs. (5.92) and (5.92) by assuming $\nabla T = 0$, where

$$\left. \begin{aligned} J_e &= e^2 K^{(0)} \vec{E} \\ J_q &= e K^{(1)} \vec{E} \end{aligned} \right\} \Rightarrow J_q = \frac{K^{(1)}}{e K^{(0)}} J_e = T Q J_e = \Pi J_e. \quad (5.110)$$

Π is called Peltier coefficient. In Fig. 5.7 we see, that one obtains a contribution to the heat current from both metals A and B,

$$J_q = (\Pi_A - \Pi_B)J_e = T_0(Q_A - Q_B)J_e. \quad (5.111)$$

This means, that we can control the heat transfer between reservoirs using electrical current.

5.8 Transport in one dimension – Anderson localization

Transport in one spatial dimension is very special, since there are only two ways to go: left and right. We introduce the transfer matrix formalism and use it to express the conductivity through the Landauer formula. We will investigate the effects of multiple scattering at different obstacles, leading to the so-called Anderson localization, which turns a metal into an insulator.

5.8.1 Landauer Formula

The transmission and reflection at an arbitrary potential with finite support in one dimension can be described by a transfer matrix. A suitable basis for the electrons in this case are left-

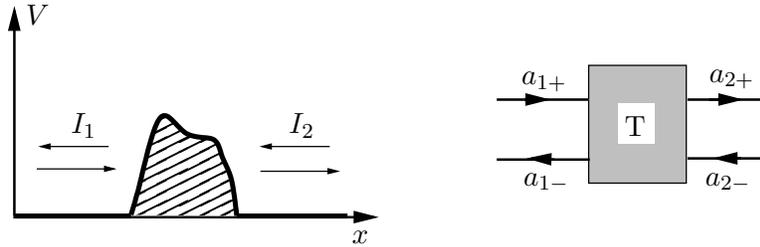


Figure 5.8: Transfer matrices describe potential scattering in one dimension.

and right moving plane waves (cf. Fig. 5.8) with wave vector $\pm k$. Only plane waves with the same $|k|$ on the left and right side of the scatterer are interconnected. Therefore, we write

$$\psi_1(x) = a_{1+}e^{ikx} + a_{1-}e^{-ikx}, \quad (5.112)$$

$$\psi_2(x) = a_{2+}e^{ikx} + a_{2-}e^{-ikx}, \quad (5.113)$$

where ψ_1 (ψ_2) is defined in the area I_1 (I_2). The vectors $a_i = (a_{i+}, a_{i-})$ are connected via a linear relation,

$$a_2 = T a_1 = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} a_1, \quad (5.114)$$

with the 2×2 transfer matrix T . The conservation of current ($J_1 = J_2$) requires $\det T = 1$, i.e., T is unimodular. Here,

$$J = \frac{i\hbar}{2m} \left\{ \frac{d\psi(x)^*}{dx} \psi(x) - \psi(x)^* \frac{d\psi(x)}{dx} \right\}, \quad (5.115)$$

such that, for a plane wave $\psi(x) = L^{-1/2}e^{ikx}$, $J = \hbar k/Lm = v/L$ (system size L).

Time reversal symmetry implies that, with $\psi(x)$, $\psi(x)^*$ is a solution of the stationary Schrödinger equation as well. From this, we find $T_{11} = T_{22}^*$ and $T_{12} = T_{21}^*$, i.e.,

$$T = \begin{pmatrix} T_{11} & T_{12} \\ T_{12}^* & T_{11}^* \end{pmatrix}. \quad (5.116)$$

It is easily shown that a shift of the scattering potential by a distance x_0 changes the coefficients of T by a phase factor, $T_{11} \rightarrow T_{11}$ and $T_{12} \rightarrow T_{12}e^{i2kx_0}$.

We can connect the coefficients of T with the transmission- and reflection amplitudes of the previous sections: With the ansatz for a right moving (unnormalized) wave,

$$\psi_1(x) = e^{ikx} + re^{-ikx}, \quad (5.117)$$

$$\psi_2(x) = te^{ikx}, \quad (5.118)$$

we find via insertion that

$$T = \begin{pmatrix} 1/t^* & -r^*/t^* \\ -r/t & 1/t \end{pmatrix}. \quad (5.119)$$

Then the conservation of currents imposes the condition $1 = |r|^2 + |t|^2$ making the matrix unimodular.

We can find a relation between r and t of a potential barrier and the electric resistivity. The incoming current density J_0 is split into a reflected and transmitted part, J_r and J_t , respectively, all given by

$$J_0 = -\frac{1}{L}ve, \quad J_r = -\frac{|r|^2}{L}ve, \quad J_t = -\frac{|t|^2}{L}ve, \quad (5.120)$$

with the velocity $v = \hbar k/m$ and the electron charge $-e$. The electron density on the two sides of the barrier is given by

$$n_1 = \frac{1 + |r|^2}{L}, \quad (5.121)$$

$$n_2 = \frac{|t|^2}{L}, \quad (5.122)$$

from which a density difference $\delta n = n_1 - n_2 = (1 + |r|^2 - |t|^2)/L = 2|r|^2/L$ results. Consequently, there is an energy difference $\delta E = -e\delta V$ between the two sides, corresponding to a voltage via

$$\delta n = \frac{dn}{dE}\delta E = -\frac{dn}{dE}e\delta V. \quad (5.123)$$

Here, $\frac{dn}{dE}dE$ is the number of states per length in the energy interval $[E, E + dE]$, i.e.,

$$\frac{dn}{dE} = \frac{1}{L} \sum_{\vec{k}, s} \delta\left(E - \frac{\hbar^2 k^2}{2m}\right) = 2 \int \frac{dk}{2\pi} \delta\left(E - \frac{\hbar^2 k^2}{2m}\right) = \frac{1}{\pi \hbar v(E)}. \quad (5.124)$$

The resistance is obtained from

$$R = \frac{\delta V}{J_t} = -\frac{\delta n}{e} \left(\frac{dn}{dE}\right)^{-1} \frac{1}{J_t} = \frac{h}{e^2} \frac{|r|^2}{|t|^2}, \quad (5.125)$$

where $h/e^2 \approx 25.8k\Omega$ is the resistance quantum. This is the famous *Landauer formula*, which is valid for all one-dimensional systems and which is often used for the description of mesoscopic systems and quantum wires.

We consider now two spatially separated scattering potentials, represented by T_1 and T_2 (r_1, t_1 bzw. r_2, t_2). The particles are multiply scattered at these potentials in a complicated manner,

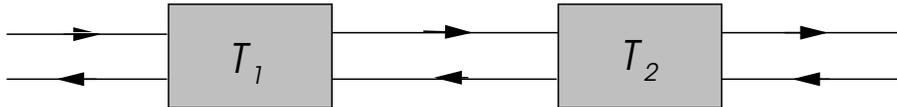


Figure 5.9: Two spatially separated scattering potentials.

but the result can, again, be expressed via a simple transfer matrix by multiplying T_1 and T_2 . All previously found properties remain valid for the new total matrix T , given by

$$T = T_1 T_2 = \begin{pmatrix} \frac{1}{t_1^* t_2^*} + \frac{r_1^* r_2}{t_1^* t_2} & -\frac{r_2^*}{t_1^* t_2^*} - \frac{r_1^*}{t_1^* t_2} \\ -\frac{r_1}{t_1 t_2^*} - \frac{r_2}{t_1 t_2} & \frac{1}{t_1 t_2} + \frac{r_1 r_2^*}{t_1 t_2^*} \end{pmatrix} = \begin{pmatrix} 1/t^* & -r^*/t^* \\ -r/t & 1/t \end{pmatrix}. \quad (5.126)$$

For the ratio between reflection and transmission probability we find

$$\begin{aligned} \frac{|r|^2}{|t|^2} &= \frac{1}{|t|^2} - 1 = \frac{1}{|t_1|^2 |t_2|^2} \left| 1 + \frac{r_1 r_2^* t_2}{t_2^*} \right|^2 - 1 \\ &= \frac{1}{|t_1|^2 |t_2|^2} \left(1 + |r_1|^2 |r_2|^2 + \frac{r_1 r_2^* t_2}{t_2^*} + \frac{r_1^* r_2 t_2^*}{t_2} \right) - 1. \end{aligned} \quad (5.127)$$

Assuming an arbitrary distance $d = x_2 - x_1$ between the two potential barriers, we average over this distance. Note, that $r_2/t_2 \propto e^{-2ikd}$ and, therefore, terms with such a factor vanish after averaging. The remainders of Eq. (5.127) can be collected to

$$\left. \frac{|r|^2}{|t|^2} \right|_{\text{avg}} = \frac{1}{|t_1|^2 |t_2|^2} (1 + |r_1|^2 |r_2|^2) - 1 = \frac{|r_1|^2}{|t_1|^2} + \frac{|r_2|^2}{|t_2|^2} + 2 \frac{|r_1|^2 |r_2|^2}{|t_1|^2 |t_2|^2}. \quad (5.128)$$

Even though two scattering potentials are added in series, a non-linear combination instead of the sum of the two ratios $|r_i|^2/|t_i|^2$ emerges. From the Landauer formula we see, that resistances add differently. Adding R_1 and R_2 serially, the total resistance is not given by $R = R_1 + R_2$, but by

$$R = R_1 + R_2 + \frac{2e^2}{h} R_1 R_2 > R_1 + R_2. \quad (5.129)$$

This is a consequence of the unavoidable multiple scattering. It is particularly prominent if $R_i \gg h/e^2$, where resistances are multiplied instead of summed.

5.8.2 Anderson localization

Let us consider a system with many arbitrarily distributed scatterers, and let ρ be a mean resistance per unit length. R_0 shall be the resistance between points 0 and ℓ_0 . The change in resistance by advancing an infinitesimal $\delta\ell$ is found from

$$dR = \rho d\ell + \frac{2e^2}{h} R \rho d\ell, \quad (5.130)$$

which yields

$$\int_{R_0}^R \frac{dR}{1 + \frac{2e^2}{h} R} = \int_{\ell_0}^{\ell} \rho d\ell, \quad (5.131)$$

and, thus,

$$\frac{h}{2e^2} \ln \left(\frac{1 + 2e^2 R(\ell)/h}{1 + 2e^2 R_0/h} \right) = \rho(\ell - \ell_0). \quad (5.132)$$

Since $R_0 \rightarrow 0$ for $\ell_0 \rightarrow 0$,

$$R(\ell) = \frac{h}{2e^2} \left(e^{2e^2 \rho \ell/h} - 1 \right). \quad (5.133)$$

Obviously, $R \rightarrow \infty$ very rapidly for increasing ℓ . This means, that this system is an insulator for arbitrarily small but finite $\rho > 0$. The reason for this is that, in one dimension, all states are *localized* (i.e., bound states) in the presence of disorder. This phenomenon is called *Anderson localization*. Even though all states are localized, the energy spectrum is continuous, as infinitely

many bound states with different energy exist. The mean localization length ξ (related to mean extension of wave functions) of individual states is found from Eq. (5.133) to be $\xi = h/e^2\rho$. The transmission amplitude is reduced on this length scale as well, since $|t| \approx 2e^{-\ell/\xi}$ for $\ell \gg \xi$.⁸ In one dimension, there is no “normal” electric resistance [$R(\ell) \approx \rho\ell$] for non-interacting particles, only two extremes: Either, the potential is perfectly periodic and the states correspond to Bloch waves. Then, coherent constructive interference produces extended states that propagate freely throughout the system and therefore a perfect conductor without resistance⁹. On the other hand, if the scattering potential is disordered, all states are localized. In this case, there is no propagation and the system is an insulator. In three-dimensional systems, multiple scattering is far less dangerous and the Ohmic law is applicable. Localization in two dimensions is very subtle and part of today’s research.

⁸For an expanded discussion of this topic, the article “*New method for a scaling theory of localization*” by P.W. Anderson, D.J. Thouless, E. Abrahams, and D.S. Fisher, Physical Review B **22**, 3519 (1980) is recommended.

⁹We have also seen in the context of chiral edge states in the Quantum Hall state, that perfect conductance in a one-dimensional channel, if there is no backscattering due to the lack of states which move in the opposite direction, i.e. in chiral states particle move only in one direction

Chapter 6

Magnetism in metals

Magnetic ordering in metals can be viewed as an instability of the Fermi liquid state. We enter this new behavior of metals through a detailed description of the Stoner ferromagnetism. The discussion of antiferromagnetism and spin density wave phases will be only brief. In Stoner ferromagnets the magnetic moment is provided by the spin of itinerant electrons. Magnetism due to localized magnetic moments will be considered in the context of Mott insulators which are subject of the next chapter.

Well-known examples of elemental ferromagnetic metals are Fe, Co and Ni belonging to the $3d$ transition metals, where the $3d$ -orbital character dominates the conduction electrons at the Fermi energy. These orbitals are rather tightly bound to the ion cores such that their mobility is reduced, enhancing the importance of interactions which essential for the formation of a magnetic state, as we will see below.

Other forms of magnetism (spin density waves and antiferromagnetism) are found in the $3d$ transition metals Cr and Mn, whereas $4d$ and $5d$ transition metals within the same columns of the periodic system are not magnetic. Their d -orbitals are more extended, leading to a higher mobility of the electrons. The $4d$ -elements Pd and Rh and the $5d$ -element Pt are, however, nearly ferromagnetic. The $4f$ -orbitals appearing in the lanthanides are nearly localized and can lead to ferromagnetism, as illustrated by the elements going from Gd through Tm in the periodic system.

Magnetism appears through a phase transition, i.e., the metal is non-magnetic at sufficiently high temperatures above a critical temperature T_c (cf. Table 6.1). In many cases magnetism appears at T_c as a continuous transition, a second order phase transition (lacking latent heat but featuring a discontinuity in the specific heat) involving the spontaneous violation of symmetry.

element	T_c (K)	type	element	T_c (K)	type
Fe	1043	ferromagnet (3d)	Gd	293	ferromagnet (4f)
Co	1388	ferromagnet (3d)	Dy	85	ferromagnet (4f)
Ni	627	ferromagnet (3d)	Cr	312	spin density wave (3d)
ZrZn ₂	22	ferromagnet	α -Mn	100	antiferromagnet
Pd	–	paramagnet	Pt	–	paramagnet
HfZn ₂	–	paramagnet			

Table 6.1: Selection of ferromagnetic materials with their respective form of magnetism and the critical temperature T_c .

6.1 Stoner instability

We focus here on the metallic ferromagnetism originating from the Stoner mechanism. In close analogy to the first Hund's rule, the exchange interaction is crucial here, i.e., the alignment of

their spins allows the electrons to reduce the energy cost due to Coulomb repulsion. According to Landau's theory of Fermi liquids, the interaction between electrons renormalizes the spin susceptibility χ to

$$\chi = \frac{m^*}{m} \frac{\chi_0}{1 + F_0^a}, \quad (6.1)$$

which diverges for $F_0^a \rightarrow -1$. This points towards a possible magnetic instability of the Fermi liquid, which we discuss within the Stoner model here.

6.1.1 Stoner model within the mean field approximation

We consider a model of conduction electrons with a repulsive contact interaction,

$$\mathcal{H} = \sum_{\mathbf{k}, s} \epsilon_{\mathbf{k}} \hat{c}_{\mathbf{k}s}^\dagger \hat{c}_{\mathbf{k}s} + U \int d^3r d^3r' \hat{\rho}_\uparrow(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}') \hat{\rho}_\downarrow(\mathbf{r}'), \quad (6.2)$$

where we use the density $\hat{\rho}_s(\mathbf{r}) = \hat{\Psi}_s^\dagger(\mathbf{r}) \hat{\Psi}_s(\mathbf{r})$. Due to the Pauli exclusion principle, the contact interaction is only active between electrons with opposite spins. This is a consequence of the exchange hole in the two-particle correlation between electrons of identical spin.

The general solution of this model is difficult. However, a mean field approximation will provide very useful insights. We rewrite,

$$\hat{\rho}_s(\mathbf{r}) = n_s + \{\hat{\rho}_s(\mathbf{r}) - n_s\}, \quad (6.3)$$

where

$$n_s = \langle \hat{\rho}_s(\mathbf{r}) \rangle. \quad (6.4)$$

We stipulate that the deviation from the mean value n_s shall be small, i.e.,

$$\langle \{\hat{\rho}_s(\mathbf{r}) - n_s\}^2 \rangle \ll n_s^2. \quad (6.5)$$

Inserting Eq. (6.3) into the Hamiltonian Eq. (6.2) we obtain

$$\begin{aligned} \mathcal{H}_{\text{mf}} &= \sum_{\mathbf{k}, s} \epsilon_{\mathbf{k}} \hat{c}_{\mathbf{k}s}^\dagger \hat{c}_{\mathbf{k}s} + U \int d^3r \{\hat{\rho}_\uparrow(\mathbf{r}) n_\downarrow + \hat{\rho}_\downarrow(\mathbf{r}) n_\uparrow - n_\uparrow n_\downarrow\} + \dots \\ &= \sum_{\mathbf{k}, s} \{\epsilon_{\mathbf{k}} + U n_{-s}\} \hat{c}_{\mathbf{k}s}^\dagger \hat{c}_{\mathbf{k}s} - U \Omega n_\uparrow n_\downarrow + \dots \end{aligned} \quad (6.6)$$

This mean field Hamiltonian describes electrons which move in the uniform background of electrons of opposite spin coupling via the (spin dependent) Coulomb interaction (exchange interaction). Fluctuations are suppressed here. The advantage of this approximation is, that we are now dealing with an effective one-particle problem, where only the mean electron interaction is taken into account (this is a generalized Hartree-Fock approximation). This enables us to calculate a few expectation values, e.g., the density of one spin species

$$\begin{aligned} n_\uparrow &= \frac{1}{\Omega} \sum_{\mathbf{k}} \langle \hat{c}_{\mathbf{k}\uparrow}^\dagger \hat{c}_{\mathbf{k}\uparrow} \rangle = \frac{1}{\Omega} \sum_{\mathbf{k}} f(\epsilon_{\mathbf{k}} + U n_\downarrow) = \int d\epsilon \underbrace{\frac{1}{\Omega} \sum_{\mathbf{k}} \delta(\epsilon - \epsilon_{\mathbf{k}} - U n_\downarrow)}_{= \frac{1}{2} N(\epsilon - U n_\downarrow)} f(\epsilon), \end{aligned} \quad (6.7)$$

and analogously, for the opposite spin direction. These mean densities are determined self-consistently, i.e., the insertion of n_s in into the mean field Hamiltonian Eq. (6.6) must provide the output according to the expectation values Eq. (6.7), under the constraint that the total number of electrons is conserved. We define the ‘‘magnetization’’ m (the real magnetization $M = \mu_B m$) via

$$n_s = \frac{1}{2} ((n_\uparrow + n_\downarrow) + s(n_\uparrow - n_\downarrow)) = \frac{n_0 + sm}{2}, \quad (6.8)$$

with the particle density n_0 . We end up with the two coupled equations

$$\begin{aligned} n_0 &= \frac{1}{2} \int d\epsilon \{N(\epsilon - Un_\downarrow) + N(\epsilon - Un_\uparrow)\} f(\epsilon) \\ &= \frac{1}{2} \sum_s \int d\epsilon N\left(\epsilon - \frac{Un_0}{2} - s\frac{Um}{2}\right) f(\epsilon), \end{aligned} \quad (6.9)$$

$$\begin{aligned} m &= \frac{1}{2} \int d\epsilon \{N(\epsilon - Un_\downarrow) - N(\epsilon - Un_\uparrow)\} f(\epsilon) \\ &= -\frac{1}{2} \sum_s \int d\epsilon s N\left(\epsilon - \frac{Un_0}{2} - s\frac{Um}{2}\right) f(\epsilon). \end{aligned} \quad (6.10)$$

Usually these equations cannot be solved analytically and have to be treated numerically.

6.1.2 Stoner criterion

An approximate solution is possible, if $m \ll n_0$. For this purpose Eqs. (6.9) and (6.10) solved adapting the chemical potential μ . For low temperatures and small magnetization we can expand μ as

$$\mu(m, T) = \epsilon_F + \Delta\mu(m, T), \quad (6.11)$$

where the constant energy shift $-Un_0/2$ in Eqs.(6.9) and (6.10) has been absorbed into ϵ_F . We use the Fermi-Dirac distribution in the form

$$f(\epsilon) = \frac{1}{e^{\beta[\epsilon - \mu(m, T)]} + 1}, \quad (6.12)$$

where $\beta = (k_B T)^{-1}$. We first expand Eq. (6.9) and obtain

$$\begin{aligned} n_0 &\approx \int d\epsilon f(\epsilon) \left\{ N(\epsilon) + \frac{1}{2} \left(\frac{Um}{2}\right)^2 N''(\epsilon) \right\} \\ &\approx \underbrace{\int_0^{\epsilon_F} d\epsilon N(\epsilon)}_{=n_0} + N(\epsilon_F) \Delta\mu + \frac{\pi^2}{6} (k_B T)^2 N'(\epsilon_F) + \frac{1}{2} \left(\frac{Um}{2}\right)^2 N'(\epsilon_F), \end{aligned} \quad (6.13)$$

where $N'(\epsilon) = dN(\epsilon)/d\epsilon$ and $N''(\epsilon) = d^2N(\epsilon)/d\epsilon^2$. Since the first term on the right side is identical to n_0 , we immediately end up with

$$\Delta\mu(m, T) \approx -\frac{N'(\epsilon_F)}{N(\epsilon_F)} \left\{ \frac{\pi^2}{6} (k_B T)^2 + \frac{1}{2} \left(\frac{Um}{2}\right)^2 \right\}, \quad (6.14)$$

since the remaining terms have to cancel. Analogously we treat Eq. (6.10) and expand in m and T ,

$$\begin{aligned} m &\approx \int d\epsilon f(\epsilon) \left[N'(\epsilon) \frac{Um}{2} + \frac{1}{3!} N'''(\epsilon) \left(\frac{Um}{2}\right)^3 \right] \\ &\approx \left\{ N(\epsilon_F) + \frac{\pi^2}{6} (k_B T)^2 N''(\epsilon_F) + \frac{1}{3!} \left(\frac{Um}{2}\right)^2 N''(\epsilon_F) + \Delta\mu N'(\epsilon_F) \right\} \frac{Um}{2}, \end{aligned} \quad (6.15)$$

and finally

$$m = \frac{m}{2} UN(\epsilon_F) \left[1 - \frac{\pi^2}{6} (k_B T)^2 \Lambda_1(\epsilon_F)^2 \right] - N(\epsilon_F) \Lambda_2(\epsilon_F)^2 \left(\frac{Um}{2}\right)^3, \quad (6.16)$$

where

$$\Lambda_1(\epsilon_F)^2 = \frac{N'(\epsilon_F)^2}{N(\epsilon_F)^2} - \frac{N''(\epsilon_F)}{N(\epsilon_F)}, \quad \Lambda_2(\epsilon_F) = \frac{1}{2} \left(\frac{N'(\epsilon_F)^2}{N(\epsilon_F)^2} - \frac{N''(\epsilon_F)}{3N(\epsilon_F)} \right). \quad (6.17)$$

The structure of Eq. (6.16) is $m = am + bm^3$, assuming $b < 0$; it thus has two types of solutions,

$$m^2 = \begin{cases} 0, & a < 1, \\ \frac{1-a}{b}, & a \geq 1. \end{cases} \quad (6.18)$$

With this, $a = 1$ corresponds to a critical value, i.e.,

$$1 = \frac{1}{2} UN(\epsilon_F) \left[1 - \frac{\pi^2}{6} (k_B T)^2 \Lambda_1(\epsilon_F)^2 \right], \quad (6.19)$$

yielding

$$k_B T_C = \frac{\sqrt{6}}{\pi \Lambda_1(\epsilon_F)} \sqrt{1 - \frac{2}{UN(\epsilon_F)}} = \frac{\sqrt{6}}{\pi \Lambda_1(\epsilon_F)} \sqrt{1 - \frac{U_c}{U}} \quad (6.20)$$

for $U > U_c = 2/N(\epsilon_F)$ (cf. Fig. 6.1). This is an instability condition for the nonmagnetic Fermi

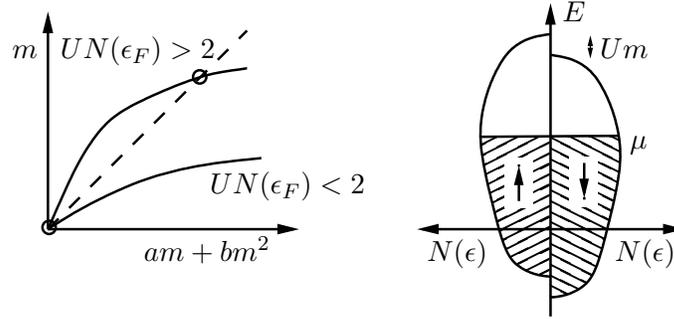


Figure 6.1: Graphical solution of Eq. (6.16) and the resulting magnetization.

liquid state with $m = 0$, and T_C is the *Curie temperature* below which the ferromagnetic state appears. The temperature dependence of the magnetization M of the ferromagnetic state is given by

$$M(T) = \mu_B m(T) \propto \sqrt{T_C - T}, \quad T < T_C, \quad (6.21)$$

if $T_C - T \ll T_C$. Note that T_C is only finite for $UN(\epsilon_F) > 2$, and $T_C \rightarrow 0$ for $UN(\epsilon_F) \rightarrow 2_+$. This condition for a finite T_C , $UN(\epsilon_F) > 2$ is known as the *Stoner criterion*.

Here, we also have obtained a simple model for a so-called *quantum phase transition*, i.e., a phase transition at $T = 0$ as a function of system parameters, e.g., the density of states or the Coulomb repulsion. While thermal fluctuations destroy the ordered state at finite temperature via entropy increase (an enhanced entropy S leads to a lowering of the free energy F at finite temperatures, $dF = -SdT + \dots$), entropy is irrelevant at $T = 0$. Instead, order is suppressed by quantum fluctuations (Heisenberg's uncertainty principle). The density of states as an internal parameter can, e.g., be changed by external pressure. Pressure by reducing the lattice constant may facilitate the motion of the conduction electrons (increased overlap of atomic orbitals) and increase the Fermi velocity. Consequently, the density of states is reduced (cf. Fig. 6.2). In fact, pressure is able to destroy ferromagnetism in weakly ferromagnetic materials as ZrZn_2 , MnSi , and UGe_2 . In other materials, the Curie temperature is high enough, such that the technologically feasible pressure is insufficient to suppress magnetism. It is, however, possible, that pressure leads to other transitions that eventually destroy magnetism, e.g., structural phase transitions. This is seen in iron (Fe), where a pressure of about 12 GPa induces a transition

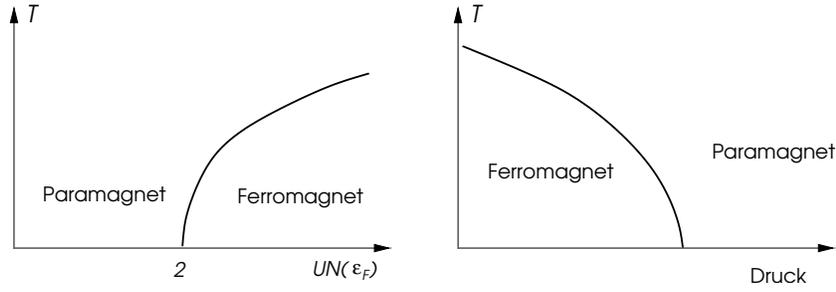


Figure 6.2: Phase diagram of a Stoner ferromagnet in the $T-UN(\epsilon_F)$ and $T-p$ plane, respectively.

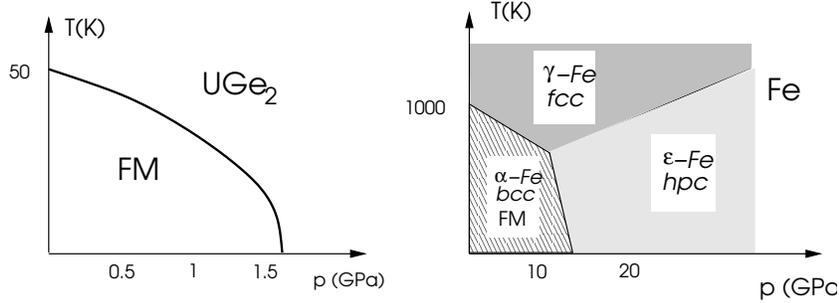


Figure 6.3: Phase diagrams of UGe_2 and Fe.

from magnetic iron with body-centered crystal (bcc) structure to a nonmagnetic, hexagonal close packed (hcp) structure (cf. Fig. 6.3). While this form of transition is a quantum phase transition as well, it appears mostly as a discontinuous transition of first order.¹ Pressure can also induce an increase in $N(\epsilon_F)$, e.g., in metals with multiple bands, where compression leads to a redistribution of charge. One example is most likely $Sr_3Ru_2O_7$ for which uniaxial pressure along the z -axis leads to magnetism.

Let us eventually turn to the question, why is Cu, being a direct neighbor of Ni in the $3d$ -row of the periodic table, not ferromagnetic? Both elemental metals even share the same crystal structure (fcc). This can be understood from the Stoner instability criterion ($UN(\epsilon_F) = 2$): While the conduction electrons at the Fermi level of Ni have $3d$ -character and belong to a narrow band with a large density of states, the Fermi energy of Cu is situated in the broad $4s$ -band and constitutes a much smaller density of states (cf. Fig. 6.4). With this, the Cu conduction electrons are much less localized and feature a weaker tendency towards ferromagnetism. Cu is known to be a better conductor than Ni for this reason.

6.1.3 Spin susceptibility for $T > T_C$

We apply an infinitesimal magnetic field H along the z -axis, which induces a spin polarization due to the Zeeman coupling. From the self-consistency equations we obtain

$$\begin{aligned}
 m &= -\frac{1}{2} \int d\epsilon f(\epsilon) \sum_s s N \left(\epsilon - \mu_B s H - s \frac{Um}{2} \right) \approx \int d\epsilon f(\epsilon) N'(\epsilon) \left(\frac{Um}{2} + \mu_B H \right) \\
 &= N(\epsilon_F) \left[1 - \frac{\pi^2}{6} (k_B T)^2 \Lambda_1(\epsilon_F)^2 \right] \left(\frac{Um}{2} + \mu_B H \right). \tag{6.22}
 \end{aligned}$$

¹The Stoner instability is a simplification of the quantum phase transition. In most cases, a discontinuous phase transition originates in the band structure or in fluctuation effects, which have been ignored here, cf. D. Belitz and T.R. Kirkpatrick, Phys. Rev. Lett. **89**, 247202 (2002).

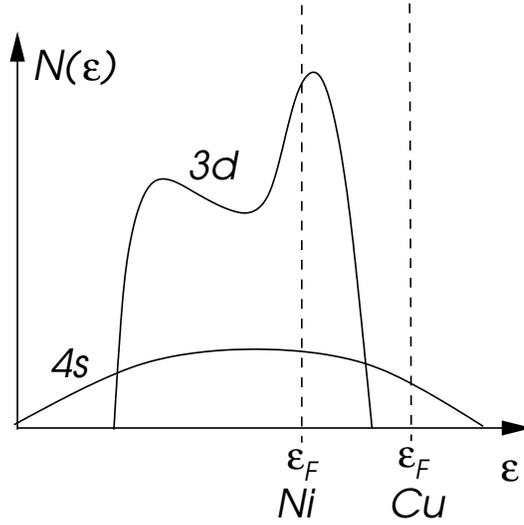


Figure 6.4: The position of the Fermi energy of Cu and Ni, respectively.

Solved for m , this yields

$$M = \mu_B m = \frac{\chi_0(T)}{1 - U\chi_0(T)/2\mu_B^2} H, \quad (6.23)$$

and, consequently,

$$\chi = \frac{M}{H} = \frac{\chi_0(T)}{1 - U\chi_0(T)/2\mu_B^2} m, \quad (6.24)$$

where

$$\chi_0(T) = \mu_B^2 N(\epsilon_F) \left[1 - \frac{\pi^2}{6} (k_B T)^2 \Lambda_1(\epsilon_F)^2 \right]. \quad (6.25)$$

We see, that the denominator of the susceptibility $\chi(T)$ vanishes exactly if the Stoner instability criterion is fulfilled, i.e., the diverging susceptibility indicates the instability and

$$\chi(T) \approx \frac{\chi_0(T_C)}{\frac{T_C^2}{T^2} - 1}. \quad (6.26)$$

Note that for $T \rightarrow T_{C+}$ the susceptibility diverges as $\chi(T) \propto |T_C - T|^{-1}$ corresponding to the mean field behavior (mean field coefficient $\gamma = 1$).

6.2 General spin susceptibility and magnetic instabilities

The ferromagnetic state is characterized by a uniform magnetization. There are, however, magnetically ordered states which do not feature a net magnetization. Examples are spin density waves, antiferromagnets and spin spiral states. Here, we analyze general instability conditions.

6.2.1 General dynamic spin susceptibility

We consider a magnetic field, oscillating in time and modulated space like

$$\mathbf{H}(\mathbf{r}, t) = \mathbf{H} e^{i\mathbf{q}\cdot\mathbf{r} - i\omega t} e^{\eta t}, \quad (6.27)$$

and calculate the resulting magnetization, for the corresponding Fourier component./ We proceed analogously as in chapter 3. We define the spin density operator $\hat{\mathbf{S}}(\mathbf{r})$,

$$\hat{\mathbf{S}}(\mathbf{r}) = \frac{\hbar}{2} \sum_{s,s'} \hat{\Psi}_s^\dagger(\mathbf{r}) \boldsymbol{\sigma}_{ss'} \hat{\Psi}_{s'}(\mathbf{r}) \quad (6.28)$$

in real space, and

$$\widehat{\mathbf{S}}_{\mathbf{q}} = \int d^3r \widehat{\mathbf{S}}(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} = \frac{\hbar}{2\Omega} \sum_{\mathbf{k}, s, s'} c_{\mathbf{k}+\mathbf{q}, s}^\dagger \boldsymbol{\sigma}_{ss'} c_{\mathbf{k}, s'} = \frac{1}{\Omega} \sum_{\mathbf{k}} \widehat{\mathbf{S}}_{\mathbf{k}, \mathbf{q}} \quad (6.29)$$

in momentum space. The Hamiltonian of the electronic system with contact interaction is given by

$$\mathcal{H} = \sum_{\mathbf{k}, s} \epsilon_{\mathbf{k}} \widehat{c}_{\mathbf{k}s}^\dagger \widehat{c}_{\mathbf{k}s} - \underbrace{\frac{g\mu_B}{\hbar} \int d^3r \mathbf{H}(\mathbf{r}, t) \cdot \widehat{\mathbf{S}}(\mathbf{r})}_{=\mathcal{H}_Z} + U \int d^3r \widehat{\rho}_\uparrow(\mathbf{r}) \widehat{\rho}_\downarrow(\mathbf{r}). \quad (6.30)$$

We investigate a magnetic field in the x - y -plane, $\mathbf{H} \rightarrow H^+(\mathbf{q}, \omega)$ with

$$\mathcal{H}_Z = -\frac{g\mu_B}{\hbar\Omega} \sum_{\mathbf{k}} H^+(\mathbf{q}, \omega) e^{i\mathbf{q}\cdot\mathbf{r}} \widehat{S}_{\mathbf{k}, -\mathbf{q}}^- e^{-i\omega t + \eta t} + h.c.. \quad (6.31)$$

We ignore the 'h.c.' in \mathcal{H}_Z in the following. Using the equations of motion, analogous to Sect. 3.2, we determine the induced magnetization, first without the interaction term, i.e.,

$$i\hbar \frac{\partial}{\partial t} \widehat{S}_{\mathbf{k}, \mathbf{q}}^+ = [\mathcal{H}, \widehat{S}_{\mathbf{k}, \mathbf{q}}^+]. \quad (6.32)$$

Thus, we obtain for the given Fourier component,

$$i\hbar \frac{\partial}{\partial t} \widehat{S}_{\mathbf{k}, \mathbf{q}}^+(t)_{\mathbf{k}, \mathbf{q}} = (\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}}) \widehat{S}_{\mathbf{k}, \mathbf{q}}^+(t) - g\hbar\mu_B (\widehat{c}_{\mathbf{k}+\mathbf{q}\uparrow}^\dagger \widehat{c}_{\mathbf{k}+\mathbf{q}\uparrow} - \widehat{c}_{\mathbf{k}\downarrow}^\dagger \widehat{c}_{\mathbf{k}\downarrow}) H^+(\mathbf{q}, \omega) e^{-i\omega t}. \quad (6.33)$$

Taking the temporal Fourier transform and the thermal average we obtain,

$$(\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}} - \hbar\omega + i\hbar\eta) \langle S_{\mathbf{k}, \mathbf{q}}^+ \rangle = -g\hbar\mu_B (n_{\mathbf{k}+\mathbf{q}\uparrow} - n_{\mathbf{k}\downarrow}) H^+(\mathbf{q}, \omega), \quad (6.34)$$

which then leads to the induced spin density (magnetization),

$$\langle S_{\text{ind}}^+(\mathbf{q}, \omega) \rangle = \frac{1}{\Omega} \sum_{\mathbf{k}} \langle S_{\mathbf{k}, \mathbf{q}}^+ \rangle = \frac{\hbar}{\mu_B} \chi_0(\mathbf{q}, \omega) H^+(\mathbf{q}, \omega), \quad (6.35)$$

with

$$\chi_0(\mathbf{q}, \omega) = -\frac{g\mu_B^2}{\Omega} \sum_{\mathbf{k}} \frac{n_{\mathbf{k}+\mathbf{q}\uparrow} - n_{\mathbf{k}\downarrow}}{\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}} - \hbar\omega + i\hbar\eta}. \quad (6.36)$$

Note that the form of $\chi_0(\mathbf{q}, \omega)$ is analogous to the Lindhard function Eq. (3.34). The form Eq. (6.35) determines the induced spin density in linear response approximation.

Albeit, we have not yet included the effects of the interaction. Analogously to the charge density in Sect.3.2., the induced spin density generates an effective field: the induced spin polarization appears in the exchange interaction as an effective magnetic field. We rewrite the contact interaction in Eq. (6.30) in the form

$$\begin{aligned} U \int d^3r \rho_\uparrow(\mathbf{r}) \rho_\downarrow(\mathbf{r}) &= \frac{U}{\Omega} \sum_{\mathbf{k}, \mathbf{k}', \mathbf{q}} \widehat{c}_{\mathbf{k}+\mathbf{q}\uparrow}^\dagger \widehat{c}_{\mathbf{k}\uparrow} \widehat{c}_{\mathbf{k}'-\mathbf{q}\downarrow}^\dagger \widehat{c}_{\mathbf{k}'\downarrow} \\ &= -\frac{U}{\Omega} \sum_{\mathbf{k}, \mathbf{k}', \mathbf{q}} \widehat{c}_{\mathbf{k}+\mathbf{q}\uparrow}^\dagger \widehat{c}_{\mathbf{k}\downarrow} \widehat{c}_{\mathbf{k}'-\mathbf{q}\downarrow}^\dagger \widehat{c}_{\mathbf{k}'\uparrow} = -\frac{U}{\Omega\hbar^2} \sum_{\mathbf{q}} \widehat{S}_{\mathbf{q}}^+ \widehat{S}_{-\mathbf{q}}^-. \end{aligned} \quad (6.37)$$

We may now consider the induced spin polarization $\widehat{S}_{\mathbf{q}}^+ \rightarrow \langle S_{\text{ind}}^+(\mathbf{q}, \omega) \rangle$ defining an effective magnetic field due to induced spin polarization coupling to the electron spins through interaction term:

$$-\frac{g\mu_B}{\hbar} H_{\text{ind}}^+(\mathbf{q}, \omega) \widehat{S}_{-\mathbf{q}}^- = -\frac{U}{\hbar^2} \langle S^+(\mathbf{q}, \omega) \rangle \widehat{S}_{-\mathbf{q}}^-, \quad (6.38)$$

the effective magnetic field H_{ind}^+ is then given by

$$H_{\text{ind}}^+ = \frac{U}{g\mu_B\hbar} \langle S^+(\mathbf{q}, \omega) \rangle. \quad (6.39)$$

This induced field acts on the spins as well, such that the total response of the spin density on the external field becomes

$$\begin{aligned} M^+(\mathbf{q}, \omega) &= \frac{\mu_B}{\hbar} \langle S^+(\mathbf{q}, \omega) \rangle = \chi_0(\mathbf{q}, \omega) \{H^+(\mathbf{q}, \omega) + H_{\text{ind}}^+(\mathbf{q}, \omega)\} \\ &= \chi_0(\mathbf{q}, \omega) H^+(\mathbf{q}, \omega) + \chi_0(\mathbf{q}, \omega) \frac{U}{g\mu_B\hbar} \langle S^+(\mathbf{q}, \omega) \rangle. \end{aligned} \quad (6.40)$$

Thus, $M^+(\mathbf{q}, \omega) = \chi(\mathbf{q}, \omega) H^+(\mathbf{q}, \omega)$, with

$$\chi(\mathbf{q}, \omega) = \frac{\chi_0(\mathbf{q}, \omega)}{1 - \frac{U}{2\mu_B^2} \chi_0(\mathbf{q}, \omega)}. \quad (6.41)$$

In analogy to Sect.3.2, this corresponds to the so-called RPA form. This form is valid for all field directions, as long as spin-orbit coupling is neglected and the spin is isotropic.

Looking at the case $\mathbf{q}, \omega \rightarrow 0$ corresponding to a uniform, static external field, we obtain

$$\chi_0(\mathbf{q}, 0) = -\frac{2\mu_B^2}{\Omega} \sum_{\mathbf{k}} \frac{n_{\mathbf{k}+\mathbf{q}\uparrow} - n_{\mathbf{k}\downarrow}}{\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}}} \xrightarrow{\mathbf{q} \rightarrow 0} -\frac{2\mu_B^2}{\Omega} \sum_{\mathbf{k}} \frac{\partial f(\epsilon_{\mathbf{k}})}{\partial \epsilon_{\mathbf{k}}} = \chi_0(T), \quad (6.42)$$

which corresponds to the Pauli susceptibility ($g = 2$). Then, $\chi(T)$ is again cast into the form that we have seen in Eq. (6.24) and describes the instability of the metal with respect to ferromagnetic spin polarization, where the denominator vanishes. It is, however, possible, that $\mathbf{q} = 0$ is not the leading instability, if $\chi_0(\mathbf{q}, 0) > \chi_0(0, 0)$. Then, another form of magnetic order would occur.

6.2.2 Instability with finite wave vector \mathbf{q}

In order to show that, indeed, the Stoner instability does not always prevail among all possible magnetic instabilities, we look first at a simple argument based on the local susceptibility. We define the local magnetic moment along the z -axis, $M(\mathbf{r}) = \mu_B \langle \hat{\rho}_\uparrow(\mathbf{r}) - \hat{\rho}_\downarrow(\mathbf{r}) \rangle$, and observe in linear response the nonlocal relation

$$M(\mathbf{r}) = \int d^3r' \tilde{\chi}_0(\mathbf{r} - \mathbf{r}') H_z(\mathbf{r}'), \quad (6.43)$$

or, in Fourier space,

$$M_{\mathbf{q}} = \chi_0(\mathbf{q}) H_{\mathbf{q}}, \quad (6.44)$$

with

$$\chi_0(\mathbf{q}) = \int d^3r \tilde{\chi}_0(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}}. \quad (6.45)$$

Let us compare $\chi_0(\mathbf{q} = 0)$ with $\overline{\chi(\mathbf{q})} = \tilde{\chi}_0(\mathbf{r} = 0)$, i.e., the uniform and the local susceptibility at $T = 0$. The local susceptibility appears as the average of $\chi_0(\mathbf{q})$ over all \mathbf{q} ,

$$\overline{\chi_0(\mathbf{q})} = \frac{2\mu_B^2}{\Omega^2} \sum_{\mathbf{k}, \mathbf{q}} \frac{n_{\mathbf{k}+\mathbf{q}} - n_{\mathbf{k}}}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}+\mathbf{q}}} = \frac{\mu_B^2}{2} \int d\epsilon N(\epsilon) \int d\epsilon' N(\epsilon') \frac{f(\epsilon) - f(\epsilon')}{\epsilon' - \epsilon}, \quad (6.46)$$

as compared to $\chi_0(\mathbf{q} = 0) = \mu_B^2 N(\epsilon_F)$. The local susceptibility depends on the density of states and the Fermi energy of the system.

A very good qualitative understanding can be obtained already by a very simple form of the density of states,

$$N(\epsilon) = \begin{cases} \frac{1}{D}, & -D \leq \epsilon \leq +D, \\ 0, & |\epsilon| > D, \end{cases} \quad (6.47)$$

i.e., $N(\epsilon)$ for a band in the form of a box with width $2D$ (band width). With this, the integral in Eq. (6.46) is easily evaluated and leads to the ratio,

$$R_0 = \frac{\overline{\chi_0(\mathbf{q})}}{\chi_0(\mathbf{q} = 0)} = \ln \left(\frac{4}{1 - \eta^2} \right) + \eta \ln \left(\frac{1 - \eta}{1 + \eta} \right), \quad (6.48)$$

with $\eta = \epsilon_F/D$ (cf. Fig. 6.5). For small and large band fillings (ϵ_F close to the band edges),

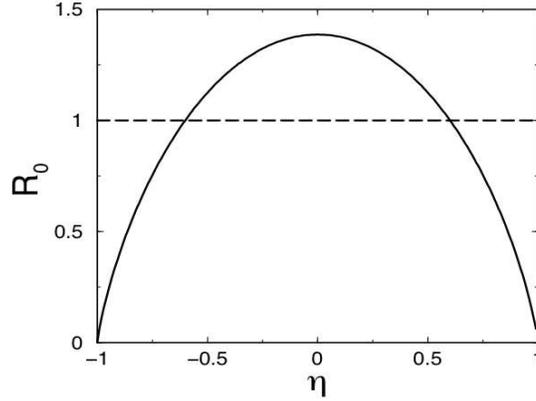


Figure 6.5: R_0 for a box-shaped band with width $\eta = \epsilon_F/D$.

the tendency towards ferromagnetism dominates. If, on the other hand, ϵ_F tends towards the middle of the band, the susceptibility $\chi_0(\mathbf{q})$ will cease to be maximal at $\mathbf{q} = 0$, and magnetic ordering with finite \mathbf{q} becomes more probable.

6.2.3 Influence of the band structure

Magnetic order at finite \mathbf{q} depends strongly on the details of the band structure. The argument with the local susceptibility is nothing more but a general indication on a possible instability at finite \mathbf{q} . A crucial feature for the importance of a given \mathbf{q} is the so-called “nesting” of the Fermi surface. Let us assume that, around certain extended areas of the Fermi surface the energy dispersion satisfies the condition,

$$\xi_{\mathbf{k}+\mathbf{Q}} = -\xi_{\mathbf{k}} \quad (6.49)$$

for a fixed vector \mathbf{Q} and for all \mathbf{k} in close proximity of the Fermi surface with $\xi_{\mathbf{k}} = \epsilon_{\mathbf{k}} - \epsilon_F$. Then, this area will dominate the susceptibility. Let us investigate the static susceptibility $\chi_0(\mathbf{q})$ for $\vec{q} = \vec{Q}$ under the assumption, that Eq.(6.49) holds for all \vec{k} . Then,

$$\begin{aligned} \chi_0(\vec{Q}; T) &= \frac{2\mu_B^2}{\Omega} \sum_{\vec{k}} \frac{n_{\vec{k}+\vec{Q}} - n_{\vec{k}}}{\xi_{\vec{k}} - \xi_{\vec{k}+\vec{Q}}} = \mu_B^2 \int \frac{d^3k}{(2\pi)^3} \frac{f(-\xi_{\vec{k}}) - f(\xi_{\vec{k}})}{\xi_{\vec{k}}} \\ &= \mu_B^2 \int \frac{d^3k}{(2\pi)^3} \frac{\tanh(\xi_{\vec{k}}/2k_B T)}{\xi_{\vec{k}}} = \frac{\mu_B^2}{2} \int d\xi N(\xi) \frac{\tanh(\xi/2k_B T)}{\xi}. \end{aligned} \quad (6.50)$$

In order to evaluate this integral approximately, we realize that the leading contributions comes from the immediate vicinity of the Fermi energy so that we replace $N(\xi)$ by the constant density

of states $N(\epsilon_F)$. Furthermore, the integral is only convergent, if we introduce a cutoff, which we do at $\epsilon_0 \sim D$, half the bandwidth. Thus,

$$\begin{aligned}\chi_0(\vec{Q}, 0; T) &\approx -\mu_B^2 N(\epsilon_F) \int_0^{\epsilon_0} d\xi \frac{\tanh(\xi/2k_B T)}{\xi} \\ &= \mu_B^2 N(\epsilon_F) \left\{ \ln\left(\frac{\epsilon_0}{2k_B T}\right) + \ln\left(\frac{4e^\gamma}{\pi}\right) \right\} \approx \mu_B^2 N(\epsilon_F) \ln\left(\frac{1.14\epsilon_0}{2k_B T}\right),\end{aligned}\quad (6.51)$$

where we assume $\epsilon_0 \ll k_B T$, and where $\gamma = 0.57777$ is Euler's constant. The non-renormalized susceptibility diverges logarithmically at low temperatures. By inserting the generalized Stoner relation, we find the instability criterion [divergence of $\chi(\vec{Q}, 0; T)$ in Eq. (6.41)] in the form

$$0 = 1 - \frac{UN(\epsilon_F)}{2} \ln\left(\frac{1.14\epsilon_0}{2k_B T_c}\right),\quad (6.52)$$

and the critical temperature

$$k_B T_c = 1.14\epsilon_0 e^{-2/UN(\epsilon_F)}.\quad (6.53)$$

Obviously, there is a finite critical temperature for arbitrarily small values of $UN(\epsilon_F) > 0$. The nesting condition results in $\chi_0(\vec{q}, 0; T)$ being maximal for $\vec{q} = \vec{Q}$ and leading to the relevant instability in Eq. (6.41) producing magnetic order with wave vector \mathbf{Q} . We speak here of a *spin density wave*. The spin density has, for example, the form

$$\mathbf{S}(\mathbf{r}) = \hat{z} S \cos(\mathbf{Q} \cdot \mathbf{r}),\quad (6.54)$$

without a uniform component. Such spin density waves are known in low-dimensional systems like organic conductors, or in transition metals like chrome (Cr). In all cases, nesting plays an important role (cf. Fig. 6.6).

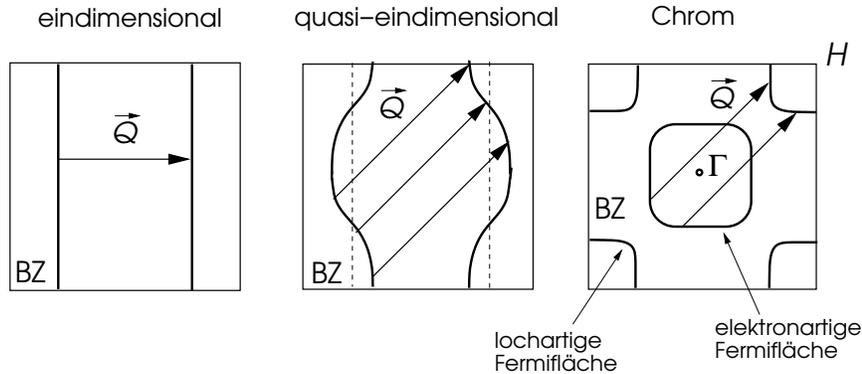


Figure 6.6: Different nesting conditions.

In quasi-one-dimensional electron systems, there is a main direction of motion and two unimportant directions with weak dispersion. In this case nesting is possible as shown in the center panel of Fig.6.6. Chrome is a three-dimensional metal, where nesting occurs between a electron-like Fermi surface around the Γ point and a hole-like Fermi surface at the zone boundary (H point). These Fermi surfaces originate in different bands (right panel in Fig.6.6). Chrome has a cubic body centered crystal structure, where the H -point at $(\pi/a, 0, 0)$ leads to the nesting vector $\mathbf{Q} \parallel (1, 0, 0)$ (and equivalent directions), which is incommensurate with the lattice.

The textbook example of nesting is found from a tight-binding model in a simple cubic lattice with nearest-neighbor hopping and half filling. The band structure is given by

$$\epsilon_{\mathbf{k}} = -2t[\cos k_x a + \cos k_y a + \cos k_z a],\quad (6.55)$$

with the chemical potential $\mu = 0$. Obviously, $\epsilon_{\mathbf{k}+\mathbf{Q}} = -\epsilon_{\mathbf{k}}$ for all \mathbf{k} , if $\mathbf{Q} = \frac{\pi}{a}(1, 1, 1)$. This corresponds to total particle-hole symmetry.

Analogously to the Peierls instability, the spin density wave induces the opening of a gap at the Fermi surface. Thus, the instability is a Fermi surface instability. The gap is confined to the areas of the Fermi surface obeying the nesting condition. Contrarily to the ferromagnetic order, the material can become insulating due to the formation of a spin density wave.

6.3 Stoner excitations

Finally, we discuss the elementary excitations of the ferromagnetic ground state, including particle-hole excitations as well as new, collective modes. We focus on spin excitations, for which we make the Ansatz

$$|\psi_{\mathbf{q}}\rangle = \sum_{\mathbf{k}} f_{\mathbf{k}} \hat{c}_{\mathbf{k}+\mathbf{q},\downarrow}^{\dagger} \hat{c}_{\mathbf{k}\uparrow} |\psi_g\rangle, \quad (6.56)$$

i.e., we extract an electron from the ground state $|\psi_g\rangle$ and replace it by one with an opposite spin. This implies a selection factor $n_{\mathbf{k}\uparrow}(1 - n_{\mathbf{k}+\mathbf{q},\downarrow})$ which takes care of a electron with $(\mathbf{k} \downarrow)$ being available, and no electron with $(\mathbf{k} + \mathbf{q}, \uparrow)$ being present.

We solve the Schrödinger equation

$$\mathcal{H}|\psi_{\mathbf{q}}\rangle = (E_g + \hbar\omega_{\mathbf{q}})|\psi_{\mathbf{q}}\rangle. \quad (6.57)$$

A straightforward calculation shows, that the eigenvalue condition has the form

$$\frac{1}{U} = \frac{1}{\Omega} \sum_{\mathbf{k}} \frac{n_{\mathbf{k}\downarrow} - n_{\mathbf{k}+\mathbf{q}\uparrow}}{\hbar\omega_{\mathbf{q}} - \epsilon_{\mathbf{k}+\mathbf{q}\downarrow} + \epsilon_{\mathbf{k}\uparrow}}, \quad (6.58)$$

corresponding to a root of the denominator of the RPA susceptibility Eq. (6.41). One notices immediately, that a part of the eigenvalues corresponds, in principle, to the continuum of electron-hole excitations with the spectrum

$$\hbar\omega_{\mathbf{q}} = \epsilon_{\mathbf{k}+\mathbf{q},\downarrow} - \epsilon_{\mathbf{k}\uparrow} = \epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}} + U(n_{\uparrow} - n_{\downarrow}), \quad (6.59)$$

where we use the definition $\epsilon_{\mathbf{k}s} = \epsilon_{\mathbf{k}} + Un_{-s}$.

In addition collective excitations exist. One can interpret it as a bound state of an electron and a hole analogous to the exciton. It is easily seen that, in the limit $\mathbf{q} \rightarrow 0$, Eq. (6.58) becomes

$$\frac{1}{U} = \frac{n_{\downarrow} - n_{\uparrow}}{\hbar\omega_0 - U(n_{\uparrow} - n_{\downarrow})}. \quad (6.60)$$

This means, that $\hbar\omega_0 = 0$ is a solution which we will interpret later. We expand the right hand side of Eq. (6.58),

$$\frac{1}{U} = \frac{1}{\Delta\Omega} \sum_{\mathbf{k}} \frac{n_{\mathbf{k}+\mathbf{q}\uparrow} - n_{\mathbf{k}\downarrow}}{1 - \frac{\hbar\omega}{\Delta} + \frac{1}{\Delta}(\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}})}, \quad (6.61)$$

with $\Delta = U(n_{\uparrow} - n_{\downarrow})$ and $\epsilon_{\mathbf{k}+\mathbf{q},\downarrow} - \epsilon_{\mathbf{k}\uparrow} = \epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}} + \Delta$. With Eq. (6.60), we obtain

$$\begin{aligned} 0 &= \frac{U}{\Delta} \sum_{\mathbf{k}} (n_{\mathbf{k}+\mathbf{q}\downarrow} - n_{\mathbf{k}\uparrow}) \left[\{\hbar\omega_{\mathbf{q}} + \epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}}\} - \frac{1}{\Delta} \{\hbar\omega_{\mathbf{q}} + \epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}}\}^2 \right] + \dots \\ &\approx \hbar\omega_{\mathbf{q}} - \frac{U}{\Delta} \sum_{\mathbf{k}} \frac{n_{\mathbf{k}\uparrow} + n_{\mathbf{k}\downarrow}}{2} (\mathbf{q} \cdot \nabla_{\mathbf{k}})^2 \epsilon_{\mathbf{k}} - \frac{U}{\Delta^2} \sum_{\mathbf{k}} (n_{\mathbf{k}\downarrow} - n_{\mathbf{k}\uparrow}) (\mathbf{q} \cdot \nabla_{\mathbf{k}} \epsilon_{\mathbf{k}})^2 + O(q^4) \end{aligned} \quad (6.62)$$

up to second order in \mathbf{q} . For the concrete evaluation we assume a simple parabolic form for the band energies ($\epsilon_{\mathbf{k}} = \hbar^2 \mathbf{k}^2 / 2m^*$), and a weak magnetization $n_{\uparrow} - n_{\downarrow} \ll n_0$. Then,

$$\hbar\omega_{\mathbf{q}} = \frac{\hbar^2 q^2}{2m^* \Delta} \left\{ Un_0 - \frac{4\epsilon_F}{3} \right\} = vq^2. \quad (6.63)$$

Note, that $v > 0$, since the instability criterion for this case reads $U_c = 4\epsilon_F/3n_0 = 2/N(\epsilon_F)$ and

$$\hbar\omega_{\mathbf{q}} = \frac{\hbar^2 q^2}{2m^*} \frac{n_0}{3\sqrt{3}} \left(\frac{U_c}{U} - 1 \right)^{1/2} \geq 0. \quad (6.64)$$

This collective excitation features a q^2 -dependence and vanishes for $\mathbf{q} \rightarrow 0$. The last point is a consequence of the ferromagnetic state breaking a continuous symmetry, i.e., the rotation symmetry which is broken by the choice of a given direction of magnetization. A uniform rotation of the magnetization does not cost any energy. This corresponds to the so-called *Goldstone theorem*.² Such an infinitesimal rotation is induced by our excitation with $\mathbf{q} = 0$ (global spin rotation), i.e.,

$$\sum_{\mathbf{k}} \hat{c}_{\mathbf{k}\downarrow}^\dagger \hat{c}_{\mathbf{k}\uparrow} = \hat{S}_{tot}^-. \quad (6.65)$$

The elementary excitations have both electron and hole contributions, with an energy gap $\approx \Delta$ at small \mathbf{q} . Thus, the collective excitations, which we call *magnons*, are well defined quasi-particles. They constitute propagating spin waves. When these excitations enter the particle-hole continuum, they are damped in the same way as plasmons (cf. Fig. 6.7). Being a bound state between an electron and a hole, magnons are, like excitons, bosonic quasi-particles.

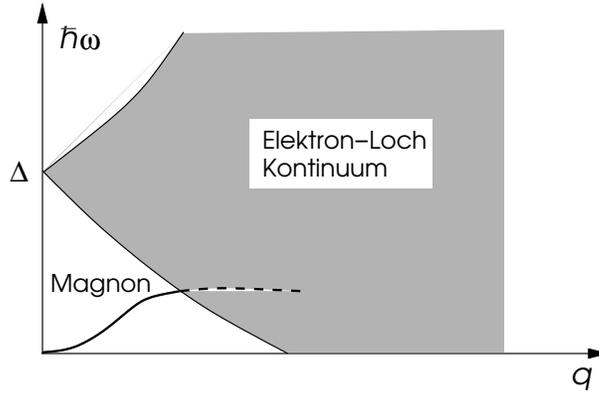


Figure 6.7: Elementary (particle-hole) and collective (magnons) excitations of the Stoner ferromagnet.

²The *Goldstone theorem* states that, in a system with a short-ranged interaction, a phase which is reached by the breaking of a continuous symmetry features collective excitation with arbitrarily small energy, so-called Goldstone modes. These modes have bosonic character. In the case of the Stoner ferromagnet, these modes are the magnons or spin waves.

Chapter 7

Mott insulators and the magnetism of localized moments

Up to now, we have mostly assumed that the interaction between electrons leads to secondary effects. This was, essentially, the message of the Fermi liquid theory, the standard model of condensed matter physics. There, the interactions of course renormalize the properties of a metal, but their description is still possible by using a language of nearly independent fermionic quasiparticles with a few modifications. Even in connection with the magnetism of itinerant electrons, where interactions proved to be crucial, the description in terms of extended Bloch states. Many properties were determined by the band structure of the electrons in the lattice, i.e., the electrons were preferably described in k -space.

However, in this chapter, we will consider situations, where it is less clear whether we should describe the electrons in momentum or in real space. The problem becomes obvious with the following Gedanken experiment: We look at a regular lattice of H-atoms. The lattice constant should be large enough such that the atoms can be considered to be independent for now. In the ground state, each H-atom contains exactly one electron in the $1s$ -state, which is the only atomic orbital we consider at the moment. The transfer of one electron to another atom would cost the relatively high energy of $E(H^+) + E(H^-) - 2E(H) \sim 15\text{eV}$, since it corresponds to an ionization. Therefore, the electrons remain localized on the individual H-atoms and the description of the electron states is obviously best done in real space. The reduction of the lattice constant will gradually increase the overlap of the electron wave functions of neighboring atoms. In analogy to the H_2 molecule, the electrons can now extend on neighboring atoms, but the cost in energy remains that of an "ionization". Thus, transfer processes are only possible virtually, there are not yet itinerant electrons in the sense of a metal.

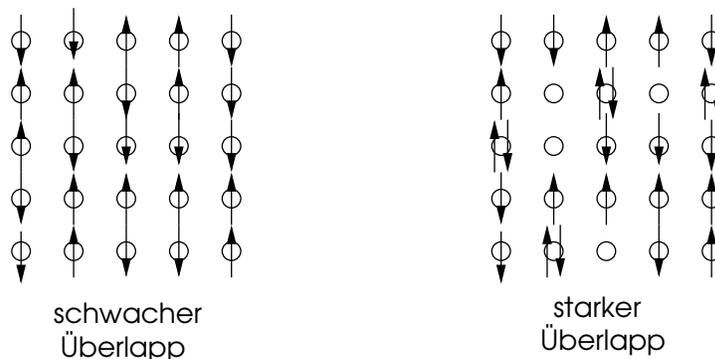


Figure 7.1: Possible states of the electrons in a lattice with weak or strong overlap of the electron wave functions, respectively.

On the other hand, we know the example of the alkali metals, which release their outermost ns -electron into an extended Bloch state and build a metallic (half-filled) band. This would actually work well for the H-atoms for sufficiently small lattice constant too.¹ Obviously, a transition between the two limiting behaviors should exist. This metal-insulator transition, which occurs, if the gain of kinetic energy surpasses the energy costs for the charge transfer. The insulating side is known as a Mott insulator.

While the obviously metallic state is reliably described by the band picture and can be sufficiently well approximated by the previously discussed methods, this point of view becomes obsolete when approaching the metal-insulator transition. According to band theory, a half-filled band must produce a metal, which definitely turns wrong when entering the insulating side of the transition. Unfortunately, no well controlled approximation for the description of this metal-insulator transition exists, since there are no small parameters for a perturbation theory.

Another important aspect is the fact, that in a standard Mott insulator each atom features an electron in the outermost occupied orbital and, hence, a degree of freedom in the form of a localized spin $s = 1/2$, in the simplest case. While charge degrees of freedom (motion of electrons) are frozen at small temperatures, the same does not apply to these spin degrees of freedom. Many interesting magnetic phenomena are produced by the coupling of these spins. Other, more general forms of Mott insulators exist as well, which include more complex forms of localized degrees of freedom, e.g., partially occupied degenerate orbital states.

7.1 Mott transition

First, we investigate the metal-insulator transition. Its description is difficult, since it does not constitute a transition between an ordered and a disordered state in the usual sense. We will, however, use some simple considerations which will allow us to gain some insight into the behavior of such systems.

7.1.1 Hubbard model

We introduce a model, which is based on the tight-binding approximation we have introduced in Chapt. 1. It is inevitable to go back to a description based on a lattice and give up continuity. The model describes the motion of electrons, if their wave functions on neighboring lattice sites only weakly overlap. Furthermore, the Coulomb repulsion, leading to an increase in energy, if a site is doubly occupied, is taken into account. We include this with the lattice analogue of the contact interaction. The model, called *Hubbard model*, has the form

$$\mathcal{H} = -t \sum_{\langle i,j \rangle, s} (\hat{c}_{is}^\dagger \hat{c}_{js} + \text{h.c.}) + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}, \quad (7.1)$$

where we consider hopping between nearest neighbors only, via the matrix element $-t$. Note, that $\hat{c}_{is}^{(\dagger)}$ are real-space field operators on the lattice (site index i) and $\hat{n}_{is} = \hat{c}_{is}^\dagger \hat{c}_{is}$ is the density operator. We focus on half filling, $n = 1$, one electron per site on average.

There are two obvious limiting cases:

- *Insulating atomic limit:* We put $t = 0$. The ground state has exactly one electron on each lattice site. This state is, however, highly degenerate. In fact, the degeneracy is 2^N (number of sites N), since each electron has spin $1/2$, i.e.,

$$|\Phi_{A0}\{s_i\}\rangle = \prod_i \hat{c}_{i,s_i}^\dagger |0\rangle, \quad (7.2)$$

¹In nature, this can only be induced by enormous pressures metallic hydrogen probably exists in the centers of the large gas planets Jupiter and Saturn due to the gravitational pressure.

where the spin configuration $\{s_i\}$ can be chosen arbitrarily. We will deal with the lifting of this degeneracy later. The first excited states feature one lattice site without electron and one doubly occupied site. This state has energy U and its degeneracy is even higher, i.e., $2^{N-2}N(N-1)$. Even higher excited states correspond to more empty and doubly occupied sites. The system is an insulator and the density of states is shown in Fig. 7.2.

- *Metallic band limit:* We set $U = 0$. The electrons are independent and move freely via hopping processes. The band energy is found through a Fourier transform of the Hamiltonian. With

$$\hat{c}_{is} = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \hat{c}_{\mathbf{k}s} e^{i\mathbf{k}\cdot\mathbf{r}_i}, \quad (7.3)$$

we can rewrite

$$-t \sum_{\langle i,j \rangle, s} (\hat{c}_{is}^\dagger \hat{c}_{js} + h.c.) = \sum_{\mathbf{k}, s} \epsilon_{\mathbf{k}} \hat{c}_{\mathbf{k}s}^\dagger \hat{c}_{\mathbf{k}s}, \quad (7.4)$$

where

$$\epsilon_{\mathbf{k}} = -t \sum_{\mathbf{a}} e^{i\mathbf{k}\cdot\mathbf{a}} = -2t \{ \cos k_x a + \cos k_y a + \cos k_z a \}, \quad (7.5)$$

and the sum runs over all vectors \mathbf{a} connecting nearest neighbors. The density of states is also shown in Fig. 7.2. Obviously, this system is metallic, with a unique ground state

$$|\Phi_{B0}\rangle = \prod_{\mathbf{k}} \Theta(-\epsilon_{\mathbf{k}}) \hat{c}_{\mathbf{k}\uparrow}^\dagger \hat{c}_{\mathbf{k}\downarrow}^\dagger |0\rangle. \quad (7.6)$$

Note, that $\epsilon_F = 0$ at half filling, whereas the bandwidth $2D = 12t$.

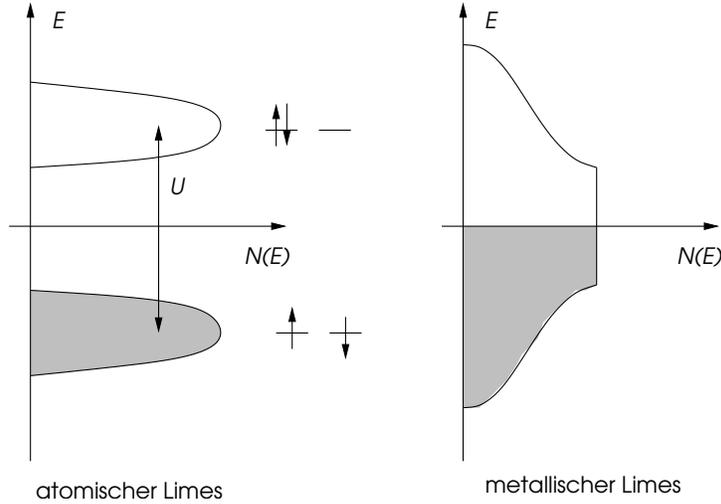


Figure 7.2: Density of states of the Hubbard model in the atomic limit (left) and in the free limit (right).

7.1.2 Insulating state

We consider the two lowest energy sectors for the case $t \ll U$. The ground state sector α has already been defined: one electrons sits on each lattice site. The lowest excited states create the sector β with one empty and one doubly occupied site (cf. Fig. 7.3). With the finite hopping matrix element, the empty (holon) and the doubly occupied (doublon) site become "mobile". A fraction of the degeneracy ($2^{N-2}N(N-1)$) is herewith lifted and the energy obtains a momentum dependence,

$$E_{\mathbf{k},\mathbf{k}'} = U + \epsilon_{\mathbf{k}} + \epsilon_{\mathbf{k}'} > U - 12t. \quad (7.7)$$

Even though ignoring the spin configurations here is a daring approximation, we obtain a qualitatively good picture of the situation.² One notices that, with increasing $|t|$, the two energy sectors approach each other, until they finally overlap. In the left panel Fig.7.2 the holon-doublon excitation spectrum is depicted by two bands, the lower and upper Hubbard bands, where the holon is a hole in the lower and the doublon a particle in the upper Hubbard band. The excitation gap is the gap between the two bands and we may interpret this system as an insulator, called a Mott insulator. (Note, however, that this band structure depends strongly on the correlation effects (e.g. spin correlation) and is not rigid as the band structure of a semiconductor.) The band overlap (closing of the gap) indicates a transition, after which a perturbative treatment is definitely inapplicable. This is, in fact, the metal-insulator transition.

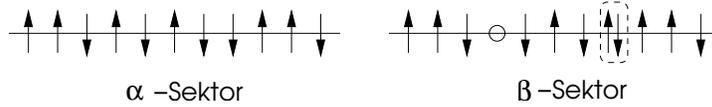


Figure 7.3: Illustration of the two energy sectors, α and β .

7.1.3 The metallic state

On the metallic side, the initial state is better defined since the ground state is a filled Fermi sea without degeneracy. The treatment of the Coulomb repulsion U turns out to become difficult, once we approach the Mott transition, where the electrons suffer a strong impediment in their mobility. In this region, there is no straight-forward way of a perturbative treatment. The so-called Gutzwiller approximation, however, provides a qualitative and very instructive insights into the properties of the strongly correlated electrons.

For this approximation we introduce the following important densities:

- 1: electron density
- s_{\uparrow} : density of the singly occupied lattice sites with spin \uparrow
- s_{\downarrow} : density of the singly occupied lattice sites with spin \downarrow
- d : density of the doubly occupied sites
- h : density of the empty sites

It is easily seen, that $h = d$ and $s_{\uparrow} = s_{\downarrow} = s/2$, as long as no uniform magnetization is present. Note, that d determines the energy contribution of the interaction term to Ud , which we regard as the index of fixed interaction energy sectors. Furthermore,

$$1 = s + 2d \tag{7.8}$$

holds. The view point of the Gutzwiller approximation is based on the renormalization of the probability of the hopping process due to the correlation of the electrons, exceeding restrictions due to the Pauli principle. With this, the importance of the spatial configuration of the electrons is enhanced. In the Gutzwiller approximation, the latter is taken into account statistically by simple probabilities for the occupation of lattice sites.

We fix the density of the doubly occupied sites d and investigate the hopping processes which keep d constant. First, we consider an electron hopping from a singly occupied to an empty

²Note that the motion of an empty site (holon) or doubly occupied site (doublon) is not independent of the spin configuration which is altered through moving these objects. As a consequence, the holon/doublon motion is not entirely free leading to a reduction of the band width. Therefore the band width seen in Fig.7.2 (left panel) is smaller than $2D$, in general. The motion of a single hole was in detail discussed by Brinkman and Rice (Phys. Rev. B 2, 1324 (1970)).

site ($i \rightarrow j$). Hopping probability depends on the availability of the initial configuration. We compare the probability to find this initial state for the correlated (P) and the uncorrelated (P_0) case and write

$$P(\uparrow 0) + P(\downarrow 0) = g_t \{P_0(\uparrow 0) + P_0(\downarrow 0)\}. \quad (7.9)$$

The factor g_t will eventually appear as the renormalization of the hopping probability and, thus, leads to an effective kinetic energy of the system due to correlations. We determine both sides statistically. In the correlated case, the joint probability for i to be singly occupied and j to be empty is obviously

$$P(\uparrow 0) + P(\downarrow 0) = sh = sd = d(1 - 2d). \quad (7.10)$$

where we used Eq.(7.8). In the uncorrelated case (where d is not fixed), we have

$$P_0(\uparrow 0) = n_{i\uparrow}(1 - n_{i\downarrow})(1 - n_{j\uparrow})(1 - n_{j\downarrow}) = \frac{1}{16}. \quad (7.11)$$

The case for \downarrow follows accordingly. In order to collect the total result for hopping processes which keep d constant, we have to do the same calculation for the hopping process $(\uparrow\downarrow, \uparrow) \rightarrow (\uparrow, \uparrow\downarrow)$, which leads to the same result. Processes of the kind $(\uparrow\downarrow, 0) \rightarrow (\uparrow, \downarrow)$ leave the sector of fixed d and are ignored.³ With this, we obtain in all cases the same renormalization factor for the kinetic energy,

$$g_t = 8d(1 - 2d), \quad (7.12)$$

i.e., $t \rightarrow g_t t$. We consider the correlations by treating the electrons as independent but with a renormalized matrix element $g_t t$. The energy in the sector d becomes

$$E(d) = g_t \epsilon_{\text{kin}} + Ud = 8d(1 - 2d)\epsilon_{\text{kin}} + Ud, \quad \epsilon_{\text{kin}} = \frac{1}{N} \int_{-D}^0 d\epsilon N(\epsilon)\epsilon. \quad (7.13)$$

For fixed U and t , we can minimize this w.r.t. d (note that this is not a variational calculation in a strict sense, the resulting energy is not an upper bound to the ground state energy), and find

$$d = \frac{1}{4} \left(1 - \frac{U}{U_c}\right) \quad \text{und} \quad g_t = 1 - \left(\frac{U}{U_c}\right)^2, \quad (7.14)$$

with the critical value

$$U_c = 8|\epsilon_{\text{kin}}| \approx 25t \sim 4D. \quad (7.15)$$

For $u \geq U_c$, double occupancy and, thus, hopping is completely suppressed, i.e., electrons become localized. This observation by Brinkman and Rice [Phys. Rev. B **2**, 4302 (1970)] provides a qualitative description of the metal-insulator transition to a Mott insulator, but takes into account only local correlations, while correlations between different lattice sites are not considered. Moreover, correlations between the spin degrees of freedom are entirely neglected. The charge excitations contain contributions between different energy scales: (1) a metallic part, described via the renormalized effective Hamiltonian

$$\mathcal{H}_{\text{ren}} = \sum_{\mathbf{k}, s} g_t \epsilon_{\mathbf{k}} \hat{c}_{\mathbf{k}s}^\dagger \hat{c}_{\mathbf{k}s} + Ud, \quad (7.16)$$

and (2) a part with higher energy, corresponding to charge excitations on the energy scale U , i.e., to excitations raising the number of doubly occupied sites by one (or more).

³This formulation is based on plausible arguments. A more rigorous derivation can be found in the literature, e.g., in D. Vollhardt, Rev. Mod. Phys. **56**, 99 (1984); T. Ogawa et al., Prog. Theor. Phys. **53**, 614 (1975); S. Huber, *Gutzwiller-Approximation to the Hubbard-Model* (Proseminar SS02, <http://www.itp.phys.ethz.ch/proseminar/condmat02>).

We can estimate the contribution to the metallic conduction. Since in the tight-binding description the current operator contains the hopping matrix element and is thus subject to the same renormalization as the kinetic energy, we obtain

$$\sigma_1(\omega) = \frac{\omega_p^{*2}}{4} \delta(\omega) + \sigma_1^{\text{high energy}}(\omega), \quad (7.17)$$

where we have used Eq. (5.8) for a perfect conductor (no residual resistivity in a perfect lattice). There is a high-energy part which we do not specify here. The plasma frequency is renormalized, $\omega_p^{*2} = g_t \omega_p^2$, such that the f -sum rule in Eq. (5.9) yields

$$I = \int_0^\infty d\omega \sigma_1(\omega) = \frac{\omega_p^2}{8} g_t + I_{\text{high energy}} = \frac{\omega_p^2}{8}. \quad (7.18)$$

For $U \rightarrow U_c$, the coherent metallic part becomes weaker and weaker,

$$\frac{\omega_p^2}{8} g_t = \left\{ 1 - \left(\frac{U}{U_c} \right)^2 \right\} \frac{\omega_p^2}{8}. \quad (7.19)$$

According to the f -sum rule, the lost weight must gradually be transferred to the “high-energy” contribution.

7.1.4 Fermi liquid properties of the metallic state

The just discussed approximation allows us to discuss a few Fermi liquid properties of the metallic state close to metal-insulator transition in a simplified way. Let us investigate the momentum distribution. According to the above definition,

$$\epsilon_{\text{kin}} = \sum_{\mathbf{k} \in \text{FS}} \epsilon_{\mathbf{k}}, \quad (7.20)$$

where the sum runs over all \mathbf{k} in the Fermi sea (FS). One can show within the above approximation, that the distribution is a constant within (n_{in}) and outside (n_{out}) the Fermi surface for finite U , such that, for \mathbf{k} in the first Brillouin zone,

$$\frac{1}{2} = \frac{1}{N} \sum_{\mathbf{k} \in \text{FS}} n_{\text{in}} + \frac{1}{N} \sum_{\mathbf{k} \notin \text{FS}} n_{\text{out}} = \frac{1}{2} (n_{\text{in}} + n_{\text{out}}) \quad (7.21)$$

and

$$g_t \epsilon_{\text{kin}} = \frac{1}{N} \sum_{\mathbf{k} \in \text{FS}} n_{\text{in}} \epsilon_{\mathbf{k}} + \frac{1}{N} \sum_{\mathbf{k} \notin \text{FS}} n_{\text{out}} \epsilon_{\mathbf{k}}. \quad (7.22)$$

Taking into account particle-hole symmetry, i.e.,

$$\sum_{\mathbf{k}} \epsilon_{\mathbf{k}} = \sum_{\mathbf{k} \in \text{FS}} \epsilon_{\mathbf{k}} + \sum_{\mathbf{k} \notin \text{FS}} \epsilon_{\mathbf{k}} = 0, \quad (7.23)$$

we are able to determine n_{in} and n_{out} ,

$$\left. \begin{array}{l} n_{\text{in}} + n_{\text{out}} = 1 \\ n_{\text{in}} - n_{\text{out}} = g_t \end{array} \right\} \Rightarrow n_{\text{in}} = \frac{1}{2}(1 + g_t), \quad n_{\text{out}} = \frac{1}{2}(1 - g_t). \quad (7.24)$$

With this, the jump in the distribution at the Fermi energy is equal to g_t , which, as previously, corresponds to the quasi-particle weight (cf. Fig. 7.4). For $U \rightarrow U_c$ it vanishes, i.e., the quasiparticles cease to exist for $U = U_c$.

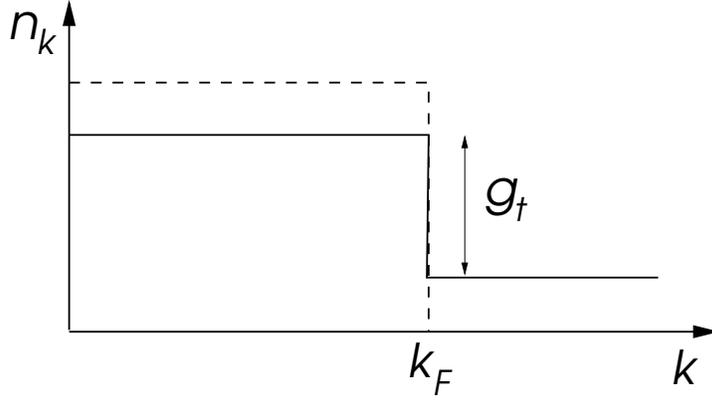


Figure 7.4: The distribution function in the Gutzwiller approximation, displaying the jump at the Fermi energy.

Without going into the details of the calculation, we provide a few Fermi liquid parameters. It is easy to see that the effective mass

$$\frac{m}{m^*} = g_t, \quad (7.25)$$

and thus

$$F_1^s = 3(g_t^{-1} - 1) = \frac{3U^2}{U_c^2 - U^2}, \quad (7.26)$$

where $t = 1/2m$ and the density of states $N(\epsilon_F)^* = N(\epsilon_F)g_t^{-1}$. Furthermore,

$$F_0^a = -\frac{UN(\epsilon_F)}{4} \frac{2U_c + U}{(U + U_c)^2} U_c, \quad \Rightarrow \quad \chi = \frac{\mu_B^2 N(\epsilon_F)^*}{1 + F_0^a}, \quad (7.27)$$

$$F_0^s = \frac{UN(\epsilon_F)}{4} \frac{2U_c - U}{(U - U_c)^2} U_c, \quad \Rightarrow \quad \kappa = \frac{N(\epsilon_F)^*}{n^2(1 + F_0^s)}. \quad (7.28)$$

It follows, that the compressibility κ vanishes for $U \rightarrow U_c$ as expected, since it becomes more and more difficult to compress the electrons or to add more electrons, respectively. The insulator is, of course, incompressible. The spin susceptibility diverges because of the diverging density of states $N(\epsilon_F)^*$. This indicates, that local spins form, which exist as completely independent degrees of freedom at $U = U_c$. Only the antiferromagnetic correlation between the spins would lead to a renormalization, which turns χ finite. This correlation is, as mentioned above, neglected in the Gutzwiller approximation. The effective mass diverges and shows that the quasiparticles are more and more localized close to the transition, since the occupation of a lattice site is getting more rigidly fixed to 1.⁴

As a last remark, it turns out that the Gutzwiller approximation is well suited to describe the strongly correlated Fermi liquid ³He [cf., e.g., D. Vollhardt, Rev. Mod. Phys. **56**, 99 (1984)].

⁴This can be observed within the Gutzwiller approximation in the form of local fluctuations of the particle number. For this, we introduce the density matrix of the electron states on an arbitrary lattice site,

$$\hat{\rho} = h|0\rangle\langle 0| + d|\uparrow\downarrow\rangle\langle\uparrow\downarrow| + \frac{s}{2}\{|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|\}, \quad (7.29)$$

from which we deduce the variance of the occupation number,

$$\langle n^2 \rangle - \langle n \rangle^2 = \langle n^2 \rangle - 1 = \text{tr}(\hat{\rho}n^2) - 1 = 4d + s - 1 = 2d. \quad (7.30)$$

The deviation from single occupation vanishes with d , i.e., with the approach of the metal-insulator transition. Note that the dissipation-fluctuation theorem connects $\langle n^2 \rangle - \langle n \rangle^2$ to the compressibility.

7.2 The Mott insulator as a quantum spin system

One of the most important characteristics of the Mott insulator is the presence of spin degrees of freedom after the freezing of the charge. This is one of the most profound features distinguishing a Mott insulator from a band insulator. In our simple discussion, we have seen that the atomic limit of the Mott insulator provides us with a highly degenerate ground state, where a spin-1/2 degree of freedom is present on each lattice site. We lift this degeneracy by taking into account the kinetic energy term \mathcal{H}_{kin} ($t \ll U$). In this way new physics appears on a low-energy scale, which can be described by an effective spin Hamiltonian. Prominent examples for such spin systems are transition-metal oxides like the cuprates La_2CuO_4 , SrCu_2O_3 or vanadates CaV_4O_9 , NaV_2O_5 .

7.2.1 The effective Hamiltonian

In order to employ our perturbative considerations, it is sufficient to observe the spins of two neighboring lattice sites and to consider perturbation theory for discrete degenerate states. Here, this is preferably done in real space. There are 4 degenerate configurations, $\{|\uparrow, \uparrow\rangle, |\uparrow, \downarrow\rangle, |\downarrow, \uparrow\rangle, |\downarrow, \downarrow\rangle\}$. The application of \mathcal{H}_{kin} yields

$$\mathcal{H}_{\text{kin}}|\uparrow, \uparrow\rangle = \mathcal{H}_{\text{kin}}|\downarrow, \downarrow\rangle = 0, \quad (7.31)$$

and

$$\mathcal{H}_{\text{kin}}|\uparrow, \downarrow\rangle = -\mathcal{H}_{\text{kin}}|\downarrow, \uparrow\rangle = -t|\uparrow\downarrow, 0\rangle - t|0, \uparrow\downarrow\rangle, \quad (7.32)$$

where, in the last two cases, the resulting states have an energy higher by U and lie outside the ground state sector. Thus, it becomes clear that we have to proceed to second order perturbation, where the states of higher energy will appear only virtually (cf. Fig. 7.5). We obtain the matrix

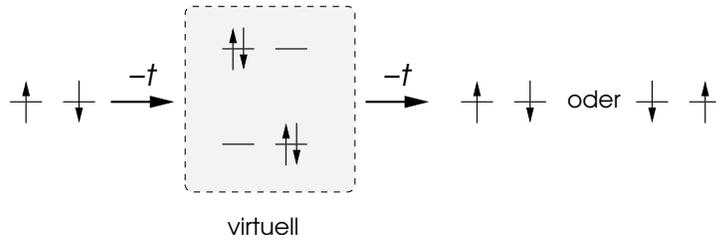


Figure 7.5: Illustration of the origin of the superexchange.

elements

$$M_{s_1, s_2; s'_1, s'_2} = - \sum_n \langle s_1, s_2 | \mathcal{H}_{\text{kin}} | n \rangle \frac{1}{\langle n | \mathcal{H}_{\text{Coul}} | n \rangle} \langle n | \mathcal{H}_{\text{kin}} | s'_1, s'_2 \rangle, \quad (7.33)$$

where $|n\rangle = |\uparrow\downarrow, 0\rangle$ or $|0, \uparrow\downarrow\rangle$, such that the denominator is always U . We end up with

$$M_{\uparrow\downarrow; \uparrow\downarrow} = M_{\downarrow\uparrow; \downarrow\uparrow} = -M_{\uparrow\downarrow; \downarrow\uparrow} = -M_{\downarrow\uparrow; \uparrow\downarrow} = -\frac{2t^2}{U}. \quad (7.34)$$

Note that the signs originates from the anticommutation properties of the Fermion operators. In the subspace $\{|\uparrow, \downarrow\rangle, |\downarrow, \uparrow\rangle\}$ we find the eigenstates of the respective secular equations,

$$\begin{aligned} \frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle) \quad E &= 0, \\ \frac{1}{\sqrt{2}}(|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle) \quad E &= -\frac{4t^2}{U}. \end{aligned} \quad (7.35)$$

Since the states $|\uparrow, \uparrow\rangle$ and $|\downarrow, \downarrow\rangle$ have energy $E = 0$, the sector with total spin $S = 1$ is degenerate (spin triplet). The spin sector $S = 0$ with the energy $-4t^2/U$ is the ground state (spin singlet).

An effective Hamiltonian with the same energy spectrum for the spin configurations can be written with the help of the spin operators $\widehat{\mathbf{S}}_1$ and $\widehat{\mathbf{S}}_2$ on the two lattice sites

$$\mathcal{H}_{\text{eff}} = J \left(\widehat{\mathbf{S}}_1 \cdot \widehat{\mathbf{S}}_2 - \frac{\hbar^2}{4} \right), \quad J = \frac{4t^2}{U\hbar^2} > 0. \quad (7.36)$$

This mechanism of spin-spin coupling is called *superexchange* and introduced by P.W. Anderson [Phys. Rev. **79**, 350 (1950)].

Since this relation is valid between all neighboring lattice sites, we can write the total Hamiltonian as

$$\mathcal{H}_H = J \sum_{\langle i,j \rangle} \widehat{\mathbf{S}}_i \cdot \widehat{\mathbf{S}}_j + \text{const.} \quad (7.37)$$

This model, reduced to spins only, is called *Heisenberg model*. The Hamiltonian is invariant under a global $SU(2)$ spin rotation,

$$U_s(\boldsymbol{\theta}) = e^{-i\widehat{\mathbf{S}} \cdot \boldsymbol{\theta}}, \quad \widehat{\mathbf{S}} = \sum_j \widehat{\mathbf{S}}_j. \quad (7.38)$$

Thus, the total spin is a good quantum number, as we have seen in the two-spin case. The coupling constant is positive and favors an antiparallel alignment of neighboring spins. The ground state is therefore not ferromagnetic.

7.2.2 Mean field approximation of the anti-ferromagnet

There are a few exact results for the Heisenberg model, but not even the ground state energy can be calculated exactly (except in the case of the one-dimensional spin chain which can be solved by means of a Bethe Ansatz). The difficulty lies predominantly in the treatment of quantum fluctuations, i.e., the zero-point motion of coupled spins. It is easiest seen already with two spins, where the ground state is a singlet and maximally entangled. The ground state of all antiferromagnetic systems is a spin singlet ($S - tot = 0$). In the so-called thermodynamic limit ($N \rightarrow \infty$) there is long-ranged anti-ferromagnetic order in the ground state for dimensions $D \geq 2$. Contrarily, the fully polarized ferromagnetic state (ground state for a model with $J < 0$) is known exactly, and as a state with maximal spin quantum number \mathbf{S}^2 it features no quantum fluctuations.

In order to describe the antiferromagnetic state anyway, we apply the mean field approximation again. We can characterize the equilibrium state of the classical Heisenberg model (spins as simple vectors without quantum properties) by splitting the lattice into two sublattices A and B , where each A -site has only B -sites as neighbors, and vice-versa.⁵ On the A -(B -)sublattice, the spins point up (down). This is unique up to a global spin rotations. Note, that this spin configuration doubles the unit cell.

We introduce the respective mean field,

$$\widehat{S}_i^z = \begin{cases} m + (\widehat{S}_i^z - m) & i \in A \\ -m + (\widehat{S}_i^z + m) & i \in B \end{cases}. \quad (7.39)$$

This leads to the mean field Hamiltonian

$$\mathcal{H}_{\text{mf}} = \mathcal{H}_A + \mathcal{H}_B = -Jzm \sum_{i \in A} \widehat{S}_i^z + Jzm \sum_{i \in B} \widehat{S}_i^z + Jz \frac{m^2}{2} N + \dots, \quad (7.40)$$

⁵Lattices which allow for such a separation are called *bipartite*. There are lattices, where this is not possible, e.g., triangular or cubic face centered lattices. There, frustration phenomena appear, a further complication of anti-ferromagnetically coupled systems.

with the coordination number z , the number of nearest neighbors ($z = 6$ in a simple cubic lattice). It is simple to calculate the partition sum of this Hamiltonian,

$$Z = \text{tr} \left\{ e^{-\beta \mathcal{H}_{\text{mf}}} \right\} = \left[\left\{ e^{\beta J m z \hbar / 2} + e^{-\beta J m z \hbar / 2} \right\} e^{-\beta J z m^2 / 2} \right]^N. \quad (7.41)$$

The free energy per spin is consequently given by

$$F(m, T) = -\frac{1}{N} k_B T \ln Z = J z \frac{m^2}{2} - k_B T \ln (2 \cosh(\beta J z m \hbar / 2)). \quad (7.42)$$

At fixed temperature, we minimize the free energy w.r.t m to determine the thermal equilibrium state,⁶ i.e., set $\partial F / \partial m = 0$ and find

$$m = \frac{\hbar}{2} \tanh \left(\frac{J z m \hbar}{2 k_B T} \right). \quad (7.43)$$

This is the self-consistency equation of the mean field theory. It provides a critical temperature T_N (Nel temperature), below which the mean moment m is finite. For $T \rightarrow T_{N-}$, m approaches 0 continuously. Thus, T_N can be found from a linearized self-consistency equation,

$$m = \frac{J z m \hbar^2}{4 k_B T} \Big|_{T=T_N}, \quad (7.44)$$

and thus

$$T_N = \frac{J z \hbar^2}{4 k_B}. \quad (7.45)$$

This means, that T_N scales with the coupling constant and with z . The larger J and the more neighbors are present, the more stable is the ordered state.⁷ For T close to T_N , we can expand the free energy in m ,

$$F(m, T) = F_0 + \frac{J z}{2} \left[\left(1 - \frac{T_N}{T} \right) m^2 + \frac{2}{3 \hbar^2} \left(\frac{T_N}{T} \right)^3 m^4 \dots \right]. \quad (7.46)$$

This is a Landau theory for a phase transition of second order, where a symmetry is spontaneously broken. The breaking of the symmetry (from the high-temperature phase with high symmetry to the low-temperature phase with low symmetry) is described by the order parameter m . The minimization of F w.r.t. m yields (cf. Fig. 7.6)

$$m(T) = \begin{cases} 0, & T > T_N, \\ \frac{\hbar}{2} \sqrt{3(T_N/T - 1)}, & T \leq T_N. \end{cases} \quad (7.47)$$

7.3 Collective modes – spin wave excitations

Besides its favorable properties, the mean field approximation also has a number of insufficiencies. Quantum and some part of thermal fluctuations are neglected, and the insight into the low-energy excitations remains vague. As a matter of fact, as in the case of the ferromagnet, collective excitations exist here. In order to investigate these, we write the Heisenberg model in its spin components, i.e.,

$$\mathcal{H}_H = J \sum_{\langle i, j \rangle} \left\{ \widehat{S}_i^z \widehat{S}_j^z + \frac{1}{2} \left(\widehat{S}_i^+ \widehat{S}_j^- + \widehat{S}_i^- \widehat{S}_j^+ \right) \right\}. \quad (7.48)$$

⁶Actually, a magnetic field pointing into the opposite direction on each site would be another equilibrium variable (next to the temperature). We set it to zero.

⁷At infinite z , the mean field approximation becomes exact.

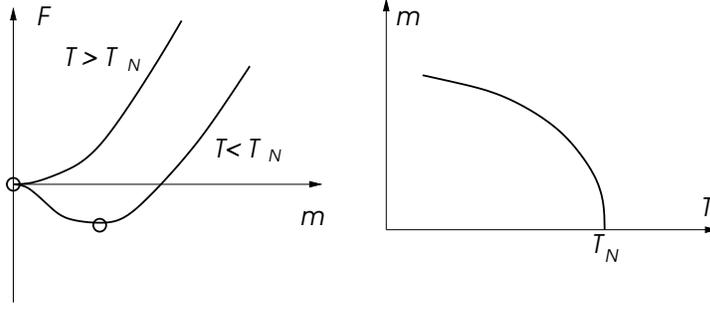


Figure 7.6: The free energy and magnetization of the anti-ferromagnet above and below T_N .

In the ordered state, the moments shall be aligned along the z -axis.

To observe the dynamics of a flipped spin, we apply the operator \widehat{S}_l^- on the ground state $|\Phi_0\rangle$, and determine the spectrum, by solving the resulting eigenvalue equation

$$(\mathcal{H}_H - E_0)\widehat{S}_l^-|\Phi_0\rangle = [\mathcal{H}_H, \widehat{S}_l^-]|\Phi_0\rangle = \hbar\omega\widehat{S}_l^-|\Phi_0\rangle, \quad (7.49)$$

with the ground state energy E_0 . Using the spin-commutation relations

$$[\widehat{S}_j^+, \widehat{S}_j^-] = 2\widehat{S}_j^z\delta_{ij}, \quad (7.50)$$

$$[\widehat{S}_j^z, \widehat{S}_j^\pm] = \pm\widehat{S}_j^\pm\delta_{ij}, \quad (7.51)$$

then yields the equation

$$\left[-J\sum'_j \widehat{S}_j^z \widehat{S}_l^- + J\sum'_j \widehat{S}_j^- \widehat{S}_l^z - \hbar\omega\widehat{S}_l^- \right] |\Phi_0\rangle = 0, \quad (7.52)$$

where \sum'_j runs over all neighbors of l . We decouple this complicated problem by replacing the operators \widehat{S}^z by their mean fields. Therefore, we have to distinguish between A and B sublattices, such that we end up with two equations,

$$\left\{ Jmz\widehat{S}_l^- + Jm\sum_{\mathbf{a}} \widehat{S}_{l+\mathbf{a}}^- - \hbar\omega\widehat{S}_l^- \right\} |\Phi_0\rangle = 0, \quad l \in A, \quad (7.53)$$

$$\left\{ -Jmz\widehat{S}_{l'}^- - Jm\sum_{\mathbf{a}} \widehat{S}_{l'+\mathbf{a}}^- - \hbar\omega\widehat{S}_{l'}^- \right\} |\Phi_0\rangle = 0, \quad l' \in B. \quad (7.54)$$

We introduce the operators

$$\widehat{S}_l^- = \sqrt{\frac{2}{N}} \sum_{\mathbf{q}} \widehat{a}_{\mathbf{q}}^\dagger e^{i\mathbf{q}\cdot\mathbf{r}_l}, \quad \widehat{S}_{l'}^- = \sqrt{\frac{2}{N}} \sum_{\mathbf{q}} \widehat{b}_{\mathbf{q}}^\dagger e^{i\mathbf{q}\cdot\mathbf{r}_{l'}}, \quad (7.55)$$

with $l \in A$ and $l' \in B$, and, vice versa,

$$\widehat{a}_{\mathbf{q}}^\dagger = \sqrt{\frac{2}{N}} \sum_{l \in A} \widehat{S}_l^- e^{-i\mathbf{q}\cdot\mathbf{r}_l}, \quad \widehat{b}_{\mathbf{q}}^\dagger = \sqrt{\frac{2}{N}} \sum_{l' \in B} \widehat{S}_{l'}^- e^{-i\mathbf{q}\cdot\mathbf{r}_{l'}}, \quad (7.56)$$

and insert them into the equation and obtain,

$$\left\{ (Jmz - \hbar\omega) \sum_{l \in A} \widehat{S}_l^- e^{-i\mathbf{q}\cdot\mathbf{r}_l} + Jm \sum_{\mathbf{a}} e^{i\mathbf{q}\cdot\mathbf{a}} \sum_{l' \in B} \widehat{S}_{l'}^- e^{-i\mathbf{q}\cdot\mathbf{r}_{l'}} \right\} |\Phi_0\rangle = 0, \quad (7.57)$$

$$\left\{ (-Jmz - \hbar\omega) \sum_{l' \in B} \widehat{S}_{l'}^- e^{-i\mathbf{q}\cdot\mathbf{r}_{l'}} - Jm \sum_{\mathbf{a}} e^{i\mathbf{q}\cdot\mathbf{a}} \sum_{l \in A} \widehat{S}_l^- e^{-i\mathbf{q}\cdot\mathbf{r}_l} \right\} |\Phi_0\rangle = 0. \quad (7.58)$$

From this follows that

$$\left\{ (Jmz - \hbar\omega)\widehat{a}_{\mathbf{q}}^\dagger + Jm\gamma_{\mathbf{q}}\widehat{b}_{\mathbf{q}}^\dagger \right\} |\Phi_0\rangle = 0, \quad (7.59)$$

$$\left\{ (-Jmz - \hbar\omega)\widehat{b}_{\mathbf{q}}^\dagger - Jm\gamma_{\mathbf{q}}\widehat{a}_{\mathbf{q}}^\dagger \right\} |\Phi_0\rangle = 0, \quad (7.60)$$

with $\gamma_{\mathbf{q}} = \sum_{\mathbf{a}} e^{i\mathbf{q}\cdot\mathbf{a}} = 2(\cos q_x a + \cos q_y a + \cos q_z a)$. This eigenvalue equation is easily solved leading to the description of spin waves in the antiferromagnet. The energy spectrum is given by

$$\hbar\omega_{\mathbf{q}} = \pm Jm\sqrt{z^2 - \gamma_{\mathbf{q}}^2}. \quad (7.61)$$

Note, that only the positive energies make sense.

It is interesting to investigate the limit of small \mathbf{q} ,

$$z^2 - \gamma_{\mathbf{q}}^2 \rightarrow z^2 \mathbf{q}^2 + O(q^4), \quad (7.62)$$

where

$$\hbar\omega_{\mathbf{q}} = Jmz|\mathbf{q}| + \dots \quad (7.63)$$

This means that, in contrast to the ferromagnet, the spin waves of the antiferromagnet have a linear low-energy spectrum (cf. Fig. 7.7). The same applies here if we expand the spectrum around $\mathbf{Q} = (1, 1, 1)\pi/a$ (folding of the Brillouin zone due to the doubling of the unit cell).

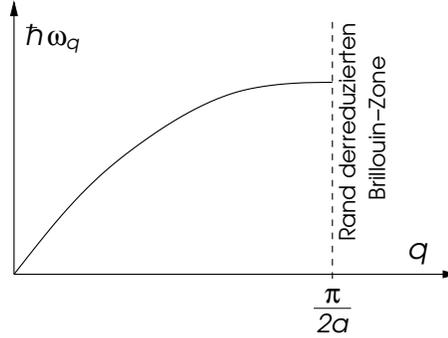


Figure 7.7: Spectrum of the spin waves in the antiferromagnet.

After a suitable normalization, the operators $\widehat{a}_{\mathbf{q}}$ and $\widehat{b}_{\mathbf{q}}$ are of bosonic nature; this comes about since, due to the mean field approximation, the \widehat{S}_i^\pm are bosonic as well,

$$[\widehat{S}_i^+, \widehat{S}_j^-] = 2\widehat{S}_i^z \delta_{ij} \approx \pm 2m\delta_{ij}, \quad (7.64)$$

where the sign depends on the sublattice. The zero-point fluctuations of these bosons yield quantum fluctuations, which reduce the moment m from its mean field value. In a one-dimensional spin chain these fluctuations are strong enough to suppress antiferromagnetically order even for the ground state.

The fact that the spectrum starts at zero has to do with the infinite degeneracy of the ground state. The ordered moments can be turned into any direction globally. This property is known under the name *Goldstone theorem*, which tells that each ordered state that breaks a continuous symmetry has collective excitations with arbitrary small (positive) energies. The linear spectrum is normal for collective excitations of this kind; the quadratic spectrum of the ferromagnet has to do with the fact that the state breaks time-inversion symmetry.

These spin excitations show the difference between a band and a Mott insulator very clearly. While in the band insulator both charge and spin excitations have an energy gap and are inert, the Mott insulator has only gapped charge excitation. However, the spin degrees of freedom for a low-energy sector which can even form gapless excitations as shown just above.