Many-Body Quantum Mechanics

Fabian Hassler

2026

Prof. Dr. Fabian Hassler
Institute for Quantum Information
RWTH Aachen

The notes are largely based on the lecture notes by G. Blatter (ETH Zurich 2006). I would like to sincerely thank Gianni Blatter for providing the script. Further thanks go to Lukas Grunwald, who contributed especially in Chapter 4 and in the appendix.

Please report typos and other errors to F. Hassler. The translation has been done with the help of Mistral Small 4. In case of ambiguities, it might be helpful to consult the German version.

This work is licensed under CC BY-NC 4.0.

Contents

1 Scattering Theory **1**
- 1.1 Wave Packets . . . 1
- 1.2 Lippmann-Schwinger Equation . . . 2
- 1.3 Differential Cross Section . . . 6
- 1.4 Partial Wave Expansion . . . 9
- 1.5 Optical Theorem . . . 12
- 1.6 Born Approximation . . . 14
- 1.7 Eikonal approximation . . . 16
- 1.8 Analytic Properties of the Scattering Matrix . . . 20
- 1.9 Low-energy scattering . . . 22

2 Identical Particles **27**
- 2.1 Description of N particles . . . 27
- 2.2 Description of N identical particles . . . 30
- 2.3 Fermions and Bosons . . . 32
- 2.4 Spin and Statistics . . . 32
- 2.5 Applications . . . 33

3 Atoms **43**
- 3.1 Atom with two electrons . . . 43
- 3.2 Hartree and Hartree-Fock approximations . . . 45
- 3.3 Thomas-Fermi screening . . . 49
- 3.4 Thomas-Fermi atom . . . 51
- 3.5 Hund's rules . . . 55
- 3.6 Zeeman Splitting in a Magnetic Field . . . 61

4 Second Quantization **65**
- 4.1 Problem formulation in first quantization . . . 66
- 4.2 Besetzungszahldarstellung and Fock Space . . . 67
- 4.3 Creation and Annihilation Operators . . . 68
- 4.4 Basis change . . . 71
- 4.5 Operators in second quantization . . . 73

4.6 Field Operators . . . 77
4.7 Correlations in the Fermi Gas . . . 82
4.8 Pair Correlations . . . 85
4.9 Equation of Motion . . . 87

5 Interacting Electron Systems **91**
5.1 Homogeneous Electron Gas . . . 91
5.2 Superconductivity . . . 99

6 Dirac Equation **111**
6.1 Klein-Gordon Equation . . . 111
6.2 Dirac Equation . . . 113
6.3 Conserved Norm . . . 118
6.4 Free Particles . . . 119
6.5 Interpretation . . . 121
6.6 Nonrelativistic limit . . . 125

A Spherical Bessel Functions **131**
A.1 Definition . . . 131
A.2 Expansion of the Plane Wave . . . 132

B Spin and Orbital Species **135**
B.1 Symmetric Group S_N . . . 135
B.2 Example: 2 Spin-1/2 Fermions . . . 138
B.3 N Spin-1/2 Particles . . . 138
B.4 Irreducible Representations of S_N . . . 139
B.5 Position-space functions with defined symmetry . . . 143
B.6 Spin functions with defined symmetry . . . 145
B.7 Spin and Orbit . . . 149

C $SU(n)$ with Young Diagrams **153**
C.1 $SU(n)$ tensors and the Schur-Weyl duality . . . 153
C.2 $SU(2)$ with Young diagrams . . . 155
C.3 $SU(3)$ with Young diagrams . . . 156
C.4 Conjugate representations . . . 158
C.5 Product representations and the quark model . . . 159

References

G. Baym
Lectures on Quantum Mechanics
(Benjamin, Menlo Park, California, 1969)
well-written, compact, modern

J.J. Sakurai
Modern Quantum Mechanics
(Addison Wesley, Reading, 1994) &
well-written, modern

W. Nolting
Theoretische Physik
5. Quantum Mechanics, 1 & 2, with exercises and solutions, (Zimmermann-Neufang, Ulmen, 1992)
introductory, clear, modern structure

A. Messiah
Quantum Mechanics I & II
(North-Holland, Amsterdam, 1976)
complete

L.D. Landau and E.M. Lifshitz
Quantum Mechanics (Nonrelativistic theory)
Vol. 3 of the Course of Theoretical Physics, (Pergamon, Oxford, 1977)
classic, contains everything

F. Schwabl
Quantenmechanik
(Springer, Berlin, 1990)
more recent book, alternative selection of topics

Chapter 1

Scattering Theory

In this chapter we will focus on 3D scattering problems. In particular, we study the scattering of particles by time-independent and radially symmetric scattering potentials. The results can be generalized *mutatis mutandis* to non-radially symmetric potentials. However, the formulas then become much more complicated due to the lack of symmetry without yielding any conceptually new results. A different situation arises when the potential becomes time-dependent or when the particles involved possess internal degrees of freedom, since then the possibility of inelastic scattering or complete absorption exists. These possibilities will not be treated here.

We investigate the unbounded motion of a particle in a scattering potential $V(\boldsymbol{r})$ centered at the coordinate origin and of short range, i.e., with the property that $r\,V(\boldsymbol{r}) \to 0$ for $r \to \infty$. The important example of the long-range Coulomb potential with $V(\boldsymbol{r}) \propto 1/r$ will be considered later in the exercises. It is generally possible that $V(\boldsymbol{r})$ has attractive regions. Bound states with $E < 0$ can then occur. Here we consider scattering states with $E > 0$ and since $V(r \to \infty) \to 0$ these states are asymptotically free.

1.1 Wave Packets

The incoming particle is described by a wave packet. In the following it will be important that the spatial extent of the wave packet ξ is larger than the range of the potential a. Naturally, however, the extent ξ should be smaller than the distance between source/detector and scattering center so that the various stages of the scattering problem are temporally separated. Moreover, we assume that the width in momentum space is sufficiently small so that we may neglect the broadening of the wave packet due to dispersion.

We consider an incoming wave packet

$$\Psi(\boldsymbol{r}; t_0) = \psi_0(\boldsymbol{r}) = \int \frac{d^3q}{(2\pi)^3}\, a_{\boldsymbol{q}}\, e^{i\boldsymbol{q}\cdot\boldsymbol{r}} \tag{1.1}$$

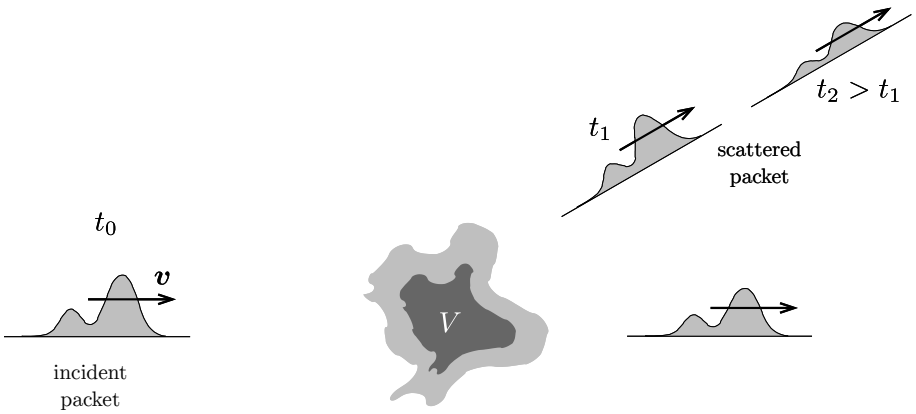


Figure 1.1: Scattering of a wave packet: The incoming packet is scattered with probability $\propto |f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})|^2$ into direction $\Omega_{\boldsymbol{r}} \equiv \hat{\boldsymbol{r}}$. Note that the amplitude of the scattered part of the wave function decays like r^{-1}, cf. (1.23).

which at time $t = t_0$ is centered far to the left of the scatterer, see Figure 1.1. The amplitude $a_{\boldsymbol{q}}$ is concentrated around $\boldsymbol{k}$ so that the packet approaches the scatterer with velocity $\boldsymbol{v} = \hbar\boldsymbol{k}/m$. [1] The time evolution of the wave function $\Psi(\boldsymbol{r}; t)$ determines the signal measured in the detector at a later time $t > t_0$. Our task is thus the determination of $\Psi(\boldsymbol{r}; t > t_0)$ in particular after it has passed through the scattering center.

Let $\Psi_{\boldsymbol{k}}$ be the exact scattering states, i.e., solutions of the eigenvalue equation

$$H\Psi_{\boldsymbol{k}}(\boldsymbol{r}) = E_k \Psi_{\boldsymbol{k}}(\boldsymbol{r}), \tag{1.2}$$

with $H = p^2/2m + V(\boldsymbol{r})$ and the definition $E_k \equiv \hbar^2 k^2/2m \geq 0$[2]. The scattering states $\Psi_{\boldsymbol{k}}$ are complete in the space of extended wave functions and we can write the time evolution of Ψ as

$$\Psi(\boldsymbol{r}; t) = \int \frac{d^3q}{(2\pi)^3} A_{\boldsymbol{q}} \Psi_{\boldsymbol{q}}(\boldsymbol{r}) e^{-iE_q(t-t_0)/\hbar}. \tag{1.3}$$

The expansion coefficients $A_{\boldsymbol{q}}$ will be determined later in Chapter 1.3 and we will show that $A_{\boldsymbol{q}} = a_{\boldsymbol{q}}$. In the expansion (1.3) only scattering states contribute, since we are interested only in the asymptotic expression for $r \to \infty$ and bound states decay exponentially in this region.

1.2 Lippmann-Schwinger Equation

Next we wish to find at least a formal solution of the time-independent Schrödinger equation (1.2) for the scattering states. The boundary conditions are given by the

[1] Concretely one could choose the amplitude $a_{\boldsymbol{q}} = (\xi^2/2\pi)^{3/2} e^{-iq_z z_0 - (\boldsymbol{q}-k\hat{\boldsymbol{z}})^2/2\xi^2}$ with $z_0 \gg a$.

[2] We write v for the norm $|\boldsymbol{v}| = \sqrt{\boldsymbol{v}^2}$ of the vector $\boldsymbol{v}$.

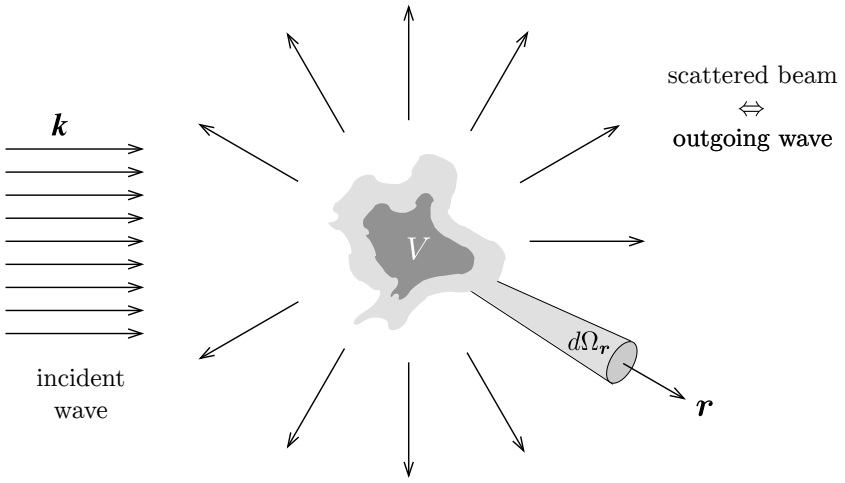


Figure 1.2: Scattering geometry: The incoming wave $\exp(i\boldsymbol{k}\cdot\boldsymbol{r})$ is scattered by the potential $V(\boldsymbol{r})$. The scattered wave asymptotically has the form of a spherical wave $f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})\exp(ikr)/r$ modulated in the solid angle $\Omega_{\boldsymbol{r}}$.

scattering geometry, see Figure 1.2. Hence the wave function $\Psi_{\boldsymbol{k}}$ should asymptotically behave for large distances r like

$$\Psi_{\boldsymbol{k}}(\boldsymbol{r}) \sim e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})\frac{e^{ik_{\mathrm{s}}r}}{r} \tag{1.4}$$

The second term proportional to $\psi_{\mathrm{s}} = e^{ik_{\mathrm{s}}r}/r$ is an outgoing spherical wave, as can be seen from the expression for the probability current

$$\boldsymbol{j}_{\mathrm{s}} = \frac{\hbar}{m}\,\mathrm{Im}(\psi_{\mathrm{s}}^{*}\partial_r\psi_{\mathrm{s}})\hat{\boldsymbol{r}} = \frac{\hbar k_{\mathrm{s}}}{mr^2}\hat{\boldsymbol{r}} = j_{\mathrm{s},r}\hat{\boldsymbol{r}} \tag{1.5}$$

. For elastic scattering processes $k_{\mathrm{s}} = k$ holds. The scattering amplitude $f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})$ depends on the wave number $\boldsymbol{k}$ (or the energy E_k and the incidence direction $\hat{\boldsymbol{k}}$) and the scattering angle $\Omega_{\boldsymbol{r}}$. Note that $\boldsymbol{k}$ is not a conserved quantity and the wave function $\Psi_{\boldsymbol{k}}$ contains momentum components $\boldsymbol{k}'$ with $\boldsymbol{k}' \neq \boldsymbol{k}$.

In a first step we solve the eigenvalue problem (1.2) with the boundary condition (1.4),

$$\underbrace{\left[\frac{\hbar^2}{2m}\nabla^2 + E_k\right]}_{\text{free propagation}}\Psi_{\boldsymbol{k}}(\boldsymbol{r}) = \underbrace{V(\boldsymbol{r})\Psi_{\boldsymbol{k}}(\boldsymbol{r})}_{\text{source}}, \tag{1.6}$$

where we have performed a sensible grouping of terms adapted to the scattering geometry. Equation (1.6) is not an eigenvalue problem in the usual sense (where E_k is unknown), but for each energy E_k solutions can be found. Equation (1.6) is an inhomogeneous partial differential equation, where the source $V(\boldsymbol{r})\Psi_{\boldsymbol{k}}(\boldsymbol{r})$ depends on the solution $\Psi_{\boldsymbol{k}}(\boldsymbol{r})$. Such driven differential equations are usually solved with the aid

of Green's functions. We solve the problem of free propagation for a point source $\delta^{(3)}(\boldsymbol{r})$,

$$(E - H_0)\, G(\boldsymbol{r}; E) = \delta^{(3)}(\boldsymbol{r}) \qquad \text{and boundary conditions for } G, \tag{1.7}$$

where we have introduced the free Hamiltonian $H_0 = p^2/2m$. If we know the solution to (1.7) we can write the solution to (1.6) as

$$\Psi_{\boldsymbol{k}}(\boldsymbol{r}) = \underbrace{e^{i\boldsymbol{k}\cdot\boldsymbol{r}}}_{\text{Solution of the hom. eq.}} + \underbrace{\int d^3r'\, G(\boldsymbol{r}-\boldsymbol{r}'; E_k) V(\boldsymbol{r}') \Psi_{\boldsymbol{k}}(\boldsymbol{r}').}_{\substack{\text{Solution of the inhomogeneous equation} \\ \text{generated by the "source" (=scattering potential)}}} \tag{1.8}$$

Here the solution of the homogeneous equation is precisely the incoming wave and the solution of the inhomogeneous equation is the scattered wave. To prove that the integral equation (1.8) is equivalent to the problem (1.6) plus boundary conditions, we apply the operator $(E_k - H_0)$ to (1.8) with the result

$$\begin{aligned}(E_k - H_0)\Psi_{\boldsymbol{k}} &= \int d^3r' \,[(E_k - H_0) G(\boldsymbol{r}-\boldsymbol{r}'; E_k)] V(\boldsymbol{r}') \Psi'_{\boldsymbol{k}}(\boldsymbol{r}') \\ &= \int d^3r' \,[\delta^{(3)}(\boldsymbol{r}-\boldsymbol{r}')] V(\boldsymbol{r}') \Psi_{\boldsymbol{k}}(\boldsymbol{r}') \\ &= V(\boldsymbol{r}) \Psi_{\boldsymbol{k}}(\boldsymbol{r}).\end{aligned} \tag{1.9}$$

The plane wave in (1.8) yields the first term of the asymptotics (1.4). The second term in (1.8) must therefore correspond to the term $\propto \exp(ikr)/r$, $r \to \infty$ in (1.4). To show this, however, we need the asymptotics of the Green's function $G(\boldsymbol{r}; E)$. We determine the Green's function $G(\boldsymbol{r}; E)$ by the Fourier transform of (1.7),

$$\int d^3r\, e^{-i\boldsymbol{q}\cdot\boldsymbol{r}} (E - H_0) G(\boldsymbol{r}; E) = \int d^3r\, \delta^{(3)}(\boldsymbol{r}) e^{-i\boldsymbol{q}\cdot\boldsymbol{r}}, \tag{1.10}$$

$$\left(E - \frac{\hbar^2 q^2}{2m}\right) G(\boldsymbol{q}; E) = 1,$$

$$G(\boldsymbol{q}; E) = \frac{1}{E - \hbar^2 q^2/2m} = \frac{1}{E - E_q}. \tag{1.11}$$

Upon back-transformation we obtain

$$\begin{aligned}G(\boldsymbol{r}; E) &= \int \frac{d^3q}{(2\pi)^3} \frac{e^{i\boldsymbol{q}\cdot\boldsymbol{r}}}{E - E_q} = -\frac{2m}{\hbar^2} \int_0^\infty \frac{q^2 dq}{4\pi^2} \int_{-1}^1 dz \frac{e^{iqzr}}{q^2 - 2mE/\hbar^2} \\ &= -\frac{2m}{4\pi^2\hbar^2} \int_0^\infty dq\, q^2 \frac{2\sin qr}{qr} \frac{1}{q^2 - 2mE/\hbar^2} \\ &= -\frac{m}{2\pi^2\hbar^2 ir} \int_{-\infty}^\infty dq\, \frac{q e^{iqr}}{q^2 - 2mE/\hbar^2}.\end{aligned} \tag{1.12}$$

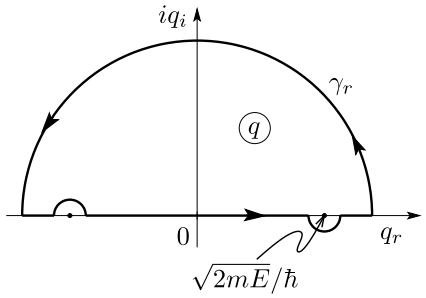


Figure 1.3: Integration path in the complex q plane. By including the pole at $q = +\sqrt{2mE}/\hbar$ we guarantee the asymptotics of an outgoing wave $\propto \exp[i(qr - Et/\hbar)]$. The pole at $q = -\sqrt{2mE}/\hbar$ generates an incoming wave $\propto \exp[-i(qr + Et/\hbar)]$. The choice of the contour in the complex q plane thus fixes the boundary conditions.

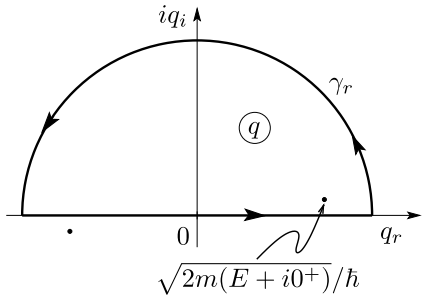


Figure 1.4: By shifting the energy into the complex plane, $E \mapsto E + i0^+$, the poles shift into the complex q plane, $q = \pm\sqrt{2mE}/\hbar \mapsto \pm\sqrt{2m(E+i0^+)}/\hbar$. Integration along $\mathbb{R}$ then automatically takes into account the correct pole in the upper half-plane.

The last integral can be solved with the aid of the residue theorem. Here it is important to obtain an outgoing wave as boundary condition. The integrand of (1.12) has poles at $q = \pm\sqrt{2mE}/\hbar$. It is $r > 0$ and thus we must close the integration path in the upper half-plane with $q_i > 0$ since $i(q_r + iq_i)r = iq_r r - q_i r$ appears in the exponent. To obtain an outgoing wave for $r \to \infty$ we must include the pole at $+\sqrt{2mE}/\hbar$. We therefore choose the integration path γ_r according to Figure 1.3 and obtain the result

$$G(\boldsymbol{r}; E) = -\frac{m}{2\pi\hbar^2}\frac{e^{i\sqrt{2mE}r/\hbar}}{r}. \tag{1.13}$$

Alternatively to the path γ_r we can integrate along the real axis and assign a small imaginary part to the energy, $E \mapsto E + i0^+$. The denominator $q^2 - 2m(E + i0^+)/\hbar^2$ then has poles at $\pm\sqrt{2m(E+i0^+)/\hbar}$, see Figure 1.4, and the integration over $q \in \mathbb{R}$ automatically includes only the pole at $+\sqrt{2mE}/\hbar$. The Green's function $G(\boldsymbol{r}; E + i0^+)$ is called the retarded Green's function G^r because of the property that it produces an outgoing wave. The alternative choice $G^a(\boldsymbol{r}, E) = G(\boldsymbol{r}; E - i0^+)$ $= -(m/2\pi\hbar^2)\exp[-i\sqrt{2mE}r/\hbar]/r$ generates an incoming wave (advanced Green's function). The choice $E = E + i0^+$ thus fixes the boundary conditions. With (1.13) we have found a formal solution of (1.6),

$$\Psi_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + \int d^3r'\, G(\boldsymbol{r} - \boldsymbol{r}'; E_k)V(\boldsymbol{r}')\Psi_{\boldsymbol{k}}(\boldsymbol{r}'), \tag{1.14}$$

$$\text{with} \quad G(\boldsymbol{r}; E_k) = -\frac{m}{2\pi\hbar^2}\frac{e^{ikr}}{r}.$$

Of course we have not really solved (1.6), we have merely rewritten the differential equation, taking into account the boundary conditions, as an integral equation.

Equation (1.14) is the Lippmann-Schwinger equation; it is physically transparent, automatically incorporates the boundary conditions, and lends itself well to the implementation of approximations. Moreover, it has the nice property that the orthogonality and completeness relations [3] of $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ carry over to $\Psi_{\boldsymbol{k}}(\boldsymbol{r})$, i.e. in particular the wave functions for distinct $\boldsymbol{k}$ are orthogonal and $\int d^3r\, \Psi_{\boldsymbol{k}'}(\boldsymbol{r})\Psi_{\boldsymbol{k}}(\boldsymbol{r}) = (2\pi)^3\delta^{(3)}(\boldsymbol{k}-\boldsymbol{k}')$ holds.

Next we examine the far-field regime $r \to \infty$ in (1.14) in order to verify the asymptotic condition (1.4) and to derive an expression for $f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})$ [which, naturally, depends on $V(\boldsymbol{r})$]. For $r \to \infty$ one has $k|\boldsymbol{r}-\boldsymbol{r}'| = kr\sqrt{(\hat{\boldsymbol{r}}-\boldsymbol{r}'/r)^2} = kr[1-2\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'/r+(r'/r)^2]^{1/2} \approx kr - k\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'$. With this, (1.14) can be rearranged to

$$\Psi_{\boldsymbol{k}}(\boldsymbol{r}) \sim e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + \left[-\frac{m}{2\pi\hbar^2}\int d^3r'\, e^{-ik\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'}V(\boldsymbol{r}')\Psi_{\boldsymbol{k}}(\boldsymbol{r}')\right]\frac{e^{ikr}}{r}, \tag{1.15}$$

and, by comparison with (1.4), we obtain the scattering amplitude in the form

$$f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}}) = -\frac{m}{2\pi\hbar^2}\int d^3r'\, e^{-ik\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'}V(\boldsymbol{r}')\Psi_{\boldsymbol{k}}(\boldsymbol{r}'). \tag{1.16}$$

The scattering amplitude has the dimension of a length and depends only on the direction $\hat{\boldsymbol{r}}$, i.e. $\Omega_{\boldsymbol{r}}$, but not on the distance from the scattering center.

1.3 Differential Cross Section

We now wish to relate the formal solution—specifically the scattering amplitude $f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})$—to the differential cross section $d\sigma/d\Omega_{\boldsymbol{r}}$ (a measurable quantity). To this end we recall the discussion of the scattering of a wave packet in Chapter 1.1. For the explicit determination of the time evolution of the wave packet in (1.3) we still need the connection between $a_{\boldsymbol{k}}$ (the initial condition) and $A_{\boldsymbol{k}}$ (the scattering solution). After the detailed discussion of the Lippmann-Schwinger equation this is no longer difficult.

At time t_0 (before the scattering) equations (1.1) and (1.3) must coincide. To obtain a condition on $A_{\boldsymbol{k}}$ we rewrite (1.1) with the help of (1.14) as

$$\Psi(\boldsymbol{r};t_0) = \psi_0(\boldsymbol{r}) = \int\frac{d^3q}{(2\pi)^3}a_{\boldsymbol{q}}\left[\Psi_{\boldsymbol{q}}(\boldsymbol{r}) + \frac{m}{2\pi\hbar^2}\int d^3r'\,\frac{e^{iq|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}V(\boldsymbol{r}')\Psi_{\boldsymbol{q}}(\boldsymbol{r}')\right]. \tag{1.17}$$

[3] The wave functions $\Psi_{\boldsymbol{k}}(\boldsymbol{r})$ are complete only asymptotically, i.e. for $r \to \infty$. To obtain a completeness relation $\delta^{(3)}(\boldsymbol{r}-\boldsymbol{r}') = \int d^3k/(2\pi)^3\Psi^*_{\boldsymbol{k}}(\boldsymbol{r})\Psi_{\boldsymbol{k}}(\boldsymbol{r}') + \sum_n \Psi^*_n(\boldsymbol{r})\Psi_n(\boldsymbol{r}')$ valid over the entire $\mathbb{R}^3$ one must in addition include possible bound states $\Psi_n(\boldsymbol{r})$.

We shall evaluate this relation in the vicinity of the source $\boldsymbol{r}_Q = -r_Q\hat{\boldsymbol{k}}$, $r_Q \gg \xi$, and use it to determine $A_{\boldsymbol{k}}$. In the second term we must compute the expression

$$\int \frac{d^3q}{(2\pi)^3}\, a_{\boldsymbol{q}} e^{iq|\boldsymbol{r}-\boldsymbol{r}'|}\Psi_{\boldsymbol{q}}(\boldsymbol{r}'). \tag{1.18}$$

We assume that $\Psi_{\boldsymbol{q}}$ is smooth over the support of $a_{\boldsymbol{q}}$, i.e. for $|\boldsymbol{q}-\boldsymbol{k}| \lesssim \xi^{-1}$ with ξ the spatial extent of the wave packet, and therefore exhibits no resonances (see also below). Hence $\Psi_{\boldsymbol{q}} \approx \Psi_{\boldsymbol{k}}$, and we may pull the scattering state out of the integral in (1.18). If we further employ the approximation $q = \boldsymbol{q}\cdot\hat{\boldsymbol{q}} \approx \boldsymbol{q}\cdot\hat{\boldsymbol{k}}$, we obtain

$$\begin{aligned}\int \frac{d^3q}{(2\pi)^3} a_{\boldsymbol{q}} e^{iq|\boldsymbol{r}-\boldsymbol{r}'|}\Psi_{\boldsymbol{q}}(\boldsymbol{r}') &\approx \Psi_{\boldsymbol{k}}(\boldsymbol{r}') \int \frac{d^3q}{(2\pi)^3}\, a_{\boldsymbol{q}} e^{i\boldsymbol{q}\cdot(\hat{\boldsymbol{k}}|\boldsymbol{r}-\boldsymbol{r}'|)} \\ &\overset{(1.1)}{=} \Psi_{\boldsymbol{k}}(\boldsymbol{r}')\psi_0(\hat{\boldsymbol{k}}|\boldsymbol{r}-\boldsymbol{r}'|).\end{aligned} \tag{1.19}$$

Here $\psi_0(\hat{\boldsymbol{k}}|\boldsymbol{r}-\boldsymbol{r}'|)$ is the amplitude (of the incoming wave packet) at the initial time at position $\hat{\boldsymbol{k}}|\boldsymbol{r}-\boldsymbol{r}'|$. In the expression (1.17) only positions with $|\boldsymbol{r}'| \lesssim a$ are required. Since at the initial time the particle is located at the source t_0 far from the scatterer, $\boldsymbol{r}_Q$, we have $\psi_0(\hat{\boldsymbol{k}}|\boldsymbol{r}-\boldsymbol{r}'|) \approx \psi_0(\hat{\boldsymbol{k}}\, r_Q) = \psi_0(-\boldsymbol{r}_Q) \approx 0$ and the second term in (1.17) vanishes. The remainder has the form

$$\Psi(\boldsymbol{r};t_0) = \psi_0(\boldsymbol{r}) = \int \frac{d^3q}{(2\pi)^3}\, a_{\boldsymbol{q}}\Psi_{\boldsymbol{q}}(\boldsymbol{r}), \tag{1.20}$$

whence a comparison of coefficients with (1.3) yields $A_{\boldsymbol{q}} = a_{\boldsymbol{q}}$.

An evaluation of $\Psi(\boldsymbol{r};t)$ for large t yields the probability distribution for the detection time. According to (1.3) this is (for $r\to\infty$)

$$\begin{aligned}\Psi(\boldsymbol{r};t) &= \int \frac{d^3q}{(2\pi)^3}\, a_{\boldsymbol{q}}\Psi_{\boldsymbol{q}}(\boldsymbol{r}) e^{-iE_q(t-t_0)/\hbar} \\ &\overset{(1.15)}{\sim} \Psi_0(\boldsymbol{r};t) + \int \frac{d^3q}{(2\pi)^3}\, a_{\boldsymbol{q}} \frac{e^{iqr}}{r} f_{\boldsymbol{q}}(\Omega_{\boldsymbol{r}}) e^{-iE_q(t-t_0)/\hbar};\end{aligned} \tag{1.21}$$

where (we use $E_q \approx -\hbar^2k^2/2m + \boldsymbol{v}\cdot\boldsymbol{q}$)

$$\Psi_0(\boldsymbol{r};t) = \overbrace{\int \frac{d^3q}{(2\pi)^3}\, a_{\boldsymbol{q}}\, e^{i\boldsymbol{q}\cdot\boldsymbol{r}}}^{=\psi_0(\boldsymbol{r}) \text{ by } (1.1)} e^{-iE_q(t-t_0)/\hbar} \approx e^{i\hbar k^2(t-t_0)/2m}\,\psi_0(\boldsymbol{r}-\boldsymbol{v}(t-t_0)) \tag{1.22}$$

(up to an unimportant phase factor) the evolution of the incoming packet without scatterer and dispersion.

With $f_{\boldsymbol{q}}$ smooth for $\boldsymbol{q}\approx\boldsymbol{k}$ (so that $f_{\boldsymbol{q}} \approx f_{\boldsymbol{k}}$ can be pulled out of the integral) and $q \approx \boldsymbol{q}\cdot\hat{\boldsymbol{k}}$ (as before) we obtain

$$\Psi(\boldsymbol{r};t) \overset{t \text{ large}}{\longrightarrow} \underbrace{\Psi_0(\boldsymbol{r};t)}_{\text{unscattered wave packet}} + \underbrace{\frac{f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})}{r}\Psi_0(\hat{\boldsymbol{k}}\, r;t)}_{\text{scattered packet}}. \tag{1.23}$$

The scattering process is sketched in Figure 1.1. According to (1.23) it involves the superposition of the unscattered wave packet and a packet scattered in the direction $\Omega_{\boldsymbol{r}}$. The latter involves the amplitude $\Psi_0(\hat{\boldsymbol{k}}\, r; t)$ of a packet propagating in the forward direction, which must be evaluated at the correct time and at the correct distance. This packet is then multiplied by the angle-dependent scattering amplitude $f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})/r$. Note that the above formula (1.23) is not applicable in two cases:

- if $V(r)$ is long-ranged, e.g. $V(r) = r^{-1}$,
- if the incident energy E_k is resonant.

The scattering amplitude $f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})$ contains the necessary information to compute the differential cross section

$$\frac{d\sigma}{d\Omega_{\boldsymbol{r}}} = \frac{dN(\Omega_{\boldsymbol{r}})}{N_{\text{in}}\, d\Omega_{\boldsymbol{r}}} = \frac{\text{number of particles scattered into } d\Omega}{\text{incident particles per cm}^2 \times d\Omega_{\boldsymbol{r}}} \tag{1.24}$$

which has the unit of an area with $[d\sigma/d\Omega_{\boldsymbol{r}}] = \text{cm}^2, \text{barn};\ 1\,\text{barn} = 10^{-24}\,\text{cm}^2$.

The density of incident particles N_{in} and the number of particles scattered into the solid angle $d\Omega$ $dN(\Omega_{\boldsymbol{r}})$ are given by

$$N_{\text{in}} = \int_{-\infty}^{\infty} dt\, j_{\text{in}}(\boldsymbol{r}_Q; t), \qquad dN(\Omega_{\boldsymbol{r}}) = \int_{-\infty}^{\infty} dt\, j_{\text{s},r}(\boldsymbol{r}; t) \underbrace{r^2\, d\Omega_{\boldsymbol{r}}}_{\text{area of the detector}}.$$

Substituting the expressions for the incident and scattered current densities, [4]

$$\boldsymbol{j}_{\text{in}}(\boldsymbol{r}; t) = \frac{\hbar}{m} \operatorname{Im}\left(\Psi_0^* \boldsymbol{\nabla} \Psi_0\right) = \frac{\hbar \boldsymbol{k}}{m} |\Psi_0(\boldsymbol{r}; t)|^2, \tag{1.25}$$

$$\begin{aligned} j_{\text{s},r}(\boldsymbol{r}; t) &= \frac{\hbar}{m} \operatorname{Im}\left[\frac{f^*}{r} \Psi_0^*(\hat{\boldsymbol{k}} r; t) \partial_r \frac{f}{r} \Psi_0(\hat{\boldsymbol{k}} r; t)\right] \\ &\approx \frac{\hbar k}{m} \frac{|f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})|^2}{r^2} |\Psi_0(\hat{\boldsymbol{k}} r; t)|^2, \end{aligned} \tag{1.26}$$

yields the differential cross section in the form

$$\frac{d\sigma}{d\Omega_{\boldsymbol{r}}} = \frac{|f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})|^2 \int dt |\Psi_0(\hat{\boldsymbol{k}} r; t)|^2}{\int dt\, |\Psi_0(\boldsymbol{r}_Q; t)|^2} = |f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})|^2, \tag{1.27}$$

with $\int dt\, |\Psi_0(\hat{\boldsymbol{k}} r; t)|^2 = \int dt\, |\Psi_0(\boldsymbol{r}_Q; t)|^2 = \int dt\, |\psi_0(\boldsymbol{v} t)|^2$ according to (1.22). Finally we define the total cross section as

$$\sigma = \int d\Omega_{\boldsymbol{r}}\, |f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})|^2\,. \tag{1.28}$$

[4] We use $\partial_r \Psi_0(\hat{\boldsymbol{k}} r; t) = \int \frac{d^3 q}{(2\pi)^3}\, i\boldsymbol{q} \cdot \hat{\boldsymbol{k}}\, a_{\boldsymbol{q}}\, e^{i\boldsymbol{q} \cdot \hat{\boldsymbol{k}} r} e^{-iE_q (t - t_0)/\hbar} \approx ik \Psi_0(\hat{\boldsymbol{k}} r; t)$.

1.4 Partial Wave Expansion

For rotationally symmetric scattering potentials $V(r)$, the Hamiltonian $H = p^2/2m + V(r)$ commutes with the angular momentum operators $\boldsymbol{L}$. Consequently, the angular problem can be separated, and we may decompose the scattering problem according to the irreducible representations of the rotation group. This partial wave decomposition reduces the problem to solving the subproblems in the various angular momentum sectors [5]: We decompose the total Hilbert space $L^2(\mathbb{R}^3)$ according to $L^2(\mathbb{R}^3) = L^2(\mathbb{R}^+) \otimes [\oplus_l \mathcal{H}_l]$ and solve the partial problems in $L^2(\mathbb{R}^+) \otimes \mathcal{H}_l$, where the angular problem is trivial (i.e. already diagonalized). We exploit the rotational symmetry fully by choosing a coordinate system with the z-axis parallel to the incoming beam, $\hat{\boldsymbol{z}} \parallel \boldsymbol{k}$. The entire problem is then rotationally symmetric under rotations about the $z-$ axis. Only quantum numbers with $m = 0$ occur, since the incoming particle cannot carry angular momentum along the direction of motion. The spherical harmonics $Y_{lm}(\theta, \varphi)$ reduce in this case to Legendre polynomials $Y_{l0}(\theta, \varphi) \propto P_l(\cos\theta)$ with the normalization

$$\int_{-1}^{1} dz\, P_l(z) P_{l'}(z) = \frac{2}{2l+1}\delta_{ll'}, \tag{1.29}$$

so that $P_l(1) = 1$. Furthermore, we must expand the incoming wave in eigenfunctions of the angular momentum operator,

$$\begin{aligned} e^{ikr\cos\theta} &= \sum_{l=0}^{\infty} i^l (2l+1) P_l(\cos\theta) j_l(kr) \\ &= \frac{1}{2}\sum_{l=0}^{\infty} i^l (2l+1) P_l(\cos\theta)[h_l^{(1)}(kr) + h_l^{(2)}(kr)], \end{aligned} \tag{1.30}$$

where we have introduced the spherical Bessel functions j_l and Hankel functions $h_l^{(1,2)}$. A brief excursus on spherical Bessel functions and a proof of (1.30) is given in Appendix A.

We can furthermore expand the sought stationary solution $\Psi_{\boldsymbol{k}}(\boldsymbol{r})$ in partial waves,

$$\Psi_{\boldsymbol{k}}(\boldsymbol{r}) = \sum_{l=0}^{\infty} i^l (2l+1) \underbrace{P_l(\cos\theta)}_{\propto Y_{l0}} R_l(r), \tag{1.31}$$

where the factor $i^l(2l+1)$ is a convention explained by comparison with (1.30). Inserting the ansatz (1.31) into the stationary Schrödinger equation $H\Psi_{\boldsymbol{k}} = E_k \Psi_{\boldsymbol{k}}$ yields the radial problem

$$\left[-\frac{p_r^2}{2m} - \frac{\hbar^2 l(l+1)}{2mr^2} + \frac{\hbar^2 k^2}{2m}\right] R_l(r) = V(r) R_l(r); \tag{1.32}$$

[5] The $2l+1$-dimensional Hilbert space $\mathcal{H}_l = \{Y_{lm}\}_{m=-l}^{m=l}$ is spanned by the spherical harmonics for l

with $p_r^2/\hbar^2 = -r^{-1}\partial_r^2 r$ this leads to the differential equation

$$\left[\partial_r^2 - \frac{l(l+1)}{r^2} + k^2\right] r\, R_l(r) = \frac{2m}{\hbar^2} r V(r) R_l(r), \tag{1.33}$$

where $\Psi_{\boldsymbol{k}}$ must satisfy the boundary condition (1.4). For $r \to \infty$ $rV(r) \to 0$ tends to zero and the right-hand side of the equation becomes vanishingly small, so that (1.33) reduces to a Bessel equation, see Appendix A. In the far field ($r \to \infty$) $R_l(r)$ must therefore have the asymptotic form

$$R_l(r) \sim \alpha_l \left[h_l^{(2)}(kr) + S_l\, h_l^{(1)}(kr)\right], \tag{1.34}$$

with

$$h_l^{(1,2)}(\rho) \sim \frac{1}{\rho} \exp[\pm i(\rho - (l+1)\pi/2)], \qquad \rho = kr, \tag{1.35}$$

the two fundamental solutions for $V \equiv 0$, the incoming and outgoing spherical waves in the form of Hankel functions.

Next we must determine the coefficients α_l and S_l, which are functions of k and E_k respectively; for $V \equiv 0$ it is obvious that

$$R_l(r) \;=\; j_l(kr) = \frac{1}{2}\left[h_l^{(2)}(kr) + h_l^{(1)}(kr)\right], \tag{1.36}$$

and hence $\alpha_l = 1/2$ and $S_l = 1$. For $V \not\equiv 0$ the incoming wave $h_l^{(2)}$ remains unchanged, but the outgoing component $h_l^{(1)}$ does change, so that the latter acquires a nontrivial weight $S_l \neq 1$. It is easy to see that for elastic scattering particle-number conservation requires $|S_l(E_k)| = 1$. This follows because in the stationary solution the total radial current density [here we use that $\int dz\, P_l^2(z) = 2/(2l+1)$]

$$\begin{aligned} j_{\mathrm{s},r}(r) &= \frac{\hbar}{m} \operatorname{Im}\left(R_l^*\, \partial_r R_l\right) \\ \downarrow \quad & 2R_l \sim \left[e^{-i(kr-\phi_l)}/kr + S_l\, e^{i(kr-\phi_l)}/kr\right], \quad \phi_l = (l+1)\pi/2 \\ &\sim \frac{\hbar}{4mkr^2}\left[|S_l|^2 - 1\right] \end{aligned} \tag{1.37}$$

must vanish so that equal numbers of particles enter and leave per unit time. We can then express the complex amplitude S_l

$$S_l(E_k) = e^{2i\delta_l(E_k)} \tag{1.38}$$

in terms of the real scattering phase $\delta_l(E_k)$. [6]

[6] In general scattering theory $S_l(E_k)$ is an eigenvalue of the S-matrix. The S-matrix is then unitary by particle-number conservation. Here, on account of the spherical symmetry of the scattering problem and the assumption of purely elastic scattering, we were able to reduce the problem to a one-dimensional subspace.

The scattering phases δ_l describe the scattering problem completely by determining the scattering amplitude $f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})$. To establish this relationship, we bring the solution (1.31) in the asymptotic region where the expression

$$\Psi_{\boldsymbol{k}}(\boldsymbol{r}) \sim \frac{1}{2}\sum_l i^l(2l+1)P_l(\cos\theta)\left[h_l^{(2)}(kr) + e^{2i\delta_l}h_l^{(1)}(kr)\right] \tag{1.39}$$

is valid, with the help of (1.30) into the form

$$\Psi_{\boldsymbol{k}}(\boldsymbol{r}) \sim e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + \underbrace{\frac{1}{2}\sum_l i^l(2l+1)P_l(\cos\theta)\left[e^{2i\delta_l}-1\right]h_l^{(1)}(kr)}_{\sim f_{\boldsymbol{k}}(\theta)e^{ikr}/r}. \tag{1.40}$$

Comparison with (1.4) $[h_l^{(1)} \sim (-i)^{l+1}e^{ikr}/kr]$ yields the scattering amplitude

$$\begin{aligned} f_{\boldsymbol{k}}(\theta) &= \frac{1}{2ik}\sum_l (2l+1)P_l(\cos\theta)\left[e^{2i\delta_l}-1\right] \\ &= \frac{1}{k}\sum_l (2l+1)P_l(\cos\theta)e^{i\delta_l}\sin\delta_l. \end{aligned} \tag{1.41}$$

The expression

$$\frac{\exp(2i\delta_l)-1}{2ik} = \frac{\exp(i\delta_l)\sin(\delta_l)}{k} = f_l \tag{1.42}$$

is called the partial-wave amplitude. The total cross section σ is obtained by an angular integration,

$$\begin{aligned} \sigma &= \int d\Omega_{\boldsymbol{r}}\,|f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})|^2 \\ &= \frac{4\pi}{k^2}\sum_l (2l+1)\sin^2\delta_l = 4\pi\sum_l (2l+1)|f_l|^2. \end{aligned} \tag{1.43}$$

The quantities

$$\sigma_l = \frac{4\pi}{k^2}(2l+1)\sin^2\delta_l \;=\; 4\pi(2l+1)|f_l|^2 \tag{1.44}$$

are called partial cross sections. Evidently $\sigma_l \le 4\pi(2l+1)/k^2$. Note that in the total cross section the various contributions from the angular-momentum sectors simply add up and there is no interference between different partial waves. It is, however, not the case that all quantities are obtained by simply adding the expressions of the partial waves; for example, for $d\sigma/d\Omega_{\boldsymbol{r}}$ one obtains interference terms from the various sectors.

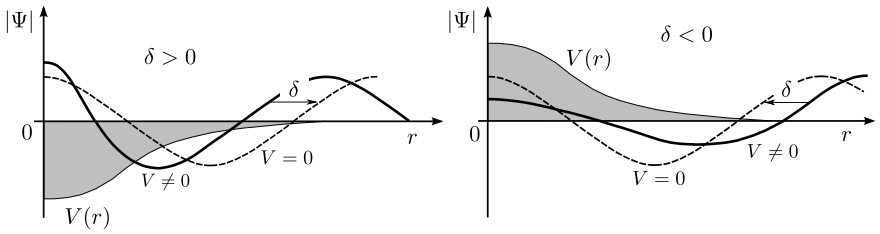


Figure 1.5: Phase shifts $\delta_l(E_k)$: An attractive potential (left) increases the kinetic energy and the wave function oscillates faster, leading to a positive phase shift. Conversely, a repulsive potential (right) slows the oscillation of the wave function and produces a negative $\delta_l(E_k)$.

The phase shift $\exp(2i\delta_l)$ can be interpreted physically in a simple way. The wave function $R_l(r)$ has the asymptotic form

$$R_l(r) \sim \frac{1}{2}\left[h_l^{(2)}(kr) + e^{2i\delta_l} h_l^{(1)}(kr)\right] \sim \frac{e^{i\delta_l}}{2}\left[\frac{(-i)^l e^{i(kr+\delta_l)}}{kr} + \frac{(+i)^l e^{-i(kr+\delta_l)}}{kr}\right]$$
$$\overset{kr\gg\delta_l}{\sim} \frac{e^{i\delta_l}}{2}\left[\frac{(-i)^l e^{i(kr+\delta_l)}}{kr+\delta_l} + \frac{(+i)^l e^{-i(kr+\delta_l)}}{kr+\delta_l}\right] \sim e^{i\delta_l} j_l(kr+\delta_l). \tag{1.45}$$

If we now compare the case

$$V \equiv 0: \qquad R_l(r) = j_l(kr) \tag{1.46}$$

with

$$V \not\equiv 0: \qquad R_l(r) \sim e^{i\delta_l} j_l(kr+\delta_l), \tag{1.47}$$

we see that a positive phase shift $\delta_l > 0$ "pulls the wave function into the potential," whereas a negative phase shift $\delta_l < 0$ "pushes the wave out," see Figure 1.5.

1.5 Optical Theorem

Consider the imaginary part of the forward scattering amplitude $f_{\boldsymbol{k}}(\theta = 0)$ in (1.41),

$$\text{Im } f_{\boldsymbol{k}}(\theta = 0) = \frac{1}{k}\sum_l (2l+1) P_l(\cos\theta) \sin^2\delta_l\Big|_{\theta=0} \tag{1.48}$$
$$= \frac{1}{k}\sum_l (2l+1)\sin^2\delta_l \;=\; \frac{k}{4\pi}\underbrace{\frac{4\pi}{k^2}\sum (2l+1)\sin^2\delta_l}_{\sigma}.$$

to obtain directly the optical theorem

$$\sigma = \frac{4\pi}{k}\,\text{Im } f_{\boldsymbol{k}}(\theta = 0). \tag{1.49}$$

The physical reason for the optical theorem is found in the conservation of particle number. The scattered particle current $\propto \sigma$ must be removed from the incident current by scattering and is thus absent from the forward amplitude ($\propto f_{\boldsymbol{k}}(\theta = 0)$). It is the interference of the scattered wave with the incident wave that reduces the unscattered wave and thereby produces a shadow of the scatterer in the forward direction. The particles missing in the shadow are precisely those that were scattered. The relation (1.49) is generally valid as long as particle number is conserved, which is the case when there is no capture or conversion of the particles.

To demonstrate the optical theorem for a general potential (not necessarily radially symmetric), consider the radial current density $j_r = (\hbar/m)\,\mathrm{Im}(\Psi^* \partial_r \Psi)$ of a wave packet after the scattering event in the form (1.23). The radial current density

$$j_r(\boldsymbol{r};t) = j_{\mathrm{in},r}(\boldsymbol{r};t) + j_{\mathrm{s},r}(\boldsymbol{r};t) + j_{r,\mathrm{int}}(\boldsymbol{r};t) \tag{1.50}$$

is composed of the incident current density, $j_{\mathrm{in},r}(\boldsymbol{r};t) = \boldsymbol{j}_{\mathrm{in}}(\boldsymbol{r};t)\cdot\hat{\boldsymbol{r}} = (\hbar\boldsymbol{k}\cdot\hat{\boldsymbol{r}}/m)|\Psi_0(\boldsymbol{r};t)|^2$ (1.25), the scattered current density $j_{\mathrm{s},r}$ (1.26), and the interference current density between the scattered and transmitted wave functions [7]

$$j_{\mathrm{int},r}(\boldsymbol{r};t) = \frac{\hbar k}{mr}\,\mathrm{Im}\left[i f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})\Psi_0^*(\boldsymbol{r};t)\Psi_0(\hat{\boldsymbol{k}}r;t)(1+\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}})\right]. \tag{1.51}$$

Next, we use the fact that $\Psi_0(\hat{\boldsymbol{k}}r;t)$ vanishes except at times when the particle has already reached the scattering center. In this case $\Psi_0(\boldsymbol{r};t)$ (and hence the interference current) is nonzero only in the forward direction. Thus in the expression for $j_{\mathrm{int},r}$ we may employ the relations $\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}} = 1$ and $f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}}) = f_{\boldsymbol{k}}(\theta = 0)$.

In essence, the interference term produces the shadow of the scatterer by reducing the current in the forward direction below the incident current density $j_{\mathrm{in},r}$ and thus lowering the forward current by the amount of the scattered current. To conserve the total particle number, the total number of scattered particles must exactly equal the number of particles removed from the incident beam. In other words: the integral of (1.51) over all solid angles and times must be precisely minus the integral of (1.26) over all solid angles and times. Now we have, however,

$$\int d\Omega_{\boldsymbol{r}}\, j_{\mathrm{int},r}(\boldsymbol{r};t) = \frac{2\hbar k}{r}\,\mathrm{Im}\left[i f_{\boldsymbol{k}}(\theta=0)\Psi_0(\hat{\boldsymbol{k}}r;t)\int d\Omega_{\boldsymbol{r}}\,\Psi_0^*(\boldsymbol{r};t)\right]. \tag{1.52}$$

For the angular integration of the wave packet Ψ_0, (1.22), we require [$q \approx \boldsymbol{q}\cdot\hat{\boldsymbol{k}}$ see (1.19)]

$$\int d\Omega_{\boldsymbol{r}}\, e^{i\boldsymbol{q}\cdot\boldsymbol{r}} = 2\pi\int_{-1}^{1} d(\cos\theta) e^{iqr\cos\theta} \overset{\boldsymbol{q}\approx\boldsymbol{k}}{\approx} \frac{2\pi(e^{i\boldsymbol{q}\cdot\hat{\boldsymbol{k}}r} - e^{-i\boldsymbol{q}\cdot\hat{\boldsymbol{k}}r})}{ikr}.$$

Hence we obtain

$$\int d\Omega_{\boldsymbol{r}}\,\Psi_0(\boldsymbol{r};t) = \frac{2\pi}{ikr}[\Psi_0(\hat{\boldsymbol{k}}r;t) - \Psi_0(-\hat{\boldsymbol{k}}r;t)],$$

[7] We use $\partial_r\Psi_0(\hat{\boldsymbol{k}}r;t) \approx ik\Psi_0(\hat{\boldsymbol{k}}r;t)$, $\partial_r\Psi_0(\boldsymbol{r};t) \approx (i\boldsymbol{k}\cdot\hat{\boldsymbol{r}})\Psi_0(\boldsymbol{r};t)$, cf. Chapter 1.3.

and (1.52) becomes

$$\int d\Omega_{\boldsymbol{r}}\, j_{\mathrm{int},r}(\boldsymbol{r};t) = -\frac{4\pi\hbar}{mr^2}\,\mathrm{Im}\, f_{\boldsymbol{k}}(\theta=0)|\Psi_0(\hat{\boldsymbol{k}}r;t)|^2, \tag{1.53}$$

since $\Psi_0(\hat{\boldsymbol{k}}r;t)\Psi_0(-\hat{\boldsymbol{k}}r;t) = 0$ in the far field. Comparing now (1.53) with the integral of (1.26) over all angles, we recover the optical theorem $\sigma = (4\pi/k)\,\mathrm{Im}\, f_{\boldsymbol{k}}(\theta=0)$. Note that this proof is independent of any possible radial symmetry of the scattering problem.

1.6 Born Approximation

Except in special cases one can determine the scattering amplitude or the phase shifts only numerically by solving (1.14) or (1.33), respectively. For weak potentials (what exactly constitutes a "weak" potential will be defined more precisely later) one may attempt to solve the Lippmann-Schwinger equation 1.14 iteratively

$$\Psi_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + \int d^3r'\, G(\boldsymbol{r}-\boldsymbol{r}';E_k)V(\boldsymbol{r}')\Psi_{\boldsymbol{k}}(\boldsymbol{r}') \tag{1.54}$$

1^{th}, 2^{th}, n^{th}

$\downarrow$

0^{th} approximation

$$\begin{aligned}
\delta\Psi^{(0)}_{\boldsymbol{k}}(\boldsymbol{r}) &= e^{i\boldsymbol{k}\cdot\boldsymbol{r}},\\
\delta\Psi^{(1)}_{\boldsymbol{k}}(\boldsymbol{r}) &= \int d^3r'\, G(\boldsymbol{r}-\boldsymbol{r}';E_k)V(\boldsymbol{r}')\overbrace{e^{i\boldsymbol{k}\cdot\boldsymbol{r}'}}^{\delta\Psi^{(0)}_{\boldsymbol{k}}(\boldsymbol{r}')},\\
&\vdots\\
\delta\Psi^{(n)}_{\boldsymbol{k}}(\boldsymbol{r}) &= \int d^3r'\, G(\boldsymbol{r}-\boldsymbol{r}';E_k)V(\boldsymbol{r}')\delta\Psi^{(n-1)}_{\boldsymbol{k}}(\boldsymbol{r}'),\\
\rightarrow \Psi^{(N)}_{\boldsymbol{k}}(\boldsymbol{r}) &= \sum_{n=0}^{N}\delta\Psi^{(n)}_{\boldsymbol{k}}(\boldsymbol{r}).
\end{aligned} \tag{1.55}$$

For $N < \infty$, (1.55) is the N^{th} Born approximation; the limiting case $N \to \infty$ yields the Born series.

Hereafter we will use only the first Born approximation,

$$\Psi^{(1)}_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + \int d^3r'\, G(\boldsymbol{r}-\boldsymbol{r}';E_k)V(\boldsymbol{r}')\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}'}, \tag{1.56}$$

and for the scattering amplitude $f^{(1)}_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})$ we find from (1.16)

$$f^{(1)}_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}}) = -\frac{m}{2\pi\hbar^2}\int d^3r'\, V(\boldsymbol{r}')\, e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}'} \tag{1.57}$$

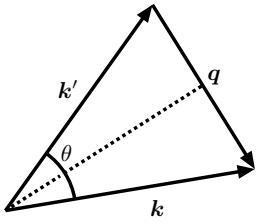


Figure 1.6: Sketch of the scattering angle θ, where $\boldsymbol{q} = \boldsymbol{k} - \boldsymbol{k}'$. Since the triangle is isosceles, the altitude (dashed line) bisects the triangle and we have $q = 2k|\sin(\theta/2)|$.

with the $\boldsymbol{k}' = \boldsymbol{k}_s = k\hat{\boldsymbol{r}}$; i.e., $f^{(1)}_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})$ is proportional to the Fourier transform $V_{\boldsymbol{k}'-\boldsymbol{k}}$ of the scattering potential.

Next we wish to obtain an explicit expression for the scattering phase shifts δ_l of a spherically symmetric potential. To this end we use the fact that in this case the scattering amplitude depends only on $q = |\boldsymbol{k} - \boldsymbol{k}'| = 2k|\sin(\theta/2)|$, see Figure 1.6. The angular integration in (1.57) can therefore be performed directly, and we obtain (Ω' is the angle between $\boldsymbol{r}'$ and $\boldsymbol{q}$)

$$\begin{aligned} f^{(1)}_{\boldsymbol{k}}(\theta) &= -\frac{m}{2\pi\hbar^2}\int_0^{2\pi} dr'\, r'^2\, V(r') \overbrace{\int d\Omega'\, e^{iqr'\cos\theta'}}^{2\pi(e^{iqr'}-e^{-iqr'})/iqr'} \\ &= -\frac{2m}{\hbar^2}\int_0^{\infty} dr'\, r'^2\, V(r')\frac{\sin(qr')}{qr'} \end{aligned} \tag{1.58}$$

Using now the addition theorem for Legendre polynomials[8]

$$\frac{\sin\Big[2x\sin(\theta/2)\Big]}{2x\sin(\theta/2)} = \sum_{l=0}^{\infty}(2l+1)P_l(\cos\theta)j_l^2(x) \tag{1.59}$$

with $x = kr'$ and (1.41), we obtain the expression for the scattering phase shift

$$\delta^{(1)}_l(E_k) \approx e^{i\delta_l}\sin(\delta_l) = -\frac{2mk}{\hbar^2}\int_0^{\infty} dr\, r^2\, V(r)\, j_l^2(kr). \tag{1.60}$$

in the Born approximation.

The validity of the approximation must be checked case by case. It is certainly necessary that $|\delta\Psi^{(1)}(r)| \ll |\delta\Psi^{(0)}(\boldsymbol{r})| = 1$; i.e.,

$$\left|\frac{m}{2\pi\hbar^2}\int d^3r' V(\boldsymbol{r}')\frac{e^{ik|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}e^{i\boldsymbol{k}\cdot\boldsymbol{r}'}\right| \ll 1. \tag{1.61}$$

For large particle energies and weak scattering potentials, such that $V_0\, ka \ll E_k$ (where $V_0 \geq 0$ denotes the strength and a the range of $V(r)$), the Born approximation

[8] See Abramowitz and Stegun 10.1.45.

is valid. To show this, we examine $r = 0$ (where $\delta\Psi^{(1)}(r)$ is largest). The correction to the wave function $|\delta\Psi^{(1)}(r=0)|$ can be written as

$$\left|\frac{m}{2\pi\hbar^2}\int d^3r'\frac{V(r')}{r'}\,e^{ikr'(1+\cos\theta')}\right| = \left|\frac{m}{\hbar^2 k}\int_0^\infty dr'\,V(r')\left(e^{2ikr'}-1\right)\right|.$$

For $ka \gg 1$ and V sufficiently smooth, the factor $\propto \exp(2ikr')$ averages out and with $V(r) \approx V_0$ (for $r < a$) we obtain

$$\left|\frac{m}{\hbar^2 k}\int_0^\infty dr\,V(r)\right| \ll 1 \Rightarrow V_0 ka \ll E_k. \tag{1.62}$$

For small energies with $ka \ll 1$ the above condition is not satisfied. However, in this limiting case one may use $\exp(2ikr') - 1 \sim 2ikr'$ so that we obtain the condition

$$\left|\frac{m}{\hbar^2}\int_0^\infty dr\,r\,V(r)\right| \ll 1 \Rightarrow V_0 a^2 \ll \frac{\hbar^2}{m} \tag{1.63}$$

i.e., $V_0 \ll E_B = \hbar^2/ma^2$, corresponding to a flat potential (a corresponding attractive flat potential has no bound states). The Born approximation is therefore applicable provided that either

$$V_0 ka \ll E_k \quad \text{or} \quad V_0 \ll E_B \tag{1.64}$$

is satisfied.

1.7 Eikonal approximation

The eikonal approximation to scattering corresponds to geometrical optics as an approximation to light propagation. It is valid as long as $V(r)$ varies little over a distance of order of a wavelength $\lambda\!\!\!^{-} = k^{-1}$. Thus $\lambda\!\!\!^{-}$ is our "small" parameter. Note that we need not assume that V itself is weak, as long as $E_k \gg V_0$, and hence the domain of validity of the eikonal approximation differs from that of the Born approximation. As already discussed in the WKB approximation, one may replace the exact wave function $\Psi_{\boldsymbol{k}}(\boldsymbol{r})$ by the semiclassical wave function

$$\Psi_{\boldsymbol{k}}(\boldsymbol{r}) \sim e^{i\mathcal{S}(\boldsymbol{r})/\hbar} \tag{1.65}$$

If in addition we assume that $\hbar|\nabla^2\mathcal{S}| \ll |\nabla\mathcal{S}|^2$, which formally corresponds to the limit $\hbar \to 0$, then the Schrödinger equation reduces to the Hamilton-Jacobi equation

$$\frac{(\boldsymbol{\nabla}\mathcal{S})^2}{2m} + V(\boldsymbol{r}) = E_k \tag{1.66}$$

for the action $\mathcal{S}(\boldsymbol{r})$. The solution of equation (1.66) consists of the solution to the classical scattering problem and is in general not explicitly determinable. Therefore

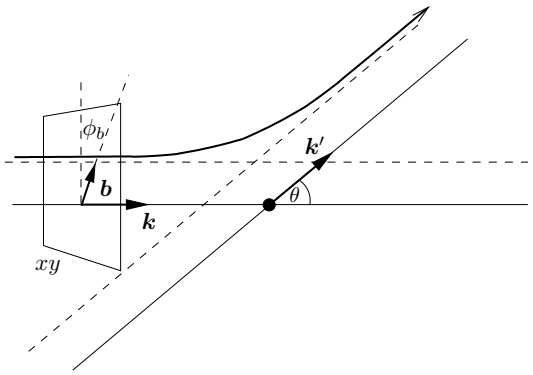


Figure 1.7: Scattering path with impact parameter $\boldsymbol{b}$ in the xy-plane. The particle approaches (quasi-) "classically" the scattering center in the direction $\boldsymbol{k} \perp \boldsymbol{b}$. It is deflected by the center in the direction $\boldsymbol{k}'$ by the scattering angle $\theta \ll 1$.

we make the additional assumption that we are interested only in small scattering angles and replace the classical trajectory by a straight line with $S(\boldsymbol{r}) = S(z)$. This reduces the partial differential equation (1.66) to an ordinary differential equation. For a particle incident along the z-axis with impact parameter $\boldsymbol{b} \perp \hat{\boldsymbol{z}}$, see Figure 1.7, we obtain

$$\frac{\mathcal{S}}{\hbar} = \int_{-\infty}^{z} dz' \left[k^2 - \frac{2m}{\hbar^2} V\left(\sqrt{b^2 + z'^2}\right)\right]^{1/2}; \tag{1.67}$$

the additive constant is fixed by the requirement $\mathcal{S}/\hbar = kz$ for $V \equiv 0$ so that the scattering solution $\Psi_{\boldsymbol{k}}$ without potential reduces to a plane wave. For $E_k \gg V_0$ we may expand the square root and obtain

$$\frac{\mathcal{S}}{\hbar} \approx kz - \frac{m}{\hbar^2 k} \int_{-\infty}^{z} dz'\, V\left(\sqrt{b^2 + z'^2}\right), \tag{1.68}$$

leading to

$$\Psi_{\boldsymbol{k}}(\boldsymbol{r}) = \Psi_{\boldsymbol{k}}(\boldsymbol{b} + z\hat{\boldsymbol{z}}) \approx e^{ikz} \exp\left[-\frac{im}{\hbar^2 k} \int_{-\infty}^{z} dz'\, V\left(\sqrt{b^2 + z'^2}\right)\right]. \tag{1.69}$$

Note that the expression (1.68) does not have the correct asymptotic form $e^{ikz} + f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})e^{ikr}/r$ since it is valid only in the forward direction (with $\theta \approx 0$). Nevertheless, we can use (1.16) to determine the scattering amplitude

$$f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}}) \approx -\frac{m}{2\pi\hbar^2} \int d^3r'\, e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}'} V\left(r'\right) \times \exp\left[-\frac{im}{\hbar^2 k} \int_{-\infty}^{z'} dz''\, V\left(\sqrt{b^2 + z''^2}\right)\right] \tag{1.70}$$

The reason is that the potential restricts the integration to the region $|\boldsymbol{r}'| \leq a$ where the relation (1.69) holds. Note that the expression without the last factor is precisely the Born approximation.

As a final step we perform in $d^3r' = d^2b\,dz'$ the integration over $d^2b = b\,db\,d\phi_b$. To this end we use the approximation [we employ $\boldsymbol{k} \perp \boldsymbol{b}$ and $(\boldsymbol{k} - \boldsymbol{k}') \cdot \hat{\boldsymbol{z}} = O(\theta^2)$]

$$(\boldsymbol{k} - \boldsymbol{k}') \cdot \boldsymbol{r}' \approx -\boldsymbol{k}' \cdot \boldsymbol{b} \approx -kb\theta \cos\phi_b \tag{1.71}$$

for small angles and obtain

$$\begin{aligned} f_{\boldsymbol{k}}^{\text{eik}}(\Omega_{\boldsymbol{r}}) &= -\frac{m}{2\pi\hbar^2} \int_0^\infty db\, b \overbrace{\int_0^{2\pi} d\phi_b e^{-ikb\theta\cos\phi_b}}^{2\pi J_0(kb\theta)} \\ &\qquad \times \int_{-\infty}^{\infty} dz \underbrace{V \exp\left[-\frac{im}{\hbar^2 k}\int_{-\infty}^{z} dz'\, V\right]}_{\frac{i\hbar^2 k}{m}\frac{d}{dz}\exp\left[-\frac{im}{\hbar^2 k}\int_{-\infty}^{z} dz'\, V\right]} \\ &= -ik \int_0^\infty db\, b J_0(kb\theta) \left[e^{2i\Delta(b)} - 1\right] \end{aligned} \tag{1.72}$$

with the scattering phase

$$\Delta(b) = -\frac{m}{2k\hbar^2} \int_{-\infty}^{\infty} dz\, V\left(\sqrt{b^2 + z^2}\right) \simeq \frac{V_0 ka}{E_k}. \tag{1.73}$$

In the derivation of the eikonal approximation we have assumed that

$$V_0 \ll E_k = E_B\,(ka)^2 \quad \text{and} \quad ka \gg 1 \tag{1.74}$$

holds; i.e., the eikonal approximation also applies when $V_0 ka \gtrsim E_k$, where the Born approximation cannot be used. In this sense the eikonal approximation is complementary to the Born approximation. In the limit $V_0 ka \ll E_k$ we may use $\exp[2i\Delta(b)] - 1 \approx 2i\Delta(b)$ so that with (1.71)

$$\begin{aligned} f_{\boldsymbol{k}}^{\text{eik}}(\theta) &= -\frac{m}{\hbar^2} \int_{-\infty}^{\infty} dz \int_0^\infty db\, b\, J_0(kb\theta) V\left(\sqrt{b^2 + z^2}\right) \\ &= -\frac{m}{2\pi\hbar^2} \int d^3r'\, V(r') e^{i(\boldsymbol{k}-\boldsymbol{k})'\cdot\boldsymbol{r}'} \end{aligned} \tag{1.75}$$

in agreement with (1.57).

We now show that, unlike the Born approximation, the eikonal approximation satisfies the optical theorem (1.49). Thus the eikonal approximation yields self-consistent results that respect particle-number conservation. To this end we compute the total

cross section

$$\begin{aligned}\sigma &= 2\pi \int_0^\infty d(\cos\theta)|f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})|^2 \\ &= 2\pi k^2 \int_0^\infty db\, b \left[e^{2i\Delta(b)} - 1\right] \int_0^\infty db'\, b' \left[e^{-2i\Delta(b')} - 1\right] \\ &\quad \times \underbrace{\int_0^\pi d\theta\, \sin\theta J_0(kb\theta) J_0(kb'\theta)}_{=I} . \end{aligned} \tag{1.76}$$

Because of the finite range of the potential we have $\Delta(b \gtrsim a) \approx 0$, so that only the region $b \lesssim a$ gives the dominant contribution to the integral in (1.76). In the angular integral I we substitute $x = ka\theta$ and obtain [9]

$$\begin{aligned} I &= \frac{1}{ka} \int_0^{ka\pi} dx\, \underbrace{\sin(x/ka)}_{\approx x/ka} J_0(bx/a) J_0(b'x/a) \\ &\approx \frac{1}{(ka)^2} \int_0^\infty dx\, x J_0(bx/a) J_0(b'x/a) = \frac{1}{k^2 b} \delta(b - b'), \end{aligned} \tag{1.77}$$

where we have used that the Bessel functions cut off the integral at $x \simeq b/a \lesssim 1$, so that we may extend the upper integration limit to ∞ and expand the sine. Inserting (1.77) into (1.76) yields the optical theorem

$$\sigma = 2\pi \int_0^\infty db\, b \left| e^{2i\Delta(b)} - 1 \right|^2 = 8\pi \int_0^\infty db\, b \sin^2 \Delta(b) = \frac{4\pi}{k} \operatorname{Im} f_{\boldsymbol{k}}(\theta = 0)$$

as the final result. Note that the optical theorem allows a simple calculation of the total cross section via the determination of $\operatorname{Im} f_{\boldsymbol{k}}(\theta = 0)$.

There is a nice connection between the eikonal approximation and the partial-wave expansion for $l \gg 1$. In this case one may regard l as a continuous variable. Equating the quantum-mechanical angular momentum $\hbar l$ with the semiclassical angular

[9] Here we use the orthogonality relation

$$\int_0^\infty dx\, x J_0(kx) J_0(k'x) = \frac{1}{k} \delta(k' - k)$$

of the Bessel functions, which can be derived from the well-known relation $\delta^{(2)}(\boldsymbol{r} - \boldsymbol{r}') = \int d^2k/(2\pi)^2 \times e^{i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')}$ by integration over the angle ϕ as follows

$$\begin{aligned} \frac{2\pi}{r} \delta(r - r') &= \int_0^{2\pi} d\phi\, \delta^{(2)}(\boldsymbol{r} - \boldsymbol{r}') = \int_0^\infty dk\, k \int_0^{2\pi} \frac{d\varphi}{2\pi} \overbrace{\int_0^{2\pi} d\phi\, e^{ikr\cos(\phi-\varphi)}}^{2\pi J_0(kr)} e^{-ikr'\cos(\phi'-\varphi)} \\ &= 2\pi \int_0^\infty dk\, k J_0(kr) J_0(kr'). \end{aligned}$$

momentum $bp = \hbar bk$ yields $l = bk$. The following relations

$$\sum_l \mapsto k \int db, \qquad P_l(\cos\theta) \overset{\theta \ll 1}{\approx} J_0(l\theta) = J_0(kb\theta), \qquad \delta_l \mapsto \Delta(b = l/k), \tag{1.78}$$

enable one to "derive" equation (1.72) directly from (1.41).

1.8 Analytic Properties of the Scattering Matrix

We consider a short-range potential that vanishes outside of $r = a$ (with more effort one can show that the results remain valid as long as $r^2 V(r) \to 0$ for $r \to \infty$). The radial solution for $r > a$ is then given by

$$R_l(r) = \frac{1}{2}\left[h_l^{(2)}(kr) + S_l\, h_l^{(1)}(kr)\right], \tag{1.79}$$

while R_l for $r < a$ must be found by integrating (1.33). The scattering phase S_l is to be chosen so that both R_l and $\partial_r R_l$ are continuous at $r = a$. Since the wave function is determined only up to a (normalization) factor, we need only require continuity of the logarithmic derivative, $\partial_r \log R_l = (\partial_r R_l)/R_l$. This yields the boundary condition

$$\alpha_l = \frac{1}{R_l}\frac{\partial R_l}{\partial r}\bigg|_{a^-} = \frac{\partial_r h_l^{(2)} + S_l\, \partial_r h_l^{(1)}}{h_l^{(2)} + S_l\, h_l^{(1)}}\bigg|_{a^+} = \frac{2\partial_r j_l + (S_l - 1)\partial_r h_l^{(1)}}{2j_l + (S_l - 1)h_l^{(1)}}\bigg|_{a^+}.$$

Solving for the scattering phase gives the result

$$S_l - 1 = \frac{2(\partial_r - \alpha_l) j_l}{(\alpha_l - \partial_r) h_l^{(1)}}\bigg|_{a^+} \tag{1.80}$$

$$\downarrow \text{ with } S_l - 1 = \frac{2i}{\cot\delta_l - i}$$

$$\cot\delta_l = \frac{(\partial_r - \alpha_l) n_l}{(\partial_r - \alpha_l) j_l}\bigg|_{a^+}, \tag{1.81}$$

whence we have determined δ_l from α_l. The partial cross section as a function of $\cot\delta_l$ is given by

$$\sigma_l = \frac{4\pi}{k^2}(2l+1)\sin^2\delta_l = \frac{4\pi}{k^2}\frac{2l+1}{1 + \cot^2\delta_l}. \tag{1.82}$$

Analysis of the expressions for $S_l(\cot\delta_l)$ and $\sigma_l(\cot\delta_l)$ shows that the following statements hold: For[10]

$$\begin{aligned} &\cot\delta_l = i \quad S_l \text{ has a pole and } \sigma_l \text{ is } \infty, \\ &\cot\delta_l = 0 \quad S_l = -1 \text{ and } \sigma_l = 4\pi(2l+1)/k^2 \text{ is maximal.} \end{aligned} \tag{1.83}$$

[10] Note that $\cot\delta_l = i$ is not possible for $\delta_l(E)$ with $E \in \mathbb{R}^+$. Here we consider the analytic continuation of δ_l to the entire complex plane. On the real axis we have $\sigma_l \le 4\pi(2l+1)/k^2$ and the maximum is attained for $\cot\delta_l = 0$.

As in one dimension (see exercises) the poles of S_l correspond precisely to bound states. To understand this, we examine S_l for a bound state with $E = -E_B < 0$. The radial wave function has in this case the asymptotic form $R_l(r) \sim h_l^{(1)}(i\kappa r) \propto e^{-\kappa r}$ with the binding energy $E_B = \hbar^2\kappa^2/2m$. The term $\propto h_l^{(2)}$ is not allowed, since it would lead to an exponentially growing wave function that is not normalizable. The continuity condition is therefore given by

$$\alpha_l = \partial_r h_l^{(1)}/h_l^{(1)}\Big|_{a^+} = \partial_r \log h_l^{(1)}\Big|_{a^+}. \tag{1.84}$$

Inserting this relation into the general formula (1.81) yields the expected result

$$\cot\delta_l = \frac{h_l^{(1)}\partial_r n_l - n_l\,\partial_r h_l^{(1)}}{h_l^{(1)}\partial_r j_l - j_l\,\partial_r h_l^{(1)}} \overset{h_l^{(1)}=j_l+in_l}{=} i. \tag{1.85}$$

Similarly we wish to show that the zeros of $\cot\delta_l$ correspond precisely to scattering resonances. To this end we expand $\cos\delta_l$ about the resonance with $E = E_r > 0$,

$$\begin{aligned}\cot\delta_l(E) &\approx \overbrace{\cot\delta_l(E_r)}^{=0} - \frac{1}{\sin^2\delta_l}\,\frac{d\delta_l}{dE}\bigg|_{E_r}(E-E_r)\\ &= -\frac{d\delta_l}{dE}\bigg|_{E_r}(E-E_r) \qquad (1.86)\\ &\downarrow \text{ with } \Gamma_r = \tfrac{2}{\delta_l'(E_r)}\\ &= -\frac{2}{\Gamma_r}(E-E_r). \qquad (1.87)\end{aligned}$$

Substituting into (1.82) we see that the partial cross section σ_l exhibits a resonance of width Γ_r about E_r,

$$\sigma_l = \frac{4\pi}{k^2}(2l+1)\frac{(\Gamma_r/2)^2}{(E-E_r)^2+(\Gamma_r/2)^2}, \tag{1.88}$$

cf. Figure 1.8. Analogously we find that the scattering amplitude $S_l(E)$ has a pole in the second Riemann sheet of the complex E plane with

$$S_l - 1 = \frac{-i\Gamma_r}{E-(E_r - i\Gamma_r/2)}. \tag{1.89}$$

Across a resonance the scattering phase $\delta_l(E)$ increases by π; the narrower the resonance, the more abrupt the increase. The value $\delta_l(E=0)$ is proportional to the number of bound states ν_l in the angular-momentum sector l. We have $\delta_l(0) = \nu_l\pi$, where an additional phase shift $+\pi/2$ in the sector $l = 0$ occurs if a "virtual" state[11] is present at $E = 0$. This is the statement of Levinson's theorem.

In the following, we examine the low-energy $ka \ll 1$ behavior for a scatterer of extension a in more detail.

[11] In the $l = 0$ sector a bound state at $E < 0$ becomes a virtual state (instead of a resonance) when the binding energy vanishes, i.e., with $E_B \to 0$.

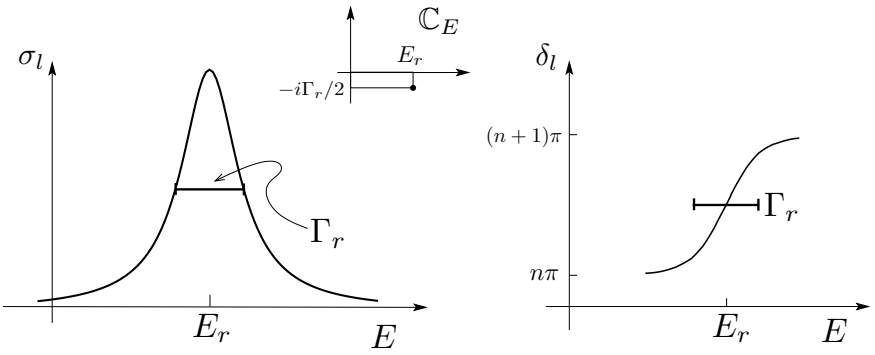


Figure 1.8: Resonance of width Γ_r: On the left the cross section $\sigma_l(E)$ and on the right the scattering phase $\delta_l(E)$; in the center the location of the pole in the complex E plane is sketched.

1.9 Low-energy scattering, $ka \ll 1$

Low-energy scattering has recently become a highly relevant topic, since it represents the primary scattering process for cold atoms in quantum atom optics. We use the expansions [12] $j_l \sim x^l/(2l+1)!!$ and $n_l \sim (2l-1)!!/x^{l+1}$ in (1.81) and obtain for $ka \ll 1$

$$\cot\delta_l \approx \frac{(2l+1)!!\,(2l-1)!!}{(ka)^{2l+1}} \frac{l+1+a\,\alpha_l(E)}{l-a\,\alpha_l(E)}. \tag{1.90}$$

For $ka \to 0$, $\cot\delta_l \to \infty$ vanishes, since $\alpha_l(E) \approx \alpha_l(0)$, yielding the result

$$\delta_l \approx \sin\delta_l \approx \tan\delta_l \sim (ak)^{2l+1}; \tag{1.91}$$

i.e., the scattering phases δ_l and the partial cross sections σ_l decrease rapidly with increasing angular momentum in the low-energy regime. The s-wave sector is therefore dominant, and we focus on this case. It is instructive to express the scattering phase δ_0 in terms of the scattering length a_s,

$$k\cot\delta_0 \overset{\alpha_0(E)\approx\alpha_0(0)}{\approx} -\frac{1+a\alpha_0(0)}{a^2\alpha_0(0)} = -\frac{1}{a_s}; \tag{1.92}$$

thereby reducing the scattering length a_s to the sole relevant parameter in the problem.

For the important case that $a\alpha_0 \gg 1$ is $a_s \approx a$. [13] The total cross section σ is

$$\sigma \approx \sigma_0 = \frac{4\pi}{k^2}\frac{1}{1+\cot^2\delta_0} \approx \frac{4\pi a_s^2}{1+(ka_s)^2}. \tag{1.93}$$

[12] The double factorial $(2l-1)!!$ of an odd number $2l-1$ is defined by $(2l-1)!! = (2l)!/(2^l l!) = 1\cdot 3\cdot 5\cdots(2l-1)$.

[13] For the hard sphere, $R_l(a) = 0$, so that $\alpha_l = \infty$, $a_s = a > 0$ and $\cot\delta_0 = -1/ka$. For small ka, therefore, $\delta_0 \approx -ka < 0$. The negative phase shift was to be expected, since the potential is repulsive.

In particular,

$$\sigma(E=0) = 4\pi a_s^2; \tag{1.94}$$

for the hard sphere with $a_s = a$ we find a fourfold larger scattering cross section than classically expected, $\sigma_{\rm kl} = \pi a^2$.

1.9.1 Bound states and resonances

We wish to examine once more the analytic properties of the scattering matrix (with bound states and scattering resonances), discussed in Chapter 1.8, using the example of low-energy s-wave scattering. The scattering phase can be written in terms of the scattering length a_s as

$$S_0 - 1 = \frac{2i}{\cot\delta_0 - i} \approx \frac{2ka_s}{i - ka_s} \qquad (\text{for } ka \ll 1). \tag{1.95}$$

The scattering matrix $S_0(E)$ has a pole at $k = i/a_s$ or $\kappa = -ik = a_s^{-1} > 0$ provided that $a_s > 0$ and $a \ll a_s$. [14] This yields a (weakly) bound state at energy $E = -E_B = -\hbar^2/2ma_s^2$. The argument can also be reversed and one may say that a weakly bound state in the potential produces a large positive scattering length $a_s \gg a > 0$, i.e., a weakly bound state repels a low-energy scattering particle. Computing the total cross section, cf. (1.93),

$$\sigma \approx \sigma_0 \approx \frac{2\pi\hbar^2/m}{E + E_B}, \quad E_B \gtrsim 0. \tag{1.96}$$

one obtains the remarkable result that the energy of the weakly bound state completely determines the low-energy scattering cross section.

Analogously, one obtains a resonance in the s channel, see (1.90), if

$$\cot\delta_0 = 0 \Rightarrow 1 + a\alpha_0(E_r) = 0. \tag{1.97}$$

For a resonance, the additional condition $\Gamma_r < E_r$ must also hold, which guarantees a (albeit broad) resonance maximum. With

$$\begin{aligned}\cot\delta_0(E) &= \frac{1}{k_r a}\frac{\overbrace{1 + a\alpha_0(E_r)}^{=0} + a\alpha_0'(E_r)(E - E_r)}{\underbrace{-a\alpha_0(E_r)}_{=1}} \\ &\approx -\frac{2}{\Gamma_r}(E - E_r)\end{aligned} \tag{1.98}$$

[14] That the condition $a \ll a_s$ must be satisfied is seen from the example of the hard sphere. We have $a = a_s$ but of course the potential has no bound state and hence $S_0 = e^{2i\delta_0}$ has no pole.

we obtain the condition

$$\Gamma_r = -\frac{2k_r a}{a\alpha_0'(E_r)} = -\frac{2k_r}{\alpha_0'(E_r)} < \frac{\hbar^2 k_r^2}{2m} = E_r, \tag{1.99}$$

$$\Rightarrow \left|\frac{\partial \alpha_0}{\partial k^2}\right| > \frac{2}{k_r} \tag{1.100}$$

for the formation of a resonance in the s channel.

Since resonances in the s channel are typically not very pronounced due to the strict condition (1.100), we next examine the case $l > 0$. We have already seen in (1.91) that the scattering for $l > 0$ is generally small,

$$\sigma_l = \frac{4\pi}{k^2}\frac{(2l+1)}{1+\cot^2\delta_l} = \frac{4\pi}{k^2}(2l+1)\sin^2\delta_l \overset{(k\to 0)}{\propto} a^2(ak)^{4l}. \tag{1.101}$$

Near a resonance, however, $\cot\delta_l \approx 0$ and σ_l passes through a maximum with

$$\sigma_l = \frac{4\pi}{k_r^2}(2l+1). \tag{1.102}$$

According to (1.90), we find resonances at

$$l + 1 + a\alpha_l(E_r) = 0. \tag{1.103}$$

For the width of the resonance we obtain [15]

$$\Gamma_r = -\frac{2k_r(ak_r)^{2l}}{[(2l-1)!!]^2\alpha_l'(E_r)} \propto k_r^{2l+1}; \tag{1.104}$$

i.e., the smaller k_r, the larger $l \geq 1$, and the larger $\alpha_l'(E_r)$, the sharper the resonance. Here it should be noted that for $l > 0$, $\Gamma_r \ll E_r \propto k_r^2$ if $k_r \to 0$. For $l = 0$ on the other hand E_r vanishes faster than Γ_r and therefore $\alpha_0'(E_r)$ must become large in the denominator to preserve the visibility of the resonance, cf. 1.100.

[15] We use that the expression $l - a\alpha_l$ in the denominator of

$$\cot\delta_l \approx \frac{(2l+1)!!(2l-1)!!}{(k_r a)^{2l+1}}\frac{l+1+a\left[\alpha_l(E_r)+\alpha_l'(E_r)(E-E_r)\right]}{l-a\alpha_l(E_r)} = -\frac{2}{\Gamma_r}(E-E_r)$$

due to (1.103) can be replaced by $2l+1$.

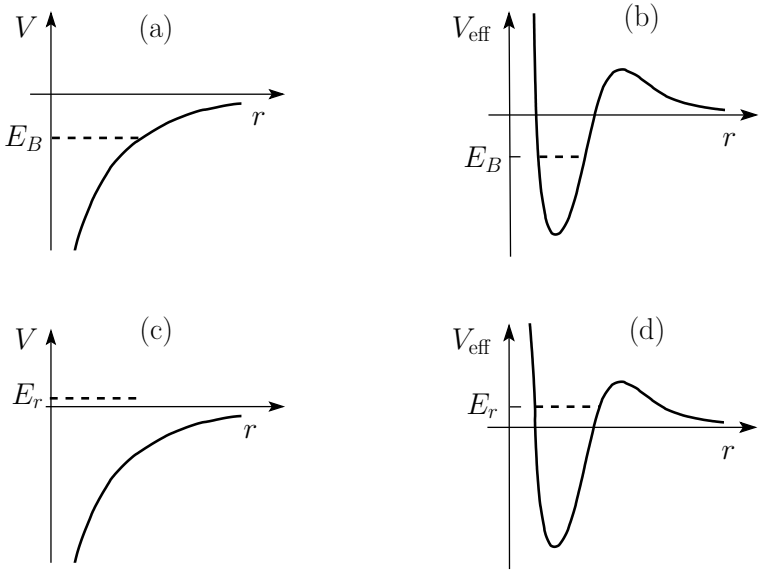


Figure 1.9: Bound states (a,b) and resonances (c,d): (a) Bound state for $l = 0$. (b) Bound state for $l > 0$, with the potential $V(r)$ replaced by the effective potential including the centrifugal barrier $\hbar^2 l(l+1)/2mr^2$. (c) In the s-wave sector the resonances are broad, mostly undefined with $\Gamma_r > E_r$. A defined resonance with $\Gamma_r < E_r$ requires that $|\alpha_0'(E_r)|$ be large. (d) For $l > 0$ sharp resonances arise, since the decay of the state is suppressed by the presence of the centrifugal barrier. The resonances are the sharper the smaller k_r, the larger l, and the larger $\alpha_l'(E_r)$ are.

Chapter 2

Identical Particles

When we wish to describe several identical particles in quantum mechanics, we must introduce the principle of the 'indistinguishability' of like particles, which has no counterpart in classical mechanics. In classical mechanics, like particles, despite identical physical properties, do not lose their individuality; in principle we can brand each particle and thereby identify it. In fact, the brands are even superfluous since we can follow the particles' trajectories and thus identify them—classically the physics of distinguishable and identical particles is the same.

In quantum mechanics we can neither affix brands nor follow trajectories, a consequence of the Heisenberg uncertainty principle. In quantum mechanics, therefore, there is in principle no way to track a single one of like particles separately and thereby distinguish the particles. A consistent account of the indistinguishability of particles yields remarkable physical effects, which we describe in the following.

2.1 Description of N particles

We start with the problem of describing N (distinguishable) particles. The concepts of wavefunctions and operators are generalized from one- to many-particle systems in the obvious way:

- the wavefunctions depend on the N particle coordinates, for example on position and spin coordinates $\boldsymbol{r}_1, s_1; \ldots; \boldsymbol{r}_N, s_N$, or in compact notation $x_1, \ldots, x_N$ with $x_i = (\boldsymbol{r}_i, s_i)$,

$$\Psi(x_1, \ldots, x_N) = \langle x_1, \ldots, x_N | \Psi \rangle; \tag{2.1}$$

  accordingly we define a Hilbert space $\mathcal{H}_N$ of complex-valued functions $\Psi(x_1, \ldots, x_N)$

of N arguments x_N with the scalar product

$$\begin{aligned}\langle\Psi|\Phi\rangle &= \sum_{s_1\cdots s_N}\int dr_1\cdots dr_N\,\Psi^*(x_1,\ldots,x_N)\Phi(x_1,\ldots,x_N)\\ &= \int dx_1\cdots dx_N\,\Psi^*(x_1,\ldots,x_N)\Phi(x_1,\ldots,x_N)\,. \end{aligned}\tag{2.2}$$

The canonical basis in $\mathcal{H}_N$ is the product basis $|\alpha_1,\ldots,\alpha_N\rangle = |\varphi_{\alpha_1}\rangle\otimes\cdots\otimes|\varphi_{\alpha_N}\rangle$ generated from a one-particle basis $\{\varphi_\alpha(x)\}$ with

$$\begin{aligned}\Psi_{\alpha_1,\ldots,\alpha_N}(x_1,\ldots,x_N) &= \langle x_1,\ldots,x_N|\alpha_1,\ldots,\alpha_N\rangle\\ &= \langle x_1|\varphi_{\alpha_1}\rangle\cdots\langle x_N|\varphi_{\alpha_N}\rangle = \varphi_{\alpha_1}(x_1)\cdots\varphi_{\alpha_N}(x_N)\,. \end{aligned}\tag{2.3}$$

Of course other bases are possible as well.

- The operators act on the N-particle wavefunctions. The one-particle operators A are given by

$$A = \sum_{i=1}^{N} a_i,\tag{2.4}$$

where a_i acts only on the i-th particle. In the product basis this looks as follows:

$$\begin{aligned}\langle x_1,\ldots,x_N|A\Psi\rangle &= A\Psi(x_1,\ldots,x_N)\\ &= [a\varphi_{\alpha_1}(x_1)]\varphi_{\alpha_2}(x_2)\ldots\varphi_{\alpha_N}(x_N)\\ &\quad+\varphi_{\alpha_1}(x_1)[a\varphi_{\alpha_2}(x_2)]\ldots\varphi_{\alpha_N}(x_N)+\cdots\,. \end{aligned}\tag{2.5}$$

Mathematically this means that the operator a_i acts as $\cdots\otimes I\otimes a\otimes I\otimes\cdots$ in the Hilbert space $\mathcal{H}_N = \mathcal{H}\otimes\mathcal{H}\otimes\cdots = \mathcal{H}^{\otimes N}$.

Operators may of course also act on two (or more) particles simultaneously and thereby generate interactions between the particles, for example pair interactions,

$$V = \frac{1}{2}\sum_{i\neq j} v_{ij},\qquad v_{ij}\text{ acts only on particle pairs }i,j.\tag{2.6}$$

As an example we consider N particles in $\mathbb{R}^3$ without spin. Then the one-particle Hilbert space is given by $\mathcal{H} = L^2(\mathbb{R}^3)$.

- The product basis is constructed from the one-particle basis $\{|\varphi_{\boldsymbol{k}_i}\rangle\}$,

$$|\boldsymbol{k}_1,\ldots,\boldsymbol{k}_N\rangle = |\varphi_{\boldsymbol{k}_1}\rangle\otimes|\varphi_{\boldsymbol{k}_2}\rangle\otimes\cdots|\varphi_{\boldsymbol{k}_N}\rangle;\tag{2.7}$$

in the position representation it reads

$$\begin{aligned}\langle\boldsymbol{r}_i|\varphi_{\boldsymbol{k}_i}\rangle &= \varphi_{\boldsymbol{k}_i}(\boldsymbol{r}_i),\\ \langle\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N|\boldsymbol{k}_1,\ldots,\boldsymbol{k}_N\rangle &= \varphi_{\boldsymbol{k}_1}(\boldsymbol{r}_1)\cdots\varphi_{\boldsymbol{k}_N}(\boldsymbol{r}_N).\end{aligned}\tag{2.8}$$

- The plane-wave basis, (Ω is the quantization volume)

$$\begin{aligned} \varphi_{\boldsymbol{k}}(\boldsymbol{r}) &= e^{i\boldsymbol{k}\cdot\boldsymbol{r}}/\sqrt{\Omega}, \\ \langle \boldsymbol{r}_1,\ldots,\boldsymbol{r}_N|\boldsymbol{k}_1,\ldots,\boldsymbol{k}_N\rangle &= \prod_i e^{i\boldsymbol{k}_i\cdot\boldsymbol{r}_i}/\sqrt{\Omega}, \end{aligned} \tag{2.9}$$

  spans the product Hilbert space

$$\mathcal{H}^{\otimes N} = L^2(\mathbb{R}^3)\otimes\cdots\otimes L^2(\mathbb{R}^3) = L^2(\mathbb{R}^3)^{\otimes N} \tag{2.10}$$

  .

- Examples of operators are (apart from the interaction all are one-particle operators)
  - the (total) momentum

$$\boldsymbol{P} = \sum_i \boldsymbol{p}_i, \quad \boldsymbol{p}_i \;=\; \frac{\hbar}{i}\boldsymbol{\nabla}_i, \tag{2.11}$$

  - the (total) energy

$$\begin{aligned} H_0 &= \sum_i \frac{p_i^2}{2m}, \quad \text{for free particles,} \\ H &= H_0 + \frac{1}{2}\sum_{i\neq j} V(\boldsymbol{r}_i-\boldsymbol{r}_j), \quad \text{for interacting particles,} \end{aligned} \tag{2.12}$$

  - the density

$$\rho(\boldsymbol{r}) = \sum_i \delta^{(3)}(\boldsymbol{r}-\boldsymbol{r}_i), \tag{2.13}$$

  - the current density

$$\boldsymbol{j}(\boldsymbol{r}) = \sum_i \left[\frac{\boldsymbol{p}_i}{2m}\delta^{(3)}(\boldsymbol{r}-\boldsymbol{r}_i) + \delta^{(3)}(\boldsymbol{r}-\boldsymbol{r}_i)\frac{\boldsymbol{p}_i}{2m}\right]. \tag{2.14}$$

  - We thus form the expectation values

$$\begin{aligned} \langle \boldsymbol{P}\,\rangle &= \sum_i \int d^3r_1\cdots d^3r_N\, \Psi^*(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N)\frac{\hbar}{i}\boldsymbol{\nabla}_i\Psi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N), \\ &= \sum_i \int d^3r_i\, \varphi^*_{\boldsymbol{k}_i}(\boldsymbol{r}_i)\frac{\hbar}{i}\boldsymbol{\nabla}_i\varphi_{\boldsymbol{k}_i}(\boldsymbol{r}_i) = \sum_i \hbar\boldsymbol{k}_i, \end{aligned} \tag{2.15}$$

$$\langle H_0\rangle = \sum_i \frac{\hbar^2 k_i^2}{2m}, \tag{2.16}$$

$$\begin{aligned}\langle H\rangle &= \langle H_0\rangle + \frac{1}{2}\sum_{i\neq j}\int d^3r_1\cdots d^3r_N \\ &\qquad \times \Psi^*(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N)V(\boldsymbol{r}_i-\boldsymbol{r}_j)\Psi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N) \\ &= \langle H_0\rangle + \frac{1}{2}\sum_{i\neq j}\int d^3r_i\, d^3r_j \\ &\qquad \times \varphi^*_{\boldsymbol{k}_i}(\boldsymbol{r}_i)\varphi^*_{\boldsymbol{k}_j}(\boldsymbol{r}_j)V(\boldsymbol{r}_i-\boldsymbol{r}_j)\varphi_{\boldsymbol{k}_i}(\boldsymbol{r}_i)\varphi_{\boldsymbol{k}_j}(\boldsymbol{r}_j) \\ &= \langle H_0\rangle + \frac{1}{2}\sum_{i\neq j}\int d^3r_i\, d^3r_j\, |\varphi_{\boldsymbol{k}_i}(\boldsymbol{r}_i)|^2 V(\boldsymbol{r}_i-\boldsymbol{r}_j)|\varphi_{\boldsymbol{k}_j}(\boldsymbol{r}_j)|^2. \end{aligned} \tag{2.17}$$

2.2 Description of N identical particles

When we write the wavefunction $\Psi(x_1,\ldots,x_N)$ of a many-particle system, we use the term "first particle" for the particle denoted by the first argument, "second particle" for the particle denoted by the second argument, and so on. This raises the question whether we thereby mark the particles and violate their fundamental indistinguishability. The answer is that we need the labels/names in order to describe the system mathematically at all. This does not, however, necessarily imply a physical distinguishability. We simply have to ensure that we treat all particles as absolutely equivalent. In particular, no (physical) operator may exist that would allow an identification of the particles. This means that all admissible operators must be symmetric. In the coordinate representation, for example, the admissible operators are symmetric in the coordinates,

$$A(x_1,\ldots,x_N) = A(x_{\pi(1)},\ldots,x_{\pi(N)}), \tag{2.18}$$

with $\pi \in S_N$ an arbitrary permutation,

$$\begin{pmatrix} 1 & 2 & 3 & \cdots & N \\ \pi(1) & \pi(2) & \pi(3) & \cdots & \pi(N)\end{pmatrix}, \tag{2.19}$$

mapping the upper line to the lower one.

Next we consider the expectation values of A,

$$\begin{aligned}\langle A\rangle &= \int dx_1\cdots dx_N\, \Psi^*(x_1,\ldots,x_N)A(x_1,\ldots,x_N)\Psi(x_1,\ldots,x_N) \\ &= \int dx_1\cdots dx_N\, \Psi^*(x_1,\ldots,x_N)A(x_{\pi(1)},\ldots,x_{\pi(N)})\Psi(x_1,\ldots,x_N) \\ &= \int dx_1\cdots dx_N\, \Psi^*(x_{\pi^{-1}(1)},\ldots,x_{\pi^{-1}(N)})A(x_1,\ldots,x_N)\Psi(x_{\pi^{-1}(1)},\ldots,x_{\pi^{-1}(N)});\end{aligned} \tag{2.20}$$

where π^{-1} is the inverse permutation to π. Evidently we may permute the arguments in Ψ arbitrarily, because A is symmetric $\langle A\rangle$ remains invariant. The invariance of $\langle A\rangle$ thus gives us no condition on the behavior of the wavefunctions Ψ under exchange of the arguments.

We can formalize the above discussion somewhat: Analogous to the translation and rotation representations in Hilbert space,

$$\begin{aligned} &\text{Translations:} && U_{\boldsymbol{a}}\,\Psi(\boldsymbol{r}) = e^{-i\boldsymbol{a}\cdot\boldsymbol{p}/\hbar}\Psi(\boldsymbol{r}) = \Psi(\boldsymbol{r}-\boldsymbol{a}),\\ &\text{Rotations:} && U_{\boldsymbol{\omega}}\,\Psi(\boldsymbol{r}) = e^{-i\boldsymbol{\omega}\cdot\boldsymbol{L}/\hbar}\Psi(\boldsymbol{r}) = \Psi(\mathbf{R}_{-\boldsymbol{\omega}}\,\boldsymbol{r}), \end{aligned} \tag{2.21}$$

we introduce the following (unitary) representation of the permutation group S_N [1]

$$U_\pi\,\Psi(x_1,\ldots,x_N) = \Psi(x_{\pi^{-1}(1)},\ldots,x_{\pi^{-1}(N)}). \tag{2.22}$$

With the definition of U_π we can define symmetric operators as follows: An operator A is symmetric if for all $\pi\in S_N$

$$U_\pi^{-1}AU_\pi = A \qquad\Leftrightarrow\qquad [A,U_\pi] \;=\; 0. \tag{2.23}$$

The central question we now address is how the permuted wavefunction $U_\pi\Psi$ relates to Ψ. Evidently U_π defines a representation of S_N in $\mathcal{H}$. Moreover, for every (admissible, i.e. symmetric) operator A—in particular also for the Hamiltonian H—we have $[H,U_\pi]=0$. This means that we can simultaneously diagonalize any operator A and U_π.

Now let $A\Psi_a = a\Psi_a$; then $AU_\pi\Psi_a = aU_\pi\Psi_a$ and $\{U_\pi\Psi_a \mid \pi\in S_N\}$ spans the eigenspace Eig_a of A [2] Within the eigenspace Eig_a with $\dim\mathrm{Eig}_a = D_a$ one can construct a D_a-dimensional representation of S_N. This representation can then be decomposed into irreducible representations of S_N. In general any irreducible representation of dimension $d\le D_a$ can occur. If $d>1$, however, all energy eigenstates of H (and likewise the eigenstates of any operator A) are d-fold degenerate. This degeneracy is called exchange degeneracy, since it arises only because of the indistinguishability of the particles. This degeneracy has measurable consequences, for example it increases the specific heat or yields a mixing entropy of two identical gases[3]. It is now an experimental fact that no exchange degeneracy of any particle occurs. Hence we need not invoke the full machinery of group theory for this problem and need only consider the one-dimensional representations. In S_N there are two of these, namely the symmetric representation S and the antisymmetric representation A. We denote by $\mathcal{H}_{\mathcal{S}/\mathcal{A}}$ the associated invariant Hilbert spaces.

[1] Since S_N is not a continuous group, no infinitesimal generator exists.

[2] In the case $A=H$ we consider Ψ_E with $H\Psi_E = E\Psi_E$ and observe that $U_\pi\Psi_E$ likewise lies in Eig_E.

[3] The absence of exchange degeneracy in mixing is known in statistical mechanics as the Gibbs paradox .

2.3 Fermions and Bosons

Let $\mathcal{H}$ be the Hilbert space of N-particle functions. Only the sectors $\mathcal{H}$

$$\begin{aligned} \mathcal{H}_{\mathcal{A}} &= \{\mathcal{A}\,\Psi | \Psi \in \mathcal{H}\}, \\ \mathcal{H}_{\mathcal{S}} &= \{\mathcal{S}\,\Psi | \Psi \in \mathcal{H}\}, \end{aligned} \tag{2.24}$$

are relevant; here

$$\mathcal{A} = \frac{1}{N!} \sum_{\pi \in S_N} (-1)^{\pi} U_{\pi}, \tag{2.25}$$

$$\mathcal{S} = \frac{1}{N!} \sum_{\pi \in S_N} U_{\pi}, \tag{2.26}$$

are the (anti) symmetrization operators with $(-1)^{\pi}$ the signature of π, see B.1. The normalization was chosen so that the operators are projectors with $\mathcal{S}^2 = \mathcal{S}$ and $\mathcal{A}^2 = \mathcal{A}$. For the functions in $\mathcal{H}_{\mathcal{A}}$ it holds

$$\Psi(x_{\pi(1)}, \ldots, x_{\pi(N)}) = (-1)^{\pi} \Psi(x_1, \ldots, x_N), \quad \forall \pi \in S_N, \tag{2.27}$$

and for those in $\mathcal{H}_{\mathcal{S}}$

$$\Psi(x_{\pi(1)}, \ldots, x_{\pi(N)}) = \Psi(x_1, \ldots, x_N). \tag{2.28}$$

In nature, therefore, only a small part of $\mathcal{H}$ is realized, namely the symmetric and antisymmetric part. We call particles whose wave functions lie in $\mathcal{H}_{\mathcal{A}}$ ($\mathcal{H}_{\mathcal{S}}$) and hence are antisymmetric (symmetric) under exchange of the arguments fermions (bosons).

$\mathcal{H}_{\mathcal{A}}$		$\mathcal{H}_{\mathcal{S}}$
antisymmetric	$\leftrightarrow$	symmetric
fermions		bosons

This raises the question which Hilbert space, $\mathcal{H}_{\mathcal{A}}$ or $\mathcal{H}_{\mathcal{S}}$, one should choose for a given species of particle, or in other words whether the particles are fermions or bosons. The answer is given by the spin-statistics theorem.

2.4 Spin and Statistics

The following connection between particle spin and the symmetry properties of the wave functions under S_N can be proved in relativistic quantum field theory:

Half-integer spin		Integer spin
Fermions	$\leftrightarrow$	Bosons

Examples of particles with the corresponding properties are:

Fermions

- Leptons $e^-,\ \mu,\ \tau$
- Lepton-neutrinos $\nu_e,\ \nu_\mu,\ \nu_\tau$
- Baryons $p, n, \Lambda, \Sigma, \Xi, \cdots$ with three quarks each
- ^{3}He

All fermions listed here have spin-1/2.

Bosons

- Higgs (spin-0)
- Mesons π, K with spin-0, or ρ, ω with spin-1, each composed of two quarks.
- Gauge bosons (Yang-Mills) of
  - the electroweak interaction: $\gamma, W^\pm, Z^0$ with spin-1.
  - the strong interaction: gluons (spin-1).
- ^{4}He (spin-0).

2.5 Applications

2.5.1 Wavefunctions of two (spin-1/2) fermions

We consider two spin-1/2 fermions, e.g. two electrons. Let $\Psi(\boldsymbol{r}_1, s_1; \boldsymbol{r}_2, s_2) \in \mathcal{H}$ be an element of the product basis

$$\Psi_{\mu,\nu}(x_1, x_2) = \varphi_\mu(x_1)\varphi_\nu(x_2). \tag{2.30}$$

The physical states are

$$\begin{aligned}\Psi^{\mathcal{A}}_{\mu,\nu}(x_1, x_2) = \sqrt{2}\mathcal{A}\Psi_{\mu,\nu} &= \frac{1}{\sqrt{2!}} \sum_{\pi\in S_2} (-1)^\pi U_\pi \Psi_{\mu,\nu}(x_1, x_2) \\ &= \frac{1}{\sqrt{2}}\big[\Psi_{\mu,\nu}(x_1, x_2) - \Psi_{\mu,\nu}(x_2, x_1)\big] \\ &= \frac{1}{\sqrt{2}}\big[\varphi_\mu(x_1)\varphi_\nu(x_2) - \varphi_\mu(x_2)\varphi_\nu(x_1)\big].\end{aligned} \tag{2.31}$$

If $\mu = \nu$, then $\mathcal{A}\Psi_{\mu,\mu} = 0$, so that two fermions cannot occupy the same state: this is the famous *Pauli exclusion principle*. The Pauli exclusion principle for fermions also holds in general for N fermions.

2.5.2 Spin and orbital wavefunctions

Next we separate the spin and orbital parts of the wavefunction Ψ of two spin-1/2 fermions,

$$\Psi(\boldsymbol{r}_1, s_1; \boldsymbol{r}_2, s_2) = \varphi(\boldsymbol{r}_1, \boldsymbol{r}_2)\chi(s_1, s_2). \tag{2.32}$$

The total wavefunction Ψ is antisymmetric if

$$\left.\begin{matrix}\varphi & \text{symmetric} \\ \chi & \text{antisymmetric}\end{matrix}\right\} \text{ or } \begin{cases}\varphi & \text{antisymmetric} \\ \chi & \text{symmetric}\end{cases} \tag{2.33}$$

holds. The spin part $\chi(s_1, s_2)$ can be reduced according to $\boldsymbol{S} = \boldsymbol{S}_1 + \boldsymbol{S}_2$, $\mathcal{H}_{1/2} \otimes \mathcal{H}_{1/2} = \mathcal{H}_0 \oplus \mathcal{H}_1$. In the spin Hilbert space $\mathcal{H}_0$ the singlet state $\chi_{0,0}(s_1, s_2)$ is antisymmetric in s_1, s_2,[4]

$$\begin{aligned}\chi_{0,0}(s_1, s_2) &= \langle s_1, s_2 | 0, 0 \rangle \\ &= \frac{1}{\sqrt{2}} \big[\langle s_1 | \uparrow \rangle \langle s_2 | \downarrow \rangle - \langle s_1 | \downarrow \rangle \langle s_2 | \uparrow \rangle \big] \\ &= \frac{1}{\sqrt{2}} \big[\chi_\uparrow(s_1) \chi_\downarrow(s_2) - \chi_\downarrow(s_1) \chi_\uparrow(s_2) \big]. \end{aligned} \tag{2.34}$$

The spin Hilbert space $\mathcal{H}_1$ is spanned by the three states $\chi_{1,1}$, $\chi_{1,0}$ and $\chi_{1,-1}$. These are the symmetric triplet states,

$$\begin{aligned}\chi_{1,1}(s_1, s_2) &= \langle s_1, s_2 | 1, 1 \rangle = \chi_\uparrow(s_1) \chi_\downarrow(s_2), \\ \chi_{1,0}(s_1, s_2) &= \langle s_1, s_2 | 1, 0 \rangle \\ &= \frac{1}{\sqrt{2}} \big[\chi_\uparrow(s_1) \chi_\downarrow(s_2) + \chi_\downarrow(s_1) \chi_\uparrow(s_2) \big], \\ \chi_{1,-1}(s_1, s_2) &= \langle s_1, s_2 | 1, -1 \rangle = \chi_\downarrow(s_1) \chi_\downarrow(s_2). \end{aligned} \tag{2.35}$$

The above symmetric and antisymmetric spin functions χ are combined with correspondingly symmetrized orbital functions φ. We therefore define the symmetrized combinations

$$\begin{aligned} &\varphi \text{ symmetric:} && \varphi_{\mathcal{S}} = \big[\varphi_1(\boldsymbol{r}_1) \varphi_2(\boldsymbol{r}_2) + \varphi_2(\boldsymbol{r}_1) \varphi_1(\boldsymbol{r}_2) \big] / \sqrt{2}, \\ &\varphi \text{ antisymmetric:} && \varphi_{\mathcal{A}} = \big[\varphi_1(\boldsymbol{r}_1) \varphi_2(\boldsymbol{r}_2) - \varphi_2(\boldsymbol{r}_1) \varphi_1(\boldsymbol{r}_2) \big] / \sqrt{2}. \end{aligned}$$

With these we can simultaneously diagonalize S^2, S_z and S_N and obtain wavefunctions Ψ with definite spin and definite symmetry. Here the spin determines the symmetry of χ and of φ:

$$\begin{aligned} S = 1 \Rightarrow \chi &= \text{symmetric}, \quad \varphi = \text{antisymmetric}; \\ S = 0 \Rightarrow \chi &= \text{antisymmetric}, \quad \varphi = \text{symmetric}. \end{aligned} \tag{2.36}$$

A similar principle holds in general for N spin-1/2 fermions; see Appendix B: from the Weyl character formula and the Pauli principle it follows that the total spin S fixes the *species* (the representation under the permutation group) of the orbital wavefunction. The species in turn essentially determines the value of the "exchange integrals," i.e. the energy of the state. Thus the total spin indirectly affects the energy of a state. Magnetic properties such as ferromagnetism are therefore consequences of the connection between spin and orbital species.

Obviously, fermions interfere with one another via the Pauli principle, and one may ask whether we must always antisymmetrize all electrons in the universe. Fortunately,

[4] It satisfies $\chi_\uparrow(s = \uparrow) = 1$ and $\chi_\uparrow(s = \downarrow) = 0$.

this is not necessary. Consider two noninteracting electrons in Aachen and Cologne, described by single-particle states $\varphi_A(\boldsymbol{r}_1, s_1)$ and $\varphi_K(\boldsymbol{r}_1, s_1)$ with $\int dx\, \varphi_A^* \varphi_K \approx 0$, which (correctly) are represented by the antisymmetric wavefunction

$$\Psi \approx \frac{1}{\sqrt{2}}\big[\varphi_A(x_1)\varphi_K(x_2) - \varphi_K(x_1)\varphi_A(x_2)\big] \tag{2.37}$$

. Let $A(x)$ be a local single-particle operator. Then its expectation value is given by

$$\begin{aligned}\langle A\rangle = \frac{1}{2}\int dx_1 dx_2 \Big[&\varphi_A^*(x_1)\varphi_K^*(x_2)\, A(x_1)\, \varphi_A(x_1)\varphi_K(x_2) \\ &+ \varphi_K^*(x_1)\varphi_A^*(x_2)\, A(x_1)\, \varphi_K(x_1)\varphi_A(x_2) \\ &-\varphi_A^*(x_1)\varphi_K^*(x_2)\, A(x_1)\, \varphi_K(x_1)\varphi_A(x_2) \\ &-\varphi_K^*(x_1)\varphi_A^*(x_2)\, A(x_1)\, \varphi_A(x_1)\varphi_K(x_2)\Big] \Big\} \approx 0 \\ \approx \langle A\rangle_A + \langle A\rangle_K.&\end{aligned} \tag{2.38}$$

It is the interference terms that carry the symmetry properties.[5] For nonoverlapping wavefunctions these terms vanish, and we may dispense with the (anti-)symmetrization: the results in Aachen depend only on the wavefunction in Aachen, and analogously for Cologne.

2.5.3 Wavefunction of two (spin-0) bosons

For two bosons the product wavefunction $\Psi_{\mu,\nu}$ must be symmetrized. The allowed wavefunction is therefore given by

$$\Psi^{\mathcal{S}}_{\mu,\nu}(x_1, x_2) = \sqrt{2}\mathcal{S}\Psi_{\mu,\nu} = \frac{1}{\sqrt{2}}\big[\varphi_\mu(x_1)\varphi_\nu(x_2) + \varphi_\mu(x_2)\varphi_\nu(x_1)\big]. \tag{2.39}$$

For bosons $\mu = \nu$ is now allowed, and the wavefunction of two bosons in the same state is given by (note the new normalization, cf. (2.53) later)

$$\Psi^{\mathcal{S}}_{\mu,\mu}(x_1, x_2) = \varphi_\mu(x_1)\varphi_\mu(x_2) \overset{x_1=x_2=x}{\longrightarrow} \varphi_\mu(x)^2 \neq 0, \tag{2.40}$$

and both bosons may occupy the same place.

Note: Fermions avoid one another, bosons like to crowd

$$\begin{aligned}\Psi^{\mathcal{S}/\mathcal{A}}_{1,2}(x_1, x_2) &= \frac{1}{\sqrt{2}}\Big[\varphi_1(x_1)\varphi_2(x_2) \mp \varphi_2(x_1)\varphi_1(x_2)\Big] \\ &\downarrow\, x_1 = x_2 = x \\ &= \begin{cases} 0, & \text{Fermionen,} \\ \sqrt{2}\varphi_1(x)\varphi_2(x), & \text{Bosonen.}\end{cases}\end{aligned} \tag{2.41}$$

□

[5] Note that for bosons the last two terms in (2.38) are added.

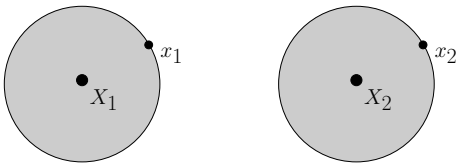


Figure 2.1: On the symmetry of the wavefunction of two hydrogen atoms: X_1, X_2 denote the coordinates of the protons, x_1, x_2 the coordinates of the electrons.

2.5.4 Composite particles

Consider two H atoms as sketched in figure 2.1. The total wavefunction $\Psi(x_1, X_1, x_2, X_2)$ must be antisymmetric in the electron coordinates x_1, x_2 and also antisymmetric in the coordinates X_1, X_2 of the protons. Exchanges of x_1, X_1; x_1, X_2; x_2, X_1; x_2, X_2 are not constrained. Since the particles being exchanged are not identical, such exchanges are physically meaningless. Exchanging the atoms (first the electrons, then the protons) yields

$$\begin{aligned}\Psi(x_2, X_2, x_1, X_1) &= -\Psi(x_1, X_2, x_2, X_1) \\ &= (-1)^2 \Psi(x_1, X_1, x_2, X_2).\end{aligned} \tag{2.42}$$

From this we obtain that H atoms, each composed of two fermions, behave like bosons. Note the consistency with the spin-statistics theorem from chapter 2.4,

$$\text{H atom} \Rightarrow \ \mathcal{H}_{1/2} \otimes \mathcal{H}_{1/2} = \mathcal{H}_0 \oplus \mathcal{H}_1; \tag{2.43}$$

so that hydrogen atoms possess integer spin and hence are bosons. The same reasoning applies to other composite particles, for example

$$\begin{aligned}{}^3\text{He} &= 2p + 1n + 2e = \text{fermion} \rightarrow \text{pairing transition}, \\ {}^4\text{He} &= 2p + 2n + 2e = \text{boson} \rightarrow \lambda - \text{transition};\end{aligned} \tag{2.44}$$

both He fluids become superfluid at low temperatures. Whereas ^{4}He "Bose-condenses" (at low temperatures $T < 2.17$ K all ^{4}He atoms occupy the same single-particle state), the fermionic ^{3}He atoms must first pair up ($\rightarrow$ bosons) and then, via a so-called BCS (Bardeen-Cooper-Schrieffer) transition at the lowest temperatures $T \sim 10$ mK, enter the superfluid state.

2.5.5 Symmetry under time evolution

Time evolution preserves the symmetry, for if

$$U_\pi \Psi = (\pm)^\pi \Psi, \qquad \begin{matrix}\text{bosons } (+), \\ \text{fermions } (-),\end{matrix} \tag{2.45}$$

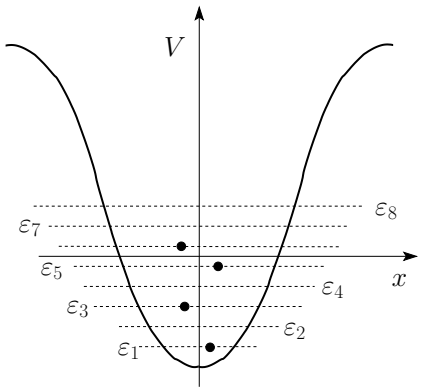


Figure 2.2: Occupation of a fermionic system in state Ψ_{1356}. The Pauli principle (the symmetry of the wave function) allows at most one particle per state.

then

$$\begin{aligned} U_\pi \Psi(t+dt) &= U_\pi \Psi(t) + U_\pi[\partial_t \Psi(t)]dt \\ &= (\pm)^\pi \Psi - iU_\pi(H/\hbar)\Psi dt \\ &\quad \downarrow_{[H,U_\pi]=0} \\ &= (\pm)^\pi \Psi - i(H/\hbar)U_\pi \Psi dt \qquad (2.46) \\ &= (\pm)^\pi[\Psi + (\partial_t \Psi)dt] = (\pm)^\pi \Psi(t+dt). \qquad (2.47) \end{aligned}$$

2.5.6 Many-particle states of non-interacting identical particles

For non-interacting particles we can write the Hamiltonian as

$$H = \sum_{i=1}^{N} H_0(x_i), \qquad x_i = (\boldsymbol{r}_i, s_i), \tag{2.48}$$

where H_0 denotes the single-particle Hamiltonian. With the single-particle solutions

$$H_0\varphi_\alpha = \varepsilon_\alpha \varphi_\alpha, \qquad \{\varphi_\alpha\} \text{ forming a complete orthonormal system (CONS)}, \tag{2.49}$$

we can define N-particle product eigenstates

$$\Psi(x_1, \ldots, x_N) = \varphi_{\alpha_1}(x_1)\varphi_{\alpha_2}(x_2)\cdots\varphi_{\alpha_N}(x_N) \tag{2.50}$$

of H. The total energy is then $E = \sum_i \varepsilon_{\alpha_i}$. Equation (2.50) describes a N-particle system in the states $\varphi_{\alpha_1}, \varphi_{\alpha_2}, \ldots \varphi_{\alpha_N}$, for example for $N = 4$ and V a potential well

$$\Psi_{1356} = \varphi_1(\boldsymbol{r}_1)\varphi_3(\boldsymbol{r}_2)\varphi_5(\boldsymbol{r}_3)\varphi_6(\boldsymbol{r}_4), \tag{2.51}$$

as shown in Figure 2.2.

Fermions: For fermions we must Ψ antisymmetrize,

$$\Psi \mapsto \Psi_{\mathcal{A}} = \frac{1}{\sqrt{N!}} \sum_{\pi \in S_N} (-1)^{\pi} \varphi_{\alpha_1}(x_{\pi(1)}) \cdots \varphi_{\alpha_N}(x_{\pi(N)})$$
$$= \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_{\alpha_1}(x_1) & \cdots & \varphi_{\alpha_1}(x_N) \\ \varphi_{\alpha_2}(x_1) & \cdots & \varphi_{\alpha_2}(x_N) \\ \vdots & \vdots & \\ \varphi_{\alpha_N}(x_1) & \cdots & \varphi_{\alpha_N}(x_N) \end{vmatrix} = \frac{1}{\sqrt{N!}} \det \phi_{\alpha_i}(x_j); \tag{2.52}$$

Here the factor $(N!)^{-1/2}$ guarantees the normalization of the many-body wave function. The determinant in expression (2.52) is called a Slater determinant. If there is a pair $i \neq j$ with $\alpha_i = \alpha_j$ then $\Psi_{\mathcal{A}} \equiv 0$. Thus the Pauli exclusion principle,[6] requires that no two fermions occupy the same state in both spin and orbital space.

Bosons: For bosons we must Ψ symmetrize the wave function,

$$\Psi \mapsto \Psi_{\mathcal{S}} = \sqrt{\frac{1}{N! \prod_j n_j!}} \sum_{\pi \in S_N} \varphi_{\alpha_1}(x_{\pi(1)}) \cdots \varphi_{\alpha_N}(x_{\pi(N)}); \tag{2.53}$$

it is a simple exercise to show that the prefactor guarantees the normalization of the wave function. The degeneracy n_j indicates how many of the indices α_i take the same value j (with $\sum_j n_j = N$).

2.5.7 Fermi Sea

In the simplest case, the electrons in a solid are described by

- Inner or valence electrons; these are electrons tightly bound to the atoms (or within the atomic lattice) and define fully occupied electron bands.
- Conduction electrons. These are quasi-free, mobile electrons in partially filled bands, cf. Figure 2.3.

The bands are filled up to the Fermi level. If the Fermi level lies within a band, the material is a metal; if the Fermi level lies in an energy gap, the material is a semiconductor or an insulator.

The conduction electrons can be most simply described as a gas of free spin-1/2 fermions with the Hamiltonian $H_0 = p^2/2m$. The ground state $|\Phi_0\rangle$ ("Fermi sea") has the form of a Slater determinant, where the low-energy states are occupied up to $k \leq k_F$, the Fermi sphere with radius k_F. The eigenstates of the single-particle Hamiltonian are plane waves $\phi_{\boldsymbol{k}} \propto e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$. In the specific case where the particles are confined to a volume $\Omega = L^3$ with periodic boundary conditions, one obtains the quantization condition $\boldsymbol{k} = (2\pi/L)\mathbb{Z}^3$. The number of states in the Fermi sphere is

$$N = \frac{\text{volume of the sphere}}{\text{volume per lattice site}} = g_s \frac{(4\pi/3)k_F^3}{(2\pi/L)^3} \tag{2.54}$$

[6] First postulated in 1925 by W. Pauli to explain the periodic table of the elements.

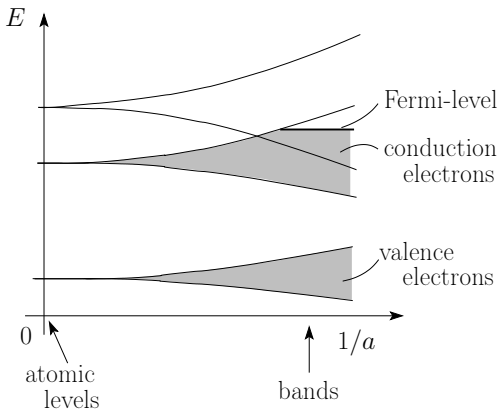


Figure 2.3: Electron states in a solid form into bands. Well-separated atoms with $a \to \infty$, a the lattice spacing, exhibit a discrete spectrum. At finite (small) separations, the electrons can hop between lattice sites, and energy bands form that are filled up to the Fermi level.

where the factor $g_s = 2$ accounts for spin degeneracy. [7] From (2.54) one obtains the relation between the electron density $n = N/\Omega$ and the parameters $\varepsilon_F = p_F^2/2m$ (Fermi energy), $p_F = \hbar k_F$ (Fermi momentum), $k_F = \sqrt{2m\varepsilon_F}/\hbar$ (Fermi wavenumber) of the Fermi gas, given by

$$k_F^3 = \frac{6\pi^2}{g_s} n \tag{2.55}$$

The corresponding ground-state energy is

$$E_0 = g_s \sum_{k<k_F} \frac{\hbar^2 k^2}{2m} = g_s \frac{L^3}{(2\pi)^3} \int_{k<k_F} d^3k \, \frac{\hbar^2 k^2}{2m} = N \frac{3}{5} \varepsilon_F. \tag{2.56}$$

This yields for the ground-state energy density $e_0 = 3\varepsilon_F n/5$, which corresponds to the Fermi pressure $p = \partial E_0/\partial\Omega = e_0$ of the quantum gas at 0 K. The ground-state energy $\propto N\varepsilon_F$ of N fermions is always high, since the particles avoid one another and thus occupy high-energy states. We find the following energy scale in typical metals: With

$$\begin{aligned} \varepsilon_F &= \frac{\hbar^2 k_F^2}{2m} = 3.81\,\text{eV} \cdot k_F^2\,[\text{Å}^{-2}] \quad \text{and} \\ k_F &\approx \frac{\pi}{a}, \qquad a = \text{ lattice constant } \approx 3\text{Å}, \end{aligned} \tag{2.57}$$

we obtain $k_F \sim 1$ Å^{-1} and $\varepsilon_F \sim$ several eV typically, that is several 10^4 K. Thus the kinetic energy of electrons in metals at room temperature is not 300 K but $\sim$ 30000 K. The Fermi pressure corresponds in this case to $\sim \varepsilon_F/a^3 \sim 10^6$ atm.

[7] If, instead of periodic boundary conditions, one assumes infinitely high potential walls such that $\phi_j(\boldsymbol{r})$ vanishes at the boundary, we obtain the single-particle wavefunctions $\phi_{\boldsymbol{k}'}(\boldsymbol{r}) = \prod_{i=1}^3 \sin k_i' r_i$ with $k_i = \pi n_i/L$ and $n_i \in \mathbb{N}$. The number of states in the Fermi sphere is then given by $N = g_s V_{k_F}^{>}/(\pi/L)^3$ where V_{k_F} is the volume of the Fermi sphere in the first octant with $V_{k_F}^{>} = (4\pi/3)k_F^3/8$. The result (2.54) (and hence all subsequent results) therefore remain unchanged.

2.5.8 Bose-Einstein condensate

The ground state of the bosonic many-body system (with spin 0) looks very different from the Fermi sea. The product state

$$\Psi_0 = \prod_{i=1}^{N} \varphi_0(\boldsymbol{r}_i) \tag{2.58}$$

with φ_0 the single-particle ground state is already symmetric and $E_0 = N\varepsilon_0$: the bosons condense into the lowest-energy state. This condensation phenomenon ultimately leads to the superfluidity of a Bose system at low temperatures. Fermions can condense only if they first pair up, for example in the form of Cooper pairs in the superconductivity of metals.[8]

2.5.9 Scattering of identical particles

Consider the collision/scattering of two identical particles. We separate into center-of-mass and relative coordinates $\boldsymbol{R} = \boldsymbol{r}_1 + \boldsymbol{r}_2$ and $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$. Thus $\boldsymbol{R}$ is symmetric and $\boldsymbol{r}$ antisymmetric in $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. We decompose the total wave function into spin and orbital parts,

$$\Psi(x_1, x_2) = e^{i\boldsymbol{P}\cdot\boldsymbol{R}}\,\psi(\boldsymbol{r})\chi(s_1, s_2). \tag{2.59}$$

For distinguishable particles the scattering solution has the asymptotic form

$$\psi(\boldsymbol{r}) \sim e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})\frac{e^{ikr}}{r}. \tag{2.60}$$

Spin-0 bosons: For bosons with spin-0 $\chi = 1$ and by the symmetry of Ψ $\psi(\boldsymbol{r}) = \psi(-\boldsymbol{r})$; accordingly we must symmetrize the scattering solution and obtain

$$\psi \sim \left(e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right) + [f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}}) + f_{\boldsymbol{k}}(\Omega_{-\boldsymbol{r}})]\frac{e^{ikr}}{r}. \tag{2.61}$$

For the differential cross section we obtain

$$\begin{aligned}\frac{d\sigma}{d\Omega_{\boldsymbol{r}}} &= \left|f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}}) + f_{\boldsymbol{k}}(\Omega_{-r})\right|^2 \\ &= \underbrace{\left|f_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})\right|^2 + \left|f_{\boldsymbol{k}}(\Omega_{-\boldsymbol{r}})\right|^2}_{\text{classical terms}} + \underbrace{2\mathrm{Re}\left[f^*_{\boldsymbol{k}}(\Omega_{\boldsymbol{r}})f_{\boldsymbol{k}}(\Omega_{-\boldsymbol{r}})\right]}_{\text{interference terms}}.\end{aligned} \tag{2.62}$$

The classical terms are the sum of the cross sections for detecting either of the two particles (the detector cannot distinguish which particle arrives).

[8] This pairing phenomenon is of a somewhat more complex nature and we refer to Chap. 5.2 and to the corresponding standard literature on BCS (Bardeen-Cooper-Schrieffer) theory.

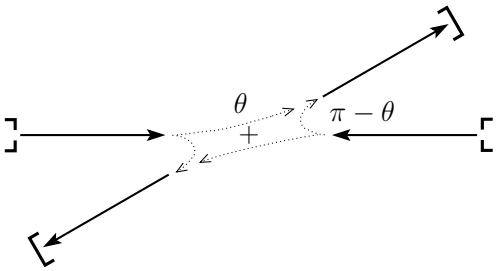


Figure 2.4: The symmetrization of the scattering wave function produces two trajectories with scattering angles θ and $\pi - \theta$ that must be added coherently.

The interference terms arise as a consequence of the symmetrization required by the particle statistics.

For a centrally symmetric potential $V(r)$, the scattering amplitude $f_{\boldsymbol{k}}$ depends only on the scattering angle θ. The inversion $\boldsymbol{r} \mapsto -\boldsymbol{r}$ in polar coordinates is given by $\theta \mapsto \pi - \theta, r \mapsto r$. This simplifies Eq. (2.62) to

$$\frac{d\sigma}{d\Omega_{\boldsymbol{r}}} = \left|f_{\boldsymbol{k}}(\theta) + f_{\boldsymbol{k}}(\pi-\theta)\right|^2 = \left|f_{\boldsymbol{k}}(\theta)\right|^2 + \left|f_{\boldsymbol{k}}(\pi-\theta)\right|^2 + 2\mathrm{Re}\left[f^*_{\boldsymbol{k}}(\theta) f_{\boldsymbol{k}}(\pi-\theta)\right].$$

The angles involved are shown in Figure 2.4. Due to the interference terms, for bosons in the case $\theta = \pi/2$ the cross section doubles compared to the classical result,

$$\theta = \frac{\pi}{2}: \qquad \frac{d\sigma}{d\Omega_{\boldsymbol{r}}} = 4\left|f_{\boldsymbol{k}}(\pi/2)\right|^2. \tag{2.63}$$

In the partial-wave expansion

$$\begin{aligned} f_{\boldsymbol{k}}(\theta) &= \sum_l i^l (2l+1) f_l P_l(\cos\theta) \\ \downarrow P_l(\cos\theta) &= (-1)^l P_l(\cos(\pi-\theta)) \\ f_{\boldsymbol{k}}(\theta) + f_{\boldsymbol{k}}(\pi-\theta) &= 2\sum_{l \text{ even}} i^l (2l+1) f_l P_l(\cos\theta), \end{aligned} \tag{2.64}$$

we find that only even angular momenta l occur (for odd l the Legendre polynomial changes sign and the contributions cancel each other out).

Spin-1/2 Fermions: For spin-1/2 fermions two cases are possible:

1. The spin-singlet state $\chi_s = \left[(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right]/\sqrt{2}$ is antisymmetric and consequently the orbital part

$$\psi(\boldsymbol{r}) = \psi(-\boldsymbol{r}) \tag{2.65}$$

must be symmetric. The cross section equals that of spin-0 bosons,

$$\left.\frac{d\sigma}{d\Omega_{\boldsymbol{r}}}\right|_s = \left|f_{\boldsymbol{k}}(\theta) + f_{\boldsymbol{k}}(\pi-\theta)\right|^2. \tag{2.66}$$

2. The symmetric spin-triplet states

$$\chi_t = \begin{cases} |\uparrow\uparrow\rangle \\ \big[|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\big]/\sqrt{2} \\ |\downarrow\downarrow\rangle \end{cases} \tag{2.67}$$

require an antisymmetric orbital wave function $\psi(\boldsymbol{r}) = -\psi(-\boldsymbol{r})$ and we obtain a scattering amplitude $f_{\boldsymbol{k}}(\theta) \to f_{\boldsymbol{k}}(\theta) - f_{\boldsymbol{k}}(\pi - \theta)$ that contains only odd angular momenta l. Hence the cross section becomes

$$\frac{d\sigma}{d\Omega_{\boldsymbol{r}}}\Big|_t = \big|f_{\boldsymbol{k}}(\theta) - f_{\boldsymbol{k}}(\pi-\theta)\big|^2 \qquad (\,=0 \text{ for } \theta = \pi/2). \tag{2.68}$$

Note that polarized fermions scatter only in odd angular-momentum channels: cold bosonic atoms exhibit a contact potential and scatter mainly in the s-wave channel, cf. (2.64)]. Cold polarized fermionic atoms, by contrast, interact much more weakly, since only the p-wave channel appears (the cross section σ_l behaves as $\sim a^2(ak)^{4l}$ at low energies).

For a statistical ensemble of an unpolarized beam the weighted average (from 3 triplet states and 1 singlet state)

$$\begin{aligned} \frac{d\sigma}{d\Omega_{\boldsymbol{r}}} &= \frac{3}{4}\frac{d\sigma}{d\Omega_{\boldsymbol{r}}}\Big|_t + \frac{1}{4}\frac{d\sigma}{d\Omega_{\boldsymbol{r}}}\Big|_s \\ &= \big|f_{\boldsymbol{k}}(\theta)\big|^2 + \big|f_{\boldsymbol{k}}(\pi-\theta)\big|^2 - \mathrm{Re}\,\big[f^*_{\boldsymbol{k}}(\theta) f_{\boldsymbol{k}}(\pi-\theta)\big]. \end{aligned} \tag{2.69}$$

Chapter 3

Atoms

3.1 Atom with two electrons

To begin, we consider the two-electron atom, helium He ($Z = 2$) or the ions H^- ($Z = 1$), Li^+ ($Z = 3$) and Be^{++} ($Z = 4$), described by the Hamiltonian

$$H = H_0 + H_{\mathrm{WW}} = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} - \frac{Ze^2}{r_1} - \frac{Ze^2}{r_2} + \frac{e^2}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}. \tag{3.1}$$

Here $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ denote the positions of the electrons relative to the atomic nucleus and $e > 0$ is the elementary charge. We have assumed that the nucleus does not move, which is a reasonable approximation given the large mass difference between the electrons and the nucleus. The Hamiltonian (3.1) is not exactly solvable because of the interaction term $H_{\mathrm{WW}} = e^2/r_{12}$, $r_{12} = |\boldsymbol{r}_1 - \boldsymbol{r}_2|$, between the two electrons. The simplest approximate solution treats this term perturbatively, i.e., we first solve the non-interacting problem $H_0\Psi_0 = E_0\Psi_0$ and then determine the correction $E_1 = \langle\Psi_0|e^2/r_{12}|\Psi_0\rangle$. The solutions of $H_0\Psi_0 = E_0\Psi_0$ are the product wavefunctions $\Psi_0(\boldsymbol{r}_1, \boldsymbol{r}_2) = \varphi_{\alpha_1}(\boldsymbol{r}_1)\varphi_{\alpha_2}(\boldsymbol{r}_2)$, where $\varphi_\alpha(\boldsymbol{r})$ is a hydrogenic wavefunction for nuclear charge Z with $\alpha = nlm$. The corresponding energies are

$$E_0 = \varepsilon_{\alpha_1} + \varepsilon_{\alpha_2} \tag{3.2}$$

with

$$\varepsilon_\alpha = -\frac{Z^2}{n^2}E_R, \qquad E_R = \frac{e^2}{2a_B} = 1\,\mathrm{Ry} \approx 13.6\,\mathrm{eV}, \qquad a_B = \frac{\hbar^2}{me^2} \approx 0.529\,\text{Å}. \tag{3.3}$$

In addition, we must take into account the Pauli principle for fermions. The total wavefunction $\Psi(x_1, x_2)$ is antisymmetric in $x_i = (\boldsymbol{r}_i, s_i)$,

$$\Psi(x_1, x_2) = -\Psi(x_2, x_1). \tag{3.4}$$

Separating spin and orbital parts yields

$$\Psi(x_1, x_2) = \Psi_0(\boldsymbol{r}_1, \boldsymbol{r}_2)\chi_0(s_1, s_2), \tag{3.5}$$

where the antisymmetry of the total wavefunction can be constructed in two ways, either with

$$\begin{cases} \Psi_0 & \text{symmetric,} \\ \chi_0 & \text{antisymmetric,} \quad (S = 0, \quad \text{singlet}) \end{cases} \tag{3.6}$$

i.e., an antisymmetric spin function (spin-singlet state), or via

$$\begin{cases} \Psi_0 & \text{antisymmetric,} \\ \chi_0 & \text{symmetric,} \qquad (S = 1, \quad \text{triplet}) \end{cases} \tag{3.7}$$

with a symmetric spin function (spin-triplet state). With Ψ_0 antisymmetric, $\alpha_1 \neq \alpha_2$ must hold, which costs us at least the additional energy

$$\Delta E = Z^2(1 - 1/2^2)E_R \approx Z^2 \times 10\,\text{eV} \tag{3.8}$$

We therefore choose the spin-singlet state as the candidate for the ground state with

$$\begin{aligned} \Psi_0(x_1, x_2) &= \varphi_{100}(\boldsymbol{r}_1)\,\varphi_{100}(\boldsymbol{r}_2)\,\chi_{00}(s_1, s_2), \\ \varphi_{100}(r) &= (Z^3/\pi a_B^3)^{1/2}\exp(-Zr/a_B), \\ \chi_{00}(s_1, s_2) &= [\chi_\uparrow(s_1)\chi_\downarrow(s_2) - \chi_\downarrow(s_1)\chi_\uparrow(s_2)]/\sqrt{2}, \\ E_0 &= -2Z^2E_R = -Z^2e^2/a_B. \end{aligned} \tag{3.9}$$

The correction E_1 follows from the integral

$$\begin{aligned} E_1 &= \int d^3r_1\, d^3r_2\, \frac{Z^6e^2}{a_B^6\pi^2}\frac{\exp[-2Z(r_1 + r_2)/a_B]}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} \\ &= \frac{5}{8}\frac{Ze^2}{a_B}. \end{aligned} \tag{3.10}$$

We obtain the ground-state energy

$$E_G = E_0 + E_1 = -\Big(Z - \frac{5}{8}\Big)\frac{Ze^2}{a_B}. \tag{3.11}$$

to first order in perturbation theory. This energy provides an upper bound for the true ground-state energy, since first-order perturbation theory is equivalent to the Rayleigh-Ritz variational principle with the unperturbed ground-state wavefunction as the trial function.

For the He atom ($Z = 2$) we obtain $E_G \approx -5.5\,\text{Ry}$. The experimentally measured value is $E_G^{\text{He}} \approx -5.8\,\text{Ry}$, about 4 eV lower than our calculated value. Using a

variational ansatz $Z \mapsto \alpha$ (α variational parameters)[1] we find the estimate

$$E_G = \min_\alpha \Big(\alpha - 2Z + \frac{5}{8}\Big)\frac{\alpha\, e^2}{a_B} = -\Big(Z - \frac{5}{16}\Big)^2 \frac{e^2}{a_B} \qquad (\alpha{=}Z{-}5/16)$$
$$\approx -5.7 \text{ Ry}, \tag{3.12}$$

an improvement by $\approx 2.7\,\text{eV}$; about $1.4\,\text{eV}$ of the total $79\,\text{eV}$ are still missing, a respectable result. The improvement brought by $Z \mapsto \alpha$ is a consequence of accounting for screening effects: each electron "sees" only a charge $2 - 5/16 = 27/16$ instead of $Z = 2$.

For the hydrogen ion H^- ($Z = 1$) we obtain

$$E_G = -\Big(\frac{11}{16}\Big)^2 \frac{e^2}{a_B} \approx -12.86\,\text{eV}. \tag{3.13}$$

The configuration $H + e^-$ corresponds to the energy -13.6 eV, so that H^- would be unstable; experimentally, however, H^- is stable, which means that our result must be off by at least $(3/4)\,\text{eV}$.

3.2 Hartree and Hartree-Fock approximations

Turning to larger atoms, we are confronted with a complicated electronic many-body problem. The general form of this many-body problem can be described by the Hamiltonian

$$H = \sum_{i=1}^{N} \Big[\frac{p_i^2}{2m} + U(\boldsymbol{r}_i)\Big] + \underbrace{\frac{1}{2}\sum_{i\neq j} \overbrace{\frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}}^{V_{\text{int}}^{ij}}}_{V_{\text{int}}} \tag{3.14}$$

(here we neglect interactions that involve the spin). The potential $U(\boldsymbol{r})$ is a given (external) potential; for atoms of charge Z we have $U(\boldsymbol{r}) = -Ze^2/r$. Naively, we might again first solve the single-particle problem

$$h = \frac{p^2}{2m} + U(\boldsymbol{r}), \qquad\qquad h\varphi_j(\boldsymbol{r}) = \varepsilon_j \varphi_j(\boldsymbol{r}), \tag{3.15}$$

and take the many-body function as an antisymmetrized product wavefunction [i.e., as a Slater determinant (2.52)] and treat the interaction term in the Hamiltonian as a perturbation. The expected results are mediocre, as we have seen in the example of the He atom. The problem is that the electrons are not only subject to the potential $U(\boldsymbol{r})$, but also move in the (average) field of the other electrons, which leads to a screening of U. This screening in turn depends on the electron orbits, so

[1] One should replace $Z \mapsto \alpha$ only in the wavefunction, not in the Hamiltonian!

that we must determine the orbits and the screening effects self-consistently. This is precisely the achievement of the Hartree and Hartree-Fock approximations: the Hartree approximation accounts for (only) the screening effects and not for the exchange effects due to the fermionic nature of the electrons. The full Hartree-Fock theory additionally incorporates the antisymmetry of the wavefunction.

3.2.1 Hartree approximation

In the Hartree approximation, the task is to find the optimal product wavefunction

$$\Psi_{\mathrm{H}} = \varphi_1(\boldsymbol{r}_1)\,\varphi_2(\boldsymbol{r}_2)\cdots\varphi_N(\boldsymbol{r}_N). \tag{3.16}$$

We can at least partially take into account the Pauli principle by orthogonalizing φ_i and occupying each state at most twice (because of $g_s = 2$). Mathematically, we seek single-particle functions φ_i such that the expectation value of the energy $\langle H\rangle_{\mathrm{H}} = \langle\Psi_{\mathrm{H}}|H|\Psi_{\mathrm{H}}\rangle$ is minimized; the wavefunctions are required to be normalized, $\|\varphi_i\| = 1$. With these constraints, which we incorporate via Lagrange multipliers, we must minimize the following functional:[2]

$$\begin{aligned}\mathcal{F}[\varphi_1(x_1),\ldots,\varphi_N(x_N)] &= \langle H\rangle_{\mathrm{H}} - \sum_{i=1}^{N}\lambda_i\langle\varphi_i|\varphi_i\rangle \\ &= \int\prod_{i=1}^{N} d^3r_i\,(\varphi_1^*\cdots\varphi_N^* H\varphi_1\cdots\varphi_N) - \sum_{i=1}^{N}\lambda_i\int dr_i^3\,\varphi_i^*\varphi_i \\ &= \sum_{i=1}^{N}\int dr_i^3\,\varphi_i^*\left[\frac{p_i^2}{2m} + U(\boldsymbol{r}_i) + \frac{1}{2}{\sum_j}'\int d^3r_j\,\varphi_j^*\frac{e^2}{|\boldsymbol{r}_i-\boldsymbol{r}_j|}\varphi_j - \lambda_i\right]\varphi_i.\end{aligned} \tag{3.17}$$

At the minimum of $\mathcal{F}$ the variation $\delta\mathcal{F} = 0$ vanishes, and we obtain the condition $(\delta\mathcal{F}/\delta\varphi_i^* = 0)$

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + U(\boldsymbol{r}) + \int d^3r'\frac{e}{|\boldsymbol{r}-\boldsymbol{r}'|}{\sum_j}'e|\varphi_j(\boldsymbol{r}')|^2\right]\varphi_i(\boldsymbol{r}) = \lambda_i\varphi_i(\boldsymbol{r}). \tag{3.18}$$

The sum $n_i(\boldsymbol{r}') = {\sum_j}'|\varphi_j(\boldsymbol{r}')|^2$ describes the average charge distribution $-e\,n_i(\boldsymbol{r}')$ of the other electrons $j \neq i$. The system (3.18) defines the Hartree equations; the optimization of the single-particle wavefunctions φ_i accounts for the presence of the other electrons through the interaction term

$$V_{\mathrm{H}}(\boldsymbol{r}) = \int d^3r'\,\frac{e^2 n_i(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}. \tag{3.19}$$

[2]The symbol $\sum_j'$ denotes the sum over j excluding the term $j = i$.

Note that the φ_i are in general not orthogonal, since $n_i(\boldsymbol{r})$ depends on i. One therefore usually writes $n_i(\boldsymbol{r}) \approx n(\boldsymbol{r})$ for the particle density, since an electron represents only a small perturbation of the $N \gg 1$ electron cloud. The density $-e\,n(\boldsymbol{r})$ then describes the charge distribution that an external test charge would perceive, and the φ_i are orthogonal. The (Hartree) expectation value of the ground-state energy $E_{\mathrm{H}} = \langle H\rangle_{\mathrm{H}} = \sum_i \lambda_i - \langle V_{\mathrm{int}}\rangle_{\mathrm{H}}$ provides an upper bound for the exact ground-state energy.

The Lagrange multipliers λ_i appear in the Hartree equation as single-particle energies ε_i. To determine their physical meaning, we multiply (3.18) by φ_i^* and integrate over $\boldsymbol{r}$. We obtain

$$\lambda_i = \int d^3r \left[\frac{p^2}{2m} + U(\boldsymbol{r})\right] |\varphi_i(\boldsymbol{r})|^2 + {\sum_j}' \int d^3r\, d^3r' \, \frac{e^2 |\varphi_i(\boldsymbol{r})|^2 |\varphi_j(\boldsymbol{r}')|^2}{|\boldsymbol{r}-\boldsymbol{r}'|}. \tag{3.20}$$

Since λ_i accounts precisely for those terms of $\langle H\rangle_{\mathrm{H}}$ that contain φ_i (see (3.17)), $-\lambda_i$ corresponds to the ionization energy under the assumption that the other wavefunctions do not change (Koopmans' theorem). In reality, the other wavefunctions relax, i.e., the relaxed state $\Psi_{\mathrm{H},i}^{(N-1)}$ (with $N-1$ particles) contains modified single-particle wavefunctions. However, the "knocking out" of the ith electron, for example in a photoemission experiment, leaves behind an exchange hole that is filled only after a finite relaxation time. A fast ionization process (e.g., via photoemission) therefore involves the unrelaxed ionization energy $-\lambda_i$ and not the relaxed quantity $\lambda_i^{\mathrm{ion}} = \langle H\rangle_{\mathrm{H}} - \langle \Psi_{\mathrm{H},i}|H|\Psi_{\mathrm{H},i}\rangle$.

3.2.2 Hartree-Fock approximation

The Hartree-Fock theory goes one step further and minimizes

$$\langle \Psi_{\mathrm{HF}}|H|\Psi_{\mathrm{HF}}\rangle - \sum_i \lambda_i \langle \varphi_i|\varphi_i\rangle \tag{3.21}$$

in the variational space spanned by the Slater determinant[3]

$$\Psi_{\mathrm{HF}} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_1(\boldsymbol{r}_1)\chi_{\sigma_1}(s_1) & \cdots & \varphi_1(\boldsymbol{r}_N)\chi_{\sigma_1}(s_N) \\ \vdots & \ddots & \vdots \\ \varphi_N(\boldsymbol{r}_1)\chi_{\sigma_N}(s_1) & \cdots & \varphi_N(\boldsymbol{r}_N)\chi_{\sigma_N}(s_N) \end{vmatrix}. \tag{3.22}$$

The single-particle expectation values of p_i^2 and $U(\boldsymbol{r}_i)$ remain unchanged, while the interaction term now produces two terms,

$$\langle V_{\mathrm{int}}\rangle_{\mathrm{HF}} = \frac{1}{2}\int d^3r\, d^3r' \, \frac{e^2}{|\boldsymbol{r}-\boldsymbol{r}'|} \sum_{i\neq j} \Big[\varphi_i^*(\boldsymbol{r})\varphi_j^*(\boldsymbol{r}')\varphi_j(\boldsymbol{r}')\varphi_i(\boldsymbol{r}) - \delta_{\sigma_i\sigma_j}\varphi_i^*(\boldsymbol{r})\varphi_j^*(\boldsymbol{r}')\varphi_j(\boldsymbol{r})\varphi_i(\boldsymbol{r}')\Big]. \tag{3.23}$$

[3] Note that the Slater determinant Ψ_{HF} in fact depends only on the variational space $\mathrm{span}\{\varphi_1,\ldots,\varphi_N\}$ and not on the choice of the basis functions φ that span this space.

The spin-dependent term $\propto \delta_{\sigma_i\sigma_j}$ is called the Fock or exchange term.

Its derivation involves the following steps: starting from the product state $\Psi_{\mathrm{H}} = \prod_i \varphi_i(\boldsymbol{r}_i)\chi_{\sigma_i}(s_i)$ we define the antisymmetrized state $\Psi_{\mathrm{HF}} = \sqrt{N!}\mathcal{A}\Psi_{\mathrm{H}} = 1/\sqrt{N!}\sum_{\pi\in S_N}(-1)^\pi U_\pi \Psi_{\mathrm{H}}$ and compute the matrix element[4]

$$\begin{aligned}
\langle V_{\mathrm{int}}\rangle_{\mathrm{HF}} &= N!\langle\Psi_{\mathrm{H}}\mathcal{A}V_{\mathrm{int}}\mathcal{A}\Psi_{\mathrm{H}}\rangle \\
&\downarrow \qquad\qquad (\mathcal{A}V_{\mathrm{int}}\mathcal{A}=V_{\mathrm{int}}\mathcal{A}^2=V_{\mathrm{int}}\mathcal{A}, \quad \text{since } [\mathcal{A},V_{\mathrm{int}}]=0) \\
&= \frac{1}{2}\sum_{j\neq i}\sum_{\pi\in S_N}(-1)^\pi\langle\Psi_{\mathrm{H}}V_{\mathrm{int}}^{ij}U_\pi\Psi_{\mathrm{H}}\rangle \quad \text{[only } \pi=1 \text{ and } \pi=(i\leftrightarrow j) \text{ do not vanish]} \\
&= \frac{1}{2}\sum_{i\neq j}\langle\Psi_{\mathrm{H}}|V_{\mathrm{int}}^{ij}[1-(i\leftrightarrow j)]|\Psi_{\mathrm{H}}\rangle\,. \qquad (3.24)
\end{aligned}$$

We have $\langle\chi_{\mathrm{H}}|\chi_{\mathrm{H}}\rangle = 1$ and

$$\langle\chi_{\mathrm{H}}|(i\leftrightarrow j)|\chi_{\mathrm{H}}\rangle = \delta_{s_i s_j}. \qquad (3.25)$$

with $\chi_{\mathrm{H}} = \prod_i \chi_{\sigma_i}$ acting on the spin space. The two terms in (3.24) then describe precisely the Hartree and exchange contributions in equation (3.23). Analogous to (3.18), we vary $\delta\mathcal{F}/\delta\varphi_i^* = 0$ and thereby obtain the Hartree-Fock integro-differential equations

$$\begin{aligned}
&\left[-\frac{\hbar^2}{2m}\nabla^2 + U(\boldsymbol{r}) + \int d^3r'\,\frac{e^2 n_i(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\right]\varphi_i(\boldsymbol{r}) \qquad (3.26)\\
&\quad - \sum_j \delta_{\sigma_i\sigma_j}\int d^3r'\left[\varphi_j^*(\boldsymbol{r}')\frac{e^2}{|\boldsymbol{r}-\boldsymbol{r}'|}\varphi_i(\boldsymbol{r}')\right]\varphi_j(\boldsymbol{r}) = \lambda_i\varphi_i(\boldsymbol{r}).
\end{aligned}$$

In the second line, we integrate over the sought eigenfunction φ_i; accordingly, the electron is subject to a nonlocal potential. The formulas (3.18) and (3.26) must be solved numerically by iteration.

Note that the Hartree-Fock approximation still does not include correlation effects[5] and therefore neither (3.18) nor (3.26) is exact. The ground-state energy $E_{\mathrm{HF}} = \langle H\rangle_{\mathrm{HF}}$ of the system includes the terms

$$E_{\mathrm{HF}} = \underbrace{\underbrace{E_0}_{\text{single-particle}} + E_{\mathrm{WW}}}_{\text{Hartree}} + \underbrace{E_{\text{Austausch.}}}_{\text{Hartree-Fock}} \qquad (3.27)$$

[4] See (2.26) for the definition of the antisymmetrization operator.

[5] Correlation energy: An electron at position $\boldsymbol{r}$ reduces the probability of finding another electron there. In particular, an electron with spin $|\uparrow\rangle$ creates an exchange and correlation hole for other electrons with spin $|\uparrow\rangle$. The HF theory accounts for the exchange part; any contributions to the exchange hole beyond the mean-field HF theory are referred to as correlation effects.

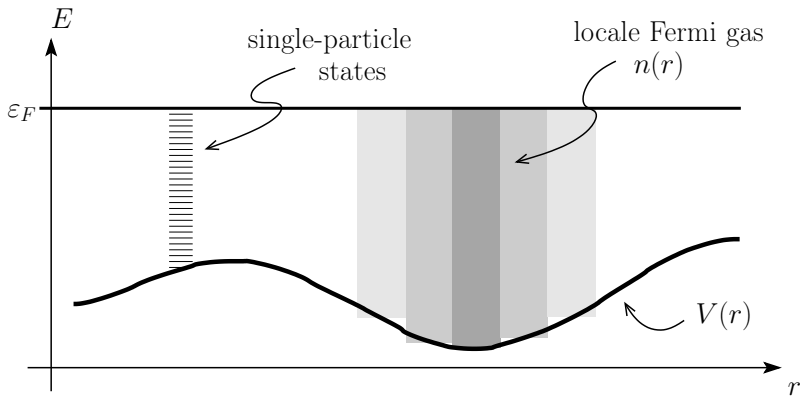


Figure 3.1: Thomas-Fermi screening: The inhomogeneous Fermi gas is approximated by a local Fermi gas of density $n(\boldsymbol{r}) = k_F^3(\boldsymbol{r})/3\pi^2$. The position-dependent Fermi wavevector $k_F(\boldsymbol{r})$ follows from the local potential $V(\boldsymbol{r})$ taking screening effects into account, cf. (3.30).

In addition, there is a correlation energy $E_{\text{Korrelation}}$ that one finds within many-body physics/quantum field theory via a systematic expansion in perturbation theory. For atoms, the Hartree-Fock approximation yields results accurate to within typical correlation energies of order $\simeq 0.5\,\text{eV}$ per electron.

3.3 Thomas-Fermi screening

We consider electrons in a slowly varying potential $U(\boldsymbol{r})$, which we may treat as locally constant, cf. Figure 3.1. The electrons then form a local Fermi gas with density [cf. (2.55) with $g_s = 2$]

$$n(\boldsymbol{r}) = \frac{k_F^3(\boldsymbol{r})}{3\pi^2}. \tag{3.28}$$

A test charge at position $\boldsymbol{r}$ experiences a local potential composed of the "bare" potential $U(\boldsymbol{r})$ and the potential generated by the inhomogeneous electron gas,

$$V(\boldsymbol{r}) = U(\boldsymbol{r}) + \int d^3r' \, \frac{e^2\, n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|}. \tag{3.29}$$

The second term is precisely the Hartree potential that we encountered in the previous chapter.

The electrons in the Fermi gas distribute themselves in equilibrium so that the addition of a test electron at any position $\boldsymbol{r}$ costs the same energy ε_F; thus

$$\varepsilon_F = \frac{\hbar^2 k_F^2(\boldsymbol{r})}{2m} + V(\boldsymbol{r}); \tag{3.30}$$

where the first term describes the kinetic energy of the added electron while the second term describes its potential energy $V(\boldsymbol{r})$. Combining (3.28) and (3.30) yields the relation between the local density $n(\boldsymbol{r})$ and the effective potential $V(\boldsymbol{r})$,

$$n(\boldsymbol{r}) = \frac{1}{3\pi^2}\left[\frac{2m}{\hbar^2}\big(\varepsilon_F - V(\boldsymbol{r})\big)\right]^{3/2}; \tag{3.31}$$

note that together with (3.29) this is an integral equation for the density $n(\boldsymbol{r})$. Equation (3.31) for $n(\boldsymbol{r})$ can be turned into a differential equation for $V(\boldsymbol{r})$: applying the Laplace operator ∇^2 to (3.29) and using $\nabla^2\, 1/|\boldsymbol{r}-\boldsymbol{r}'| = -4\pi\delta^{(3)}(\boldsymbol{r}-\boldsymbol{r}')$ gives

$$\begin{aligned}\nabla^2 V(\boldsymbol{r}) &= \nabla^2 U(\boldsymbol{r}) - 4\pi e^2 n(\boldsymbol{r}) \\ &= -4\pi e^2\left[-\rho_{\text{ext}}(\boldsymbol{r})/e + n(\boldsymbol{r})\right],\end{aligned} \tag{3.32}$$

where $\rho_{\text{ext}}(\boldsymbol{r})$ is the charge density generating the external potential $U(\boldsymbol{r})$ and $n(\boldsymbol{r})$ is the screening electron density. With the usual definition of the dielectric constant

$$\varepsilon(\boldsymbol{q}) = \frac{U(\boldsymbol{q})}{V(\boldsymbol{q})} \tag{3.33}$$

we obtain [with $q^2[V(\boldsymbol{q}) - U(\boldsymbol{q})] = 4\pi e^2 n(\boldsymbol{q})$ from (3.32)]

$$\varepsilon(\boldsymbol{q}) = \frac{q^2(V - V + U)}{q^2 V} = 1 - \frac{(V-U)q^2}{q^2 V} = 1 - \frac{4\pi e^2}{q^2}\frac{n}{V}; \tag{3.34}$$

for small $V(\boldsymbol{q})$, $\boldsymbol{q} \neq 0$ the screening (induced) charge density is given by $n \approx (\partial n/\partial V)V = (-\partial n/\partial \varepsilon_F)_{V=0} V$. Inserting into (3.33) yields the Thomas-Fermi result for the dielectric constant of a metal,

$$\varepsilon(\boldsymbol{q}) = 1 + \frac{4\pi e^2}{q^2}\frac{\partial n}{\partial \varepsilon_F} \equiv 1 + \frac{q_{\text{TF}}^2}{q^2}. \tag{3.35}$$

To see that the result (3.35) describes screening, we compute the inverse transform of (3.33)

$$V(\boldsymbol{r}) = \int d^3r' \frac{e^{-q_{\text{TF}}|\boldsymbol{r}-\boldsymbol{r}'|}}{4\pi|\boldsymbol{r}-\boldsymbol{r}'|} U(\boldsymbol{r}'). \tag{3.36}$$

Indeed, the external potential U is screened on the length scale q_{TF}^{-1} The Thomas-Fermi wavevector is determined by the electron density via k_F,

$$q_{\text{TF}}^2 = 4\pi e^2 \frac{\partial n}{\partial \varepsilon_F} = \frac{4e^2 m k_F}{\pi\hbar^2} = \frac{4}{\pi}\frac{k_F}{a_B}. \tag{3.37}$$

In a typical metal $k_F \approx 1\,\text{Å}^{-1}$ and hence $q_{\text{TF}} \simeq 1\,\text{Å}^{-1}$: a charge in a metal is screened by other electrons within a distance of $\simeq 1\,\text{Å} = 10^{-10}\,\text{m}$.

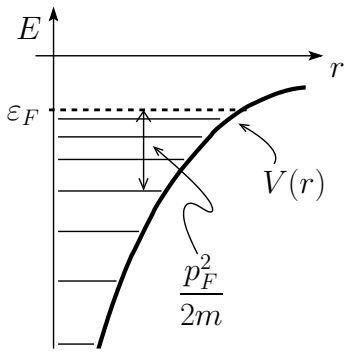


Figure 3.2: Potential with bound states in the Thomas-Fermi atom. The states are filled up to the Fermi level ε_F; their charge, via screening, contributes to the effective potential.

3.4 Thomas-Fermi atom

We wish to study atoms with nuclear charge $Z \gg 1$. The solution of the hydrogen problem (with a single electron) yields the radius $r_K = a_B/Z$. This result applies to the description of the innermost $1s$ shell, because there the nuclear potential is still "bare" and not yet screened by the presence of other electrons. For the outermost electron (in the valence shell), by contrast, only an effective nuclear charge $Z_{\text{eff}} \approx 1$ remains. Thus the outermost shell has radius $R = a_B$ (independent of the nuclear charge). To treat the electron cloud of the "middle" shells with typical radius $\bar{r}$, we must take screening effects into account more carefully. The simplest model for this is the Thomas-Fermi atom.

We apply the Thomas-Fermi screening principle to the treatment of atoms and start from (3.32) with $\rho_{\text{ext}}(\boldsymbol{r})/e = Z\delta^{(3)}(\boldsymbol{r})$ (due to the nuclear charge). We approximate $V(\boldsymbol{r})$ as spherically symmetric and obtain for $r > 0$

$$\nabla^2 V(r) = \frac{1}{r^2}\partial_r r^2 \partial_r V(r) = -\frac{4e^2}{3\pi\hbar^3}\Big[2m\big(\varepsilon_F - V(r)\big)\Big]^{3/2}. \tag{3.38}$$

For $r \to 0$ the screening must vanish and the electrons should be subject to the "bare" nuclear potential with $V(r) \sim -Ze^2/r$; this provides a boundary condition for (3.38). The atomic radius R is fixed by the condition $n(R) = 0$. From this follows with (3.31) the second boundary condition $V(R) = \varepsilon_F$. For a neutral atom $V(r > R) = \text{konst.} = 0$ and hence $\varepsilon_F = 0$; for a charged ion we have $V(r > R) = -(Z - N)e^2/r$, with N the number of electrons, cf. Figure 3.2.

In the following, we consider the neutral case with $N = Z$ and define the dimensionless scales x and Φ for the length and for the energy (the potential),

$$r = \left(\frac{3\pi}{4\sqrt{Z}}\right)^{2/3} \frac{a_B x}{2} = \beta \frac{a_B x}{Z^{1/3}}, \quad V(r) = -\frac{Ze^2}{r}\Phi(x) \qquad (\text{with } \beta = 0.8853). \tag{3.39}$$

The universal atom[6] is then characterized by a dimensionless potential $\Phi(x)$ that

[6]The potential $\Phi(x)$ is independent of Z.

solves the Thomas-Fermi equation

$$\sqrt{x}\,\frac{d^2\Phi(x)}{dx^2} = \Phi(x)^{3/2}, \qquad \Phi(0) = 1, \tag{3.40}$$

The potential satisfies the asymptotic behavior

$$\Phi(x) = \begin{cases} 1 - 1.59\,x, & x \to 0, \\ 144/x^3, & x \to \infty, \end{cases} \tag{3.41}$$

and otherwise must be found by numerical integration of the differential equation (3.40). Note that the potential of the universal Thomas-Fermi atom extends to infinity since $\Phi(x)$ decays only algebraically.

In fact, the Thomas-Fermi approximation is applicable only for intermediate radii (with $r_{\min} \ll r \ll r_{\max}$). The local-density approximation assumes that the potential does not change strongly on the (local) Fermi wavelength. Here the typical scale d of potential variation is given by the distance r, since $d \simeq V/\partial_r V \simeq r$. On the other hand, the Fermi wavevector is approximately given by $k_F^2 \simeq 2mV(r)/\hbar^2 \simeq Z\Phi(x)/a_B r$. One thus obtains the condition

$$1 \ll (k_F d)^2 \qquad \Leftrightarrow \qquad 1 \ll Z\,\Phi(x)\,r/a_B \tag{3.42}$$

for the applicability of the Thomas-Fermi approximation. For small distances we have $\Phi(x) \approx 1$ and hence $r_{\min} \simeq r_K = a_B/Z$. For large distances we use $\Phi(x) \sim x^{-3} \simeq (a_B/r)^3/Z$ and find that $r_{\max} \simeq R = a_B$ (independent of Z). Thus the Thomas-Fermi approximation holds from the innermost shell $r_K = a_B/Z$ to the outer shell $R = a_B$. In this range there are

$$\begin{aligned} \int_{r_K \le r \le R} d^3r\, n(r) &= -\frac{1}{e^2}\int_{r_K}^{R} dr\, r^2\, \nabla^2 V(\boldsymbol{r}) = \left.\frac{r^2 \partial_r V(r)}{e^2}\right|_{r=R}^{r_K} \\ &\simeq -\frac{r_K\, V(r_K)}{e^2} \simeq Z\Phi(1) \propto Z \end{aligned}$$

electrons. Hence most of the electrons (except those in the outer and inner shells) contribute to the formation of the Thomas-Fermi atom.

We find the following scaling laws as a function of Z for the Thomas-Fermi atom:

- Radius: $\bar{r} \simeq Z^{-1/3} a_B$, heavy atoms are smaller,[7]
- Potential: $V \simeq Ze^2\Phi(x)/r \simeq Z^{4/3} E_R \Phi(x)/x \propto Z^{4/3}$ at fixed $x \propto Z^{1/3} r$,
- e^- density: $n \simeq \nabla^2 V/e^2 \simeq V/e^2 r^2 \simeq Z\Phi(x)/r^3 \simeq (Z^2/a_B^3)\Phi(x)/x^3 \propto Z^2$ at fixed x,
- Kinetic energy: $p^2/2m = \hbar^2 k_F^2/2m \simeq \hbar^2 n^{2/3}/m \simeq Z^{4/3} E_R$,

[7] Numerically one finds that 50% of the atoms lie within a radius $1.33\, Z^{-1/3} a_B$.

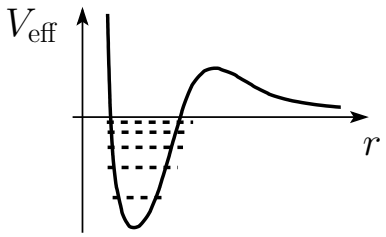


Figure 3.3: Effective potential of the Thomas-Fermi atom including the centrifugal barrier. The TF atom binds a state with angular momentum l if V_{eff} has a minimum with negative energy.

- Velocity:[8] $v \simeq p/m \simeq Z^{2/3}e^2/\hbar \propto Z^{2/3}$.

The electrons in the Thomas-Fermi atom arrange themselves in shells. The Thomas-Fermi atom binds a state with angular momentum l if the effective potential [including the centrifugal potential $\hbar^2 l(l+1)/2mr^2$]

$$V_{\text{eff}}(r) = V(r) + \frac{\hbar^2 l(l+1)}{2mr^2} = \frac{Z^{2/3}e^2}{\beta\, a_B}\left[-Z^{2/3}\frac{\Phi(x)}{x} + \frac{l(l+1)}{2\beta\, x^2}\right] \tag{3.43}$$

has a local minimum with $V_{\text{eff}} < 0$, cf. Figure 3.3.

With the TF potential $\Phi(x)$ according to (3.40) one finds for the Thomas-Fermi atom shells in the following l and Z ranges:

	$l = 0$	$l \le 1$	$l \le 2$	$l \le 3$
TF-Atom	$Z \le 4$	$5 \le Z \le 19$	$20 \le Z \le 53$	$54 \le Z$
	only s	s & p	s, p & d	s, p, d & f
Experimental		p-electrons	d-e^-	f-e^-
		for $Z \ge 5$	for $Z \ge 20$	for $Z \ge 58$

Shells in the Thomas-Fermi atom

Note that because of $V_{\text{eff}}(r) \neq Ze^2/r$ the single-particle energies ε_i in the Thomas-Fermi approximation depend on n and l (but, of course, not on m owing to radial symmetry). In particular, in an atom there are shells nl at energy ϵ_{nl} with degeneracy $(2l+1)(2s+1) = 4l+2$, i.e.

$$s\text{-shell (ent. 2)}, \qquad p\text{- (ent. 6)}, \qquad d\text{- (ent. 10)} \quad \text{and} \quad f\text{- (ent. 14)}. \tag{3.44}$$

Determining the shell sequence (increase of l or of n) requires a numerical analysis. One finds the sequence

$$\begin{aligned} 1s < 2s < 2p < 3s < 3p < 4s < 3d < 4p < 5s < 4d < 5p \\ < 6s < 4f < 5d < 6p < 7s < 5f. \end{aligned} \tag{3.45}$$

This arrangement can be memorized with the help of the scheme 3.4. Correspondingly one finds the periodic system of the elements in the form 3.5.

We go through the periodic table with some comments:

[8] We have $e^2/\hbar = \alpha c$ with $\alpha \approx 1/137$ the fine-structure constant.

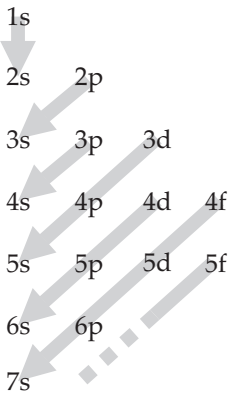


Figure 3.4: Sequence of shells nl for the energy of the last-occupied electron in the Thomas-Fermi atom. n = principal quantum number, l = orbital angular-momentum quantum number. Note that the energetically deep-lying/inner electrons exhibit a different sequence (namely the trivial one): the sequence of the so-called X-ray shells is $1s$ (the 1K-shell), $2s$, $2p$ (2L-shell), $3s$, $3p$, $3d$ (3M-shell), $4s$, $4p$, $4d$, $4f$ (4N-shell), etc.

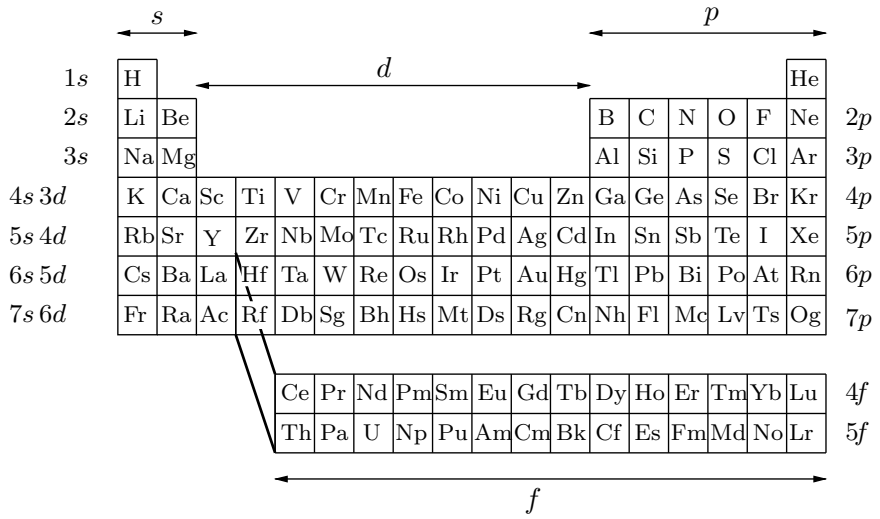


Figure 3.5: Periodic table of the elements.

- H, He: the $1s$ shell is filled
- Li, Be: the $2s$ shell; Be-Ne: the $2p$ shell is filled. $2s$ and $2p$ have different energies; orbitals with small l lie closer to the nucleus and therefore have lower energies, $\varepsilon_{2s} < \varepsilon_{2p}$.
- Na-A: as before, first the $3s$, then the $3p$ shell is filled.
- K-Kr: the energies of $4s$ are nearly equal to those of $3d$. One obtains the complicated filling (we give only the number of electrons in the $4s$ shell):

  K $4s^1$, Ca $4s^2$, $\cdots$, V $4s^2$, Cr $4s^1$, Mn $4s^2$, $\cdots$, Ni $4s^2$, Cu $4s^1$, Zn $4s^2$

  With Ga-Kr the p-shell is filled. The $3d$-shells are compact and therefore chemically less important; they do, however, provide the magnetic moments that generate the ferro- or antiferromagnetism of the elements Cr-Ni.

- Rb-Xe: essentially a repetition of the K-Kr sequence. Again the s orbitals have roughly the same energy as the d orbitals of the preceding shell, leading to the irregular filling

  $$\text{Rb } 5s^1,\ \text{Sr } 5s^2,\ \cdots,\ \text{Zr } 5s^2,\ \text{Nb } 5s^1, \cdots,\ \text{Rh } 5s^1,\ \text{Pd } 4d^{10}\,5s^0,$$
  $$\text{Ag } 4d^{10}\,5s^1,\ \text{Cd } 4d^{10}\,5s^2$$

  . With In-Xe the p-shell is filled.

- Cs-Rn: Cs and Ba fill the $6s$ orbital. La fills the first $5d^1$ orbital, then follow the lanthanides (rare earths) La-Lu with the filling of the $4f$ shell, which lies energetically close to $5d$. Again there are shifts in the occupations,

  $$\text{Ce } 5d^0\,4f^2, \cdots,\ \text{Eu } 5d^0\,4f^7,\ \text{Gd } 5d^1\,4f^7,\ \text{Tb } 5d^0\,4f^9, \cdots,\ \text{Yb } 5d^0\,4f^{14}.$$

  The compact f orbitals are passive in chemical bonds but yield large magnetic moments, especially for Ce, Pr, Nd, Sm-Yb. After filling the f-shell the $5d$ and $6p$ shells are filled; Pt and Au again show a rearrangement with $5s^1$.

- Fr-Rg: Fr and Ra fill the $7s$ orbital; Ac has the configuration $6d^1$, then follow the actinides with the filling of the f orbital. The energies for $6d$ and $5f$ lie close together and again there are corresponding rearrangements with $6d^{0,1}$ up to Lr and $6d^{2,3}$ for Rf and Db. There are no stable elements with filled $6d$, $7p$ orbitals.

For the chemical properties of the elements the outermost electrons are relevant. Correspondingly, elements in one column of the periodic table have similar chemical properties. For example,

- the noble gases He-Rn with their full orbitals are very inert,
- the alkali metals (H), Li-Fr with one electron, and
- the halogens F-At with one "hole" are very reactive.

3.5 Hund's rules

For electrons in partially filled p, d, or f shells the question arises how these electrons arrange themselves in the ground state. This question can be illustrated with the carbon atom ($1s^2\,2s^2\,2p^2$) with two $2p$ electrons. The electrons in the p shell have the six states $\{\phi_{p_x}, \phi_{p_y}, \phi_{p_z}\} \otimes \{\chi_\uparrow, \chi_\downarrow\}$ available. The two electrons can "choose" among $\binom{6}{2} = 15$ different configurations. The energy of the states is determined by the Hamiltonian $H = H_0 + H_{\text{WW}} + H_{\text{SO}}$, where

$$H_0 = \sum_{i=1}^{N} \left[\frac{p_i^2}{2m} + V_{\text{H}}(\boldsymbol{r}_i) - \frac{Ze^2}{r_i} \right] \tag{3.46}$$

with $V_{\text{H}}(\boldsymbol{r})$ the effective potential in the Hartree approximation, see Eq. (3.19). Because of rotational symmetry the $2p$ states of H_0 are degenerate.

The remaining terms partially lift this degeneracy: the (residual) interaction energy is, to leading order, given by

$$H_{\rm WW} = \frac{1}{2}\sum_{i\neq j}\frac{e^2}{|\boldsymbol{r}_i-\boldsymbol{r}_j|} - \sum_{i=1}^{N} V_{\rm H}(\boldsymbol{r}_i) \tag{3.47}$$

and the (relativistic) spin-orbit (Russell-Saunders or LS coupling) has the form

$$H_{\rm SO} = \frac{1}{2m^2c^2}\frac{1}{r}\frac{dV}{dr}\sum_{i=1}^{N}\boldsymbol{l}_i\cdot\boldsymbol{s}_i \tag{3.48}$$

with $\boldsymbol{l}_i$ and $\boldsymbol{s}_i$ the orbital angular momentum and spin of the i-th electron. Adding these energies splits the 15 states. This raises the question which of the resulting states is the ground state and what its residual degeneracy is.

To answer this question we treat the terms $H_{\rm WW}$ and $H_{\rm SO}$ in perturbation theory. $H_0 + H_{\rm WW}$ is still $SO(3)$-symmetric and spin-independent. Hence $\boldsymbol{L} = \sum_i \boldsymbol{l}_i$ and $\boldsymbol{S} = \sum_i \boldsymbol{s}_i$ are good quantum numbers.[9] The spin-orbit coupling destroys this symmetry and $\boldsymbol{L}$, $\boldsymbol{S}$ are no longer separately conserved. Therefore one passes to the total angular momentum $\boldsymbol{J}$ with

$$\boldsymbol{J} = \sum_i \boldsymbol{j}_i, \qquad \boldsymbol{j}_i = \boldsymbol{l}_i + \boldsymbol{s}_i, \tag{3.49}$$

which is still conserved.

3.5.1 Construction of the perturbation theory

To describe the splitting of the LS multiplets in perturbation theory we need to know which of the two perturbation terms $H_{\rm WW}$ and $H_{\rm SO}$ is the more important. Hence we begin with a rough estimate of the two terms (each per electron in neutral atoms with $N = Z$):[10]

$$E_{\rm WW} \simeq \frac{\langle (H_{\rm WW})^2\rangle^{1/2}}{Z} \simeq \frac{e^2}{\bar r} \simeq 0.1\frac{e^2}{a_B}Z^{1/3} \simeq \text{eV}\, Z^{1/3} \tag{3.50}$$

$$E_{\rm SO} \simeq \frac{\langle H_{\rm SO}\rangle}{Z} \simeq \frac{1}{2m^2c^2}\Big\langle \underbrace{\frac{1}{r}\frac{dV}{dr}}_{Ze^2/\bar r^3}\underbrace{\boldsymbol{l}\cdot\boldsymbol{s}}_{\hbar^2}\Big\rangle \qquad [\text{mit}\ \ \bar r = a_B/Z^{1/3}]$$

$$\simeq \underbrace{\frac{\hbar^2}{2ma_B^2}}_{\simeq \text{Ry}}\ \underbrace{\frac{1}{mc^2}}_{\simeq 1/\text{MeV}}\ \underbrace{\frac{e^2}{a_B}}_{\simeq \text{Ry}} Z^2 \simeq 10^{-3}\,\text{eV}\, Z^2. \tag{3.51}$$

[9] Note that the total wave function Ψ must be antisymmetric, i.e. the mixing of the spin wave functions implies that the individual spins $\boldsymbol{s}_i$ are not good quantum numbers. For example, for $n = 2$ the singlet spin function $\chi_S = [|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle]/\sqrt{2}$ is $S_z\chi_S = 0$ but $s_{1z}\chi_S \not\propto \chi_S$.

[10] One has $\langle H_{\rm WW}\rangle = 0$, since the single-particle states satisfy the Hartree equation. Hence the interaction is a fluctuation term and we obtain the scaling $\langle (H_{\rm WW})^2\rangle \simeq Z^2(e^2/\bar r)^2$ (we have Z^2 terms of order $(e^2/\bar r)^2$). The numerical prefactor is estimated from the He result with $Z = 2$.

The relative size of the two corrections is therefore given by

$$\begin{aligned} E_{\mathrm{WW}} &> E_{\mathrm{SO}}, \quad \text{for} \quad Z < 10^{9/5} \approx 60, \\ E_{\mathrm{WW}} &< E_{\mathrm{SO}}, \quad \text{for} \quad Z > 60; \end{aligned} \tag{3.52}$$

a more precise estimate places the crossing of the two energies at $Z \simeq 80$, which corresponds to the atomic number of lead Pb. We now discuss the case of light atoms with $E_{\mathrm{WW}} > E_{\mathrm{SO}}$.

3.5.2 Atoms lighter than Pb

We first address the problem $H_0 + H_{\mathrm{WW}}$ and then treat the spin-orbit coupling H_{SO} as a perturbation. The strategy is as follows: without H_{SO}, L and S remain good quantum numbers, so we find the values allowed by the configuration and the combinations compatible with the Pauli principle. The highly degenerate configuration level is then split by the differing interactions for the various L, S combinations into LS multiplets. Next we discuss the further splitting of the LS multiplets by H_{SO}, which leads to the fine structure. The remaining $2J+1$-fold degeneracy in the fine structure, arising from $SO(3)$ symmetry, is finally lifted by a magnetic field.

The symmetry of the problem is $SO(3)$. We combine the degenerate configuration states into states of sharp total angular momentum and spin $\boldsymbol{L}^2$, L_z, $\boldsymbol{S}^2$, S_z and identify the lowest-energy state. This state is still $(2L+1)(2S+1)$-fold degenerate. In our concrete example of carbon C with configuration $1s^2\,2s^2\,2p^2$, the $2p$ orbitals can be combined into

$$L_{\max} = 2 \text{ bis } L_{\min} = 0 \qquad (\mathcal{H}_1 \otimes \mathcal{H}_1 = \mathcal{H}_0 \oplus \mathcal{H}_1 \oplus \mathcal{H}_2), \tag{3.53}$$

$$S_{\max} = 1 \text{ bis } S_{\min} = 0 \qquad (\mathcal{H}_{1/2} \otimes \mathcal{H}_{1/2} = \mathcal{H}_0 \oplus \mathcal{H}_1). \tag{3.54}$$

The Pauli principle then allows the combinations

$$\begin{array}{llllll} L=2 & \text{symmetrisch} & \mapsto & S=0 & \text{antisymm.} & \mapsto J=2, \\ L=0 & \text{symmetrisch} & \mapsto & S=0 & \text{antisymm.} & \mapsto J=0, \\ L=1 & \text{antisymm.} & \mapsto & S=1 & \text{symmetrisch} & \mapsto J=0,1,2; \end{array}$$

the so-called LS multiplets. To characterize them one uses the term symbol

$$^{2S+1}L_J. \tag{3.55}$$

Here S is the spin, L the orbital angular momentum, and J the total angular momentum of the atom; the latter can take different values for a fixed LS multiplet.

Under H_{WW} the configuration states split. In Figure 3.6 this splitting is illustrated for carbon C. The sequence of this splitting is nontrivial—the L, S, (J) quantum numbers for the atomic ground state are given by Hund's rules, see Section 3.5.4.

In the next step we consider the spin-orbit coupling H_{SO}. H_{SO} couples $\boldsymbol{L}$ and $\boldsymbol{S}$, and only the total angular momentum $\boldsymbol{J}$ remains as a good quantum number. Hence we

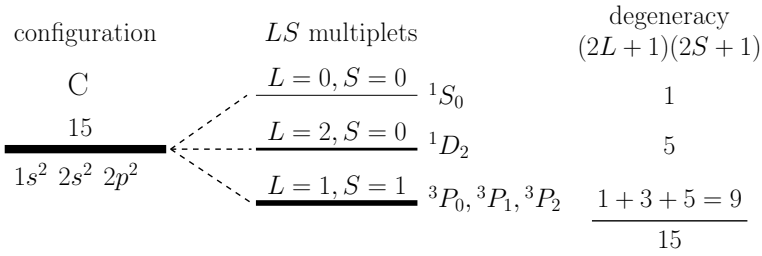


Figure 3.6: Configuration of the carbon atom C with splitting into LS multiplets and their degeneracies. The splitting of the highly degenerate configuration level involves that part of the Coulomb energy not already accounted for in the single-particle energies.

pass in the $(2L+1)(2S+1)$-dimensional LS multiplet to the $\boldsymbol{J}$ basis (the "quantum number" K symbolizes the unspecified electron configuration) with

$$|K\,L\,S\,M_L\,M_S\rangle \mapsto |K\,L\,S\,J\,M\rangle \tag{3.56}$$

When we express H_{SO} in the LS-multiplet basis, we obtain [11]

$$a\sum_{i=1}^{N} \boldsymbol{l}_i\cdot\boldsymbol{s}_i \mapsto A\,\boldsymbol{L}\cdot\boldsymbol{S} = \frac{A}{2}\left(\boldsymbol{J}^2-\boldsymbol{L}^2-\boldsymbol{S}^2\right) \tag{3.57}$$

$$= \frac{A}{2}\big[J(J+1)\underbrace{-L(L+1)-S(S+1)}_{\text{constant in the } LS\text{-multiplet}}\big] \tag{3.58}$$

where A depends on L and S. One finds that A is positive in our case, so that $J=0$ corresponds to the lowest energy state. Thus we obtain the splitting of the LS multiplet into levels separated by typical energies $\simeq (\mathrm{Ry}^2/mc^2)Z^2\cdot J$ in the sub-eV range. The term scheme for carbon with the LS splitting and the fine structure from spin-orbit coupling is sketched in Figure 3.7.

The remaining $(2J+1)$-fold degeneracy can be lifted by a (weak) magnetic field; see the discussion in Section 3.6. The resulting Zeeman splitting for the carbon atom is sketched in Figure 3.8.

3.5.3 Atoms heavier than Pb

For heavy atoms we first treat H_0+H_{SO} and consider H_{WW} as a perturbation. We combine the single-particle orbitals $|n\,l\,m\,s\rangle = \varphi_{nlm}\chi_s = R_{nl}Y_{lm}\chi_s$ to fixed l and s into those of fixed $\boldsymbol{j}$, $j=l\pm 1/2$, and obtain the new orbitals $|n\,l\,j\,m_j\rangle$. The new orbitals already incorporate the spin-orbit interaction at the single-particle level. Finally we take H_{WW} into account by combining states of fixed total angular

[11] We will argue later that $a\sum_i \boldsymbol{l}_i\cdot\boldsymbol{s}_i = A\,\boldsymbol{L}\cdot\boldsymbol{S}$. A proof follows from the Wigner-Eckart theorem, see G. Baym.

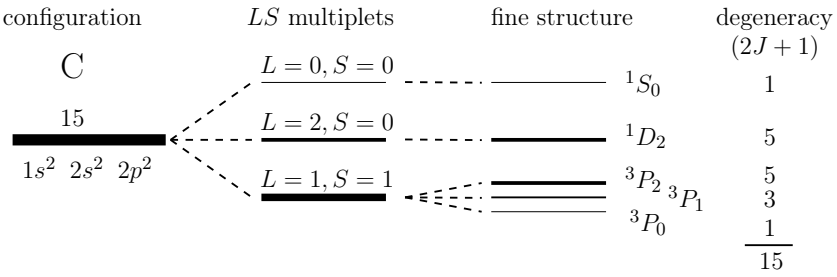


Figure 3.7: Configuration of the carbon atom C with splitting into LS multiplets and their fine structure (splitting on the order of 10 meV). The latter is induced by the (relativistic) spin-orbit coupling, which mixes orbital and spin angular momenta and thus respects only the total angular momentum J as a symmetry quantum number.

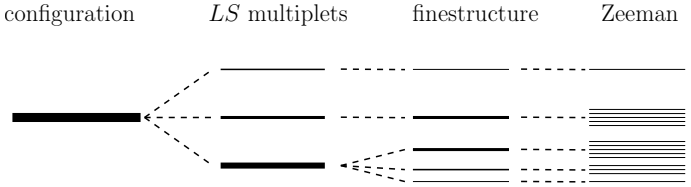


Figure 3.8: Configuration of the carbon atom C with splitting into LS-multiplets, their fine structure, and the Zeeman splitting in the magnetic field. The corresponding Hamiltonian with the various splitting terms is given by $H_0 + H_{\mathrm{ww}} + H_{\mathrm{so}} + \mu_B \sum_i \boldsymbol{H} \cdot (\boldsymbol{l}_i + 2\boldsymbol{s}_i)/\hbar$. In addition to the splitting in the external magnetic field (Zeeman), the nuclear spin also produces an effective magnetic field for the electrons, leading to the hyperfine structure.

momentum $\boldsymbol{J}$[12] and H_{ww} and diagonalize in each of these $(2J+1)$-dimensional spaces, thereby obtaining a ground state of fixed j_i^2, J^2, J_z.

The exact calculation of the term schemes and of the ground state ${}^{2S+1}L_J$ involves a complicated computation. On the other hand, for light atoms the task of finding the ground-state configurations (or the corresponding term symbol ${}^{2S+1}L_J$) can easily be solved by the empirical Hund's rules.

3.5.4 Hund's rules

1. Closed shells make no contribution to the orbital/spin angular momenta L and S.
2. The LS-multiplet with the largest S has the lowest energy.
3. For several L with equal S the largest L has the lowest energy.
4. If the shell is less than half full or exactly half full

[12] Correspondingly we construct Slater determinants of fixed J rather than of fixed L & S.

$$J = |L - S|, \text{ i.e. minimal,}$$

if the shell is more than half full, then

$$J = L + S, \text{ i.e. maximal.}$$

These empirical rules are supported by experimental observations and the following arguments:

1. In closed shells the momenta cancel to zero and the ground-state configuration corresponds to the term symbol ${}^1S_0$.
2. If the total spin S is maximal, then the spin function χ_{Spin} is symmetric; the orbital function φ_{Bahn} must, by the Pauli principle, be antisymmetric, so that the electrons avoid one another. This configuration accordingly yields the smallest Coulomb energy.
3. If the orbital angular momentum L is maximal, then the radii of the electron orbits are large and the electron density is correspondingly small, which in turn minimizes the Coulomb energy.
4. For spin-orbit coupling we first consider the case where the shell is half full or less than half full. Then all spins are parallel, i.e. in this Hilbert subspace

$$a \sum_{i=1}^{N} \boldsymbol{l}_i \cdot \boldsymbol{s}_i \approx a \sum_{i=1}^{N} \boldsymbol{l}_i \cdot \frac{\boldsymbol{S}}{N} \approx \frac{a}{N} \boldsymbol{L} \cdot \boldsymbol{S} \tag{3.59}$$

so that $A \approx a/N > 0$ and $J = |L - S|$ minimize the energy ($2S = N =$ number of electrons in the shell). For shells more than half full we can apply the following reasoning: we add the contribution of the missing (unoccupied) spins so as formally to obtain a completely filled shell, and then subtract that contribution again. Since the filled shell has vanishing spin-orbit coupling (spin singlet), only the subtracted term $-a \sum_i \boldsymbol{l}_i \cdot \boldsymbol{s}_i$ remains. Now the spins and angular momenta can be viewed as "hole" spin and angular momentum with $\boldsymbol{S} = -\sum_i \boldsymbol{s}_i$ and $\boldsymbol{L} = -\sum_i \boldsymbol{l}_i$. As these are again aligned, we obtain $A \approx -a/N < 0$ with (2S = N = number of missing electrons in the shell). Hence naturally $J = L + S$ is the total angular momentum of minimal energy.

These arguments are based on the viewpoint for light atoms. Hund's rules however apply to all atoms. We apply them here to two examples:

Nitrogen (N) with the electron configuration $1s^2\,2s^2\,2p^3$ ($Z = 7$): this is a half-filled p shell. (2) the maximum spin is $S = 3/2$ and (3) the maximum possible angular momentum is $L = 0$ ($= 1 + 0 - 1$). From this, with spin-orbit coupling (4) $J = |L - S| = 3/2$. The resulting configuration is ${}^4S_{3/2}$.

Promethium (Pm) with the electron configuration $[\text{Xe}]\,6s^2\,4f^5$ ($Z = 61$): this is a less-than-half-filled shell. (2) the maximum spin is $S = 5 \times 1/2 = 5/2$ and (3) the maximum possible angular momentum is $L = 5$ ($= 3 + 2 + 1 + 0 - 1$). (4) spin-orbit coupling yields $J = |L - S| = 5/2$. The configuration is therefore ${}^6H_{5/2}$.

3.6 Zeeman Splitting in a Magnetic Field

We consider atoms in a weak ($< 10^5$ gauss) magnetic field $\boldsymbol{H} = (0, 0, H_z)$; the term generated by the magnetic field (with the elementary charge $e > 0$; it is $H_Z = -\boldsymbol{\mu} \cdot \boldsymbol{H}$, $\boldsymbol{\mu}_L = -(e/2mc)\boldsymbol{L}$, $\boldsymbol{\mu}_S = -g_e(e/2mc)\boldsymbol{S}$, $g_e = 2$)

$$H_Z = \frac{e}{2mc}(\boldsymbol{L} + 2\boldsymbol{S}) \cdot \boldsymbol{H} = \frac{e}{2mc}(\boldsymbol{J} + \boldsymbol{S}) \cdot \boldsymbol{H} \tag{3.60}$$

is treated in perturbation theory. The $(2J + 1)$-fold degeneracy remaining in the fine structure is lifted by the magnetic field. The term proportional to $\boldsymbol{J}$ in (3.60) gives the contribution $e\hbar M H_z/2mc$ with M of the magnetic quantum number to J_z. It remains to compute $\langle S_z \rangle = \langle K\, L\, S\, J\, M | S_z | K\, L\, S\, J\, M \rangle$. It turns out that $\langle S_z \rangle \propto \langle J_z \rangle = M$. The result follows directly from the Wigner-Eckart theorem (which we do not wish to introduce here). For an alternative proof, we use the fact that $\boldsymbol{L} \cdot \boldsymbol{S}$ commutes with $\boldsymbol{J}$, so that the states $|K\, L\, S\, J\, M\rangle$ are eigenstates of $\boldsymbol{L} \cdot \boldsymbol{S}$ [the result follows immediately from $\boldsymbol{L} \cdot \boldsymbol{S} = \frac{1}{2}(\boldsymbol{J}^2 - \boldsymbol{L}^2 - \boldsymbol{S}^2)$]. Note also the following relation

$$[\boldsymbol{S}, \boldsymbol{L} \cdot \boldsymbol{S}] = \boldsymbol{S}(\boldsymbol{L} \cdot \boldsymbol{S}) - (\boldsymbol{L} \cdot \boldsymbol{S})\boldsymbol{S} \overset{13}{=} \boldsymbol{L} \times (\boldsymbol{S} \times \boldsymbol{S}) = i\hbar \boldsymbol{L} \times \boldsymbol{S}, \tag{3.61}$$

where we have used that $\boldsymbol{S} \times \boldsymbol{S} = i\hbar \boldsymbol{S}$ (angular-momentum algebra) and $[\boldsymbol{L}, \boldsymbol{S}] = 0$. By the vector product with $\boldsymbol{J}$ (from the right) we obtain

$$\begin{aligned}(\boldsymbol{S} \times \boldsymbol{J})(\boldsymbol{L} \cdot \boldsymbol{S}) - (\boldsymbol{L} \cdot \boldsymbol{S})(\boldsymbol{S} \times \boldsymbol{J}) &= i\hbar(\boldsymbol{L} \times \boldsymbol{S}) \times \boldsymbol{J} \overset{14}{=} i\hbar[\overbrace{\boldsymbol{S}(\boldsymbol{L} \cdot \boldsymbol{J})}^{+\boldsymbol{S}(\boldsymbol{S}\cdot\boldsymbol{J})} - \overbrace{\boldsymbol{L}(\boldsymbol{S} \cdot \boldsymbol{J})}^{+\boldsymbol{S}(\boldsymbol{S}\cdot\boldsymbol{J})}] \\ &= i\hbar[\boldsymbol{S}\boldsymbol{J}^2 - \boldsymbol{J}(\boldsymbol{S} \cdot \boldsymbol{J})]. \end{aligned} \tag{3.62}$$

Since $|K\, L\, S\, J\, M\rangle$ is an eigenstate of $\boldsymbol{L} \cdot \boldsymbol{S}$, the expectation value of (3.61) (and hence also of (3.62)) must vanish. It follows that

$$\underbrace{\langle \boldsymbol{S}\boldsymbol{J}^2 \rangle}_{=\hbar^2 J(J+1)\langle \boldsymbol{S} \rangle} = \langle \boldsymbol{J}(\boldsymbol{S} \cdot \boldsymbol{J}) \rangle. \tag{3.63}$$

With $\boldsymbol{J} \cdot \boldsymbol{S} = \boldsymbol{S}^2 + \boldsymbol{L} \cdot \boldsymbol{S} = \frac{1}{2}(\boldsymbol{J}^2 + \boldsymbol{S}^2 - \boldsymbol{L}^2)$ we obtain the desired relation

$$\langle S_z \rangle = \langle J_z \rangle \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}\,. \tag{3.64}$$

For the Zeeman splitting we find

$$\langle H_Z \rangle = g\,\frac{e\hbar}{2mc} M H_z = g\mu_B M H_z \tag{3.65}$$

[13] For operators the Grassmann identity $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b}(\boldsymbol{a} \cdot \boldsymbol{c}) - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c}$ holds if $[a_j, b_k] = 0$.

[14] Here we use the Grassmann identity in the form $(\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c} = \boldsymbol{b}(\boldsymbol{a} \cdot \boldsymbol{c}) - \boldsymbol{a}(\boldsymbol{b} \cdot \boldsymbol{c})$ if $[a_j, b_k] = 0$.

with the gyromagnetic factor g [also Landé g factor]

$$g = 1 + \frac{J(J+1) - L(L+1) + S(S+1)}{2J(J+1)}. \tag{3.66}$$

The elementary moment of the atom is

$$g\,\mu_B = g\frac{e\hbar}{2mc}. \tag{3.67}$$

For $S = 0$ we have $L = J$ and $g = 1$; conversely, for $L = 0$ the total angular momentum is given by the spin, $J = S$ and $g = 2$, which corresponds to the g factor of the electron spin. Note that $g < 1$ is also possible; e.g. promethium $g = 2/7$.[15] The LSJ multiplet splits into $2J + 1$ sublevels with

$$\Delta E_Z = \mu_B g\, H_z, \tag{3.68}$$

Since the g factor depends on the quantum numbers LSJ, complicated line splittings can arise in the spectrum. One distinguishes between the normal and the anomalous Zeeman effect, the anomalous Zeeman effect being the generic case. In the normal Zeeman effect the special case occurs that for both involved levels $S = 0$ and hence $g = 1$, so that an optical line splits into three components and one observes a "Zeeman triplet." In the anomalous Zeeman effect the g factors of the two levels differ. Correspondingly, more complicated line splittings result. In Figure 3.9 the two cases are shown for the Na D_1 transition [$^2S_{1/2}$ ($g = 2$) - $^2P_{1/2}$ ($g = 2/3$)], anomalous Zeeman effect, and for the Cd transition $^1P_1$ ($g = 1$) - $^1D_2$ ($g = 1$), normal Zeeman effect.

For strong magnetic fields ($> 10^5$ gauss) we start directly from the LS multiplet since $H_Z > H_{\mathrm{SO}}$. The spin-orbit coupling H_{SO} is then included at the end as a perturbation. With the basis $|KLSM_LM_S\rangle$ in the LS multiplet we immediately obtain the Zeeman splitting

$$\langle H_Z\rangle = \frac{e\hbar}{2mc}(M_L + 2M_S)H_z. \tag{3.69}$$

The perturbation H_{SO} generates the correction

$$\langle H_{\mathrm{SO}}\rangle = AM_LM_S, \tag{3.70}$$

and the resulting splitting is

$$\langle H_Z + H_{\mathrm{SO}}\rangle \approx \mu_B(M_L + 2M_S)H_z \pm |A|M_LM_S, \tag{3.71}$$

where the upper (lower) sign applies for $\leq$ ($>$) half-filled shells. This is the Paschen-Back effect for the splitting of the levels in a strong magnetic field; the mixing of the

[15] Negative g factors are also possible (e.g. in the term scheme $^6F_{1/2}$). However, this does not correspond to the ground state of a neutral atom according to Hund's rules.

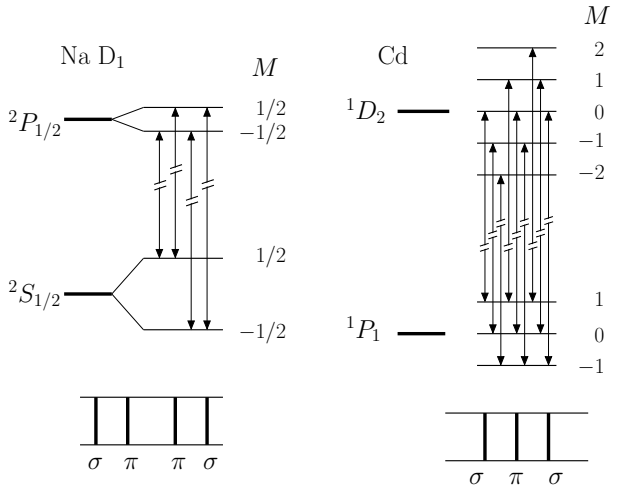


Figure 3.9: Anomalous and normal Zeeman effect for the Na D_1 transition [$^2S_{1/2}$ ($g = 2$) - $^2P_{1/2}$ ($g = 2/3$)] and for the Cd [$^1P_1$ ($g = 1$) - $^1D_2$ ($g = 1$)] transition. The π and σ transitions obey the selection rules $\Delta m = 0$ and $\Delta m = \pm 1$.

angular momenta $\boldsymbol{L}$ and $\boldsymbol{S}$ by the spin-orbit coupling H_{SO} is suppressed by the strong magnetic field and both angular momenta behave classically[16] (up to the $g = 2$ factor of the spin). Both the normal and the anomalous Zeeman effect go over into the Paschen-Back effect at high magnetic fields. Conversely, the nontrivial anomalous Zeeman effect at weak fields arises from the mixing of $\boldsymbol{S}$ and $\boldsymbol{L}$ into $\boldsymbol{J}$ by H_{SO}.

[16]The angular momenta $\boldsymbol{L}$ and $\boldsymbol{S}$ precess independently in the field.

Chapter 4

Second Quantization

The second quantization is a new formalism that provides us with an elegant and compact formulation of many-body problems. The second quantization has nothing to do with performing a "second quantization";[1] all quantum mechanics is already contained in the theory we know, which consists of wave functions, Schrödinger equations (for the dynamics), and symmetrization conditions for bosons and fermions. In second quantization, however, one switches from the product-wave-function basis to an occupation-number basis: it is no longer specified which particle is in which state, but only how many particles occupy a given single-particle state. In this way, one already incorporates much of the indistinguishability of the particles through the choice of basis. The remainder—especially the statistics—is then expressed via the commutation relations of the creation and annihilation operators.

As the name suggests, the creation and annihilation operators effect the creation and annihilation of particles. The theory therefore cannot be formulated in the Hilbert space with a fixed particle number; instead one must go over to the so-called Fock space, which is already familiar from the grand-canonical description of statistical physics. Formally, processes that do not conserve particle number thus become possible. The second-quantization formalism leads to a considerable simplification, but ultimately represents only a reformulation of the many-body problem and is by no means already a ready-made solution concept. An example of second quantization we already know: we found the spectrum of the harmonic oscillator in two alternative ways,

[1] The name is therefore somewhat misleading.

a) Conventional approach → First quantization.

$$H = p^2/2m + f\,q^2/2,$$
$$\big[p, q\big] = -i\hbar;$$
$$H\psi(q) = E\psi(q),$$
$$E_n = \hbar\omega(n + \tfrac{1}{2}),$$
$$\omega = \sqrt{f/m},$$
$$\Psi_n(q) = N_n\, H_n(q) e^{-m\omega q^2/2\hbar}.$$

b) Elegant approach → opera"-tor Instead **technique → Second quantization.**

$$H = \hbar\omega(a^\dagger a + 1/2)$$
$$\big[a, a^\dagger\big] = 1;$$
$$\text{Vacuum } |0\rangle,\ a|0\rangle = 0,$$
$$\text{states, } |n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle,$$
$$E_n = \hbar\omega(n + \tfrac{1}{2}),$$
$$\Psi_n(q) = \langle q|n\rangle.$$

of the wave function Ψ_n we have introduced the raising and lowering operators $a^\dagger$ and a and expressed the entire theory in terms of the latter. The "quanta" generated and annihilated by $a^\dagger$ and a then play the same role as "particles."[2]

We will now first introduce the occupation-number representation and then the formalism of second quantization for both bosons and fermions.

4.1 Problem formulation in first quantization

In this section we briefly summarize the results of Chapter 2. The Hamiltonian is given by

$$H = \sum_{i=1}^{N} h_i + \frac{1}{2}\sum_{i\neq j}^{N} v_{ij}\,. \tag{4.1}$$

Here h_i is a one-particle operator (QM) and v_{ij} describes the interaction of the particles i and j. Because the particles are indistinguishable, the Hamiltonian must be symmetric under exchange of the particles, see (2.18).

Let $\{|\varphi_\alpha\rangle\}$ be an orthonormal basis (ONS). Then the product states $|\varphi_{\alpha_1}, \ldots, \varphi_{\alpha_N}\rangle = |\varphi_{\alpha_1}\rangle \otimes \cdots \otimes |\varphi_{\alpha_N}\rangle$ span the Hilbert space $\mathcal{H}^{\otimes N}$ of distinguishable particles. The Hilbert space for indistinguishable particles $\mathcal{H}_{\mathcal{S},\mathcal{A}}$ is obtained by projection onto the appropriate subspace. We write the projectors compactly as (cf. (2.26))

$$\mathcal{P}_\pm = \frac{1}{N!}\sum_{\pi\in S_N} \zeta^\pi \pi \tag{4.2}$$

with $\pi|\varphi_{\alpha_1}, \ldots, \varphi_{\alpha_N}\rangle = |\varphi_{\alpha_{\pi(1)}}, \ldots, \varphi_{\alpha_{\pi(N)}}\rangle$, i.e. π acts by permuting the wave functions.[3] For bosons $\zeta = +1$, so that the states are symmetrized, and for fermions

[2] In general the particle number is conserved in nonrelativistic quantum mechanics. Here, however, the particle number n can be changed by $a^\dagger$ and a; hence the "particles" corresponding to excitations are not conserved in the case of the harmonic oscillator.

[3] This corresponds to exchanging the positions x with the inverse permutation π^{-1}.

$\zeta = -1$, which leads to antisymmetric states.

We then define basis states in $\mathcal{H}_{\mathcal{S},\mathcal{A}}$ by

$$|\varphi_{\alpha_1}, \dots, \varphi_{\alpha_N}\} = \sqrt{N!}\, \mathcal{P}_\pm |\varphi_{\alpha_1}, \dots, \varphi_{\alpha_N}\rangle = \frac{1}{\sqrt{N!}} \sum_{\pi \in S_N} \zeta^\pi |\varphi_{\alpha_{\pi(1)}}, \dots, \varphi_{\alpha_{\pi(N)}}\rangle \quad (4.3)$$

Because of the (anti-)symmetrization we have

$$|\dots, \varphi_{\alpha_k}, \dots, \varphi_{\alpha_l}, \dots\} = \zeta\, |\dots, \varphi_{\alpha_l}, \dots, \varphi_{\alpha_k}, \dots\} \quad (4.4)$$

for arbitrary k and l. These states are unnormalized for bosons. We will determine the normalization factor in the next section. For the proofs and derivations in this chapter it is, however, advantageous to work occasionally with the unnormalized (but symmetrized!) states $|\cdot\}$.

4.2 Besetzungszahldarstellung and Fock Space

Analogous to the example of the harmonic oscillator, we now introduce the occupation number basis. To do so, we first fix an ordering of the (orthonormal) one-particle states with $\varphi_1, \varphi_2, \dots$. [4] A state in $\mathcal{H}_{\mathcal{S},\mathcal{A}}$ is, by virtue of the symmetry, completely specified by stating how many particles occupy the states φ_j. As an example, consider the $N = 10$-particle state[5]

$$|\underbrace{\varphi_1, \varphi_1, \varphi_1, \varphi_1, \varphi_1, \varphi_1}_{6 \text{ times}}, \underbrace{\varphi_2, \varphi_2, \varphi_2, \varphi_2}_{4 \text{ times}}\}\,. \quad (4.5)$$

This state is determined by the specification n_j of how many particles occupy the states φ_j. In the example we have $n_1 = 6$ and $n_2 = 4$.

In general, we define the occupation number representation by

$$|n_1, n_2, \dots\rangle = \mathcal{N} |\overbrace{\varphi_1, \dots, \varphi_1}^{n_1 \text{ times}}, \overbrace{\varphi_2, \dots, \varphi_2}^{n_2 \text{ times}}, \dots\} \quad (4.6)$$

with $\sum_j n_j = N$. Note the ordering of the states on the right-hand side: the states φ_j are ordered by the index j. The Pauli principle for fermions requires $n_j = 0, 1$, while for bosons all occupation numbers $n_j = 0, 1, \dots$ are allowed. To determine the correct normalization factor $\mathcal{N}$, we compute the overlap between two states using the projector property $\mathcal{P}_\pm^2 = \mathcal{P}_\pm$

$$\begin{aligned} \{\varphi_{i_1}, \dots, \varphi_{i_N} | \varphi_{j_1}, \dots, \varphi_{j_N}\} &= N!\, \langle \varphi_{i_1}, \dots, \varphi_{i_N} | \mathcal{P}_\pm | \varphi_{j_1}, \dots, \varphi_{j_N} \rangle \\ &= \sum_{\pi \in S_N} \zeta^\pi \langle \varphi_{i_1} | \varphi_{j_{\pi(1)}} \rangle \dots \langle \varphi_{i_N} | \varphi_{j_{\pi(N)}} \rangle. \end{aligned} \quad (4.7)$$

[4] We require a discrete and ordered basis, i.e. the basis functions φ_j of the CONS can be labeled by natural numbers. A typical example are plane waves e^{ikx} with $k \in (2\pi/L)\mathbb{Z}$. In the end one takes the length L of the system to infinity, with the "usual" continuum limit $\sum_k \mapsto (L/2\pi) \int dk$.

[5] We use the (anti-)symmetry to order the particles according to the states φ_j.

By orthonormality of $|\varphi_j\rangle$, the only nonvanishing terms in the sum are those for which $(i_1, \ldots, i_N)$ is a permutation of $(j_1, \ldots, j_N)$. There are $\prod_j n_j!$ such permutations (one may permute the n_j states φ_j with fixed j arbitrarily), and hence

$$\mathcal{N} = \frac{1}{\sqrt{\prod_j n_j!}} \tag{4.8}$$

Note that for fermions $\mathcal{N} = 1$, independent of the state.

Thus every state in $\mathcal{H}_{\mathcal{S},\mathcal{A}}$ can be written as

$$|\Psi\rangle = \sum_{\substack{n_1,n_2,\ldots \\ \sum_i n_i = N}} c_{n_1,n_2,\ldots} |n_1, n_2, \ldots\rangle. \tag{4.9}$$

with

$$\langle n_1', n_2', \ldots | n_1, n_2, \ldots \rangle = \delta_{n_1',n_1} \delta_{n_2',n_2} \cdots , \tag{4.10}$$

To work, analogously to the harmonic oscillator, with creation and annihilation operators, we must relax the requirement of a fixed particle number N and introduce the Fock space

$$\mathcal{F} = \bigoplus_{N=0}^{\infty} \mathcal{H}^{\otimes N}; \tag{4.11}$$

The corresponding (anti-)symmetric subspace $\mathcal{F}_{\mathcal{S},\mathcal{A}}$ then provides the correct Hilbert space for describing bosons (fermions) in a grand canonical setting without fixed particle number. States with different particle numbers are, by construction, mutually orthogonal. The above direct sum contains the one-dimensional space $\mathcal{H}^{\otimes 0}$ (for $N = 0$), which consists only of the vacuum state $|0\rangle$ with no particles. The occupation number states span the Fock space $\mathcal{F}_{\mathcal{S},\mathcal{A}}$, with the completeness relation

$$\sum_{n_1,n_2,\ldots} |n_1, n_2, \ldots\rangle\langle n_1, n_2, \ldots| = I_{\mathcal{F}_{\mathcal{S},\mathcal{A}}}. \tag{4.12}$$

4.3 Creation and Annihilation Operators

All states in the Fock space can be generated from the vacuum state by introducing creation and annihilation operators. For each one-particle state φ_j we define the creation operator $a_j^\dagger \colon \mathcal{H}^{\otimes N} \to \mathcal{H}^{\otimes (N+1)}$ through its action

$$a_j^\dagger \, |\psi_1, \ldots, \psi_N\} = |\varphi_j, \psi_1, \ldots, \psi_N\}. \tag{4.13}$$

on an (anti-)symmetric product state $|\psi_1, \ldots, \psi_N\} = \sqrt{N!}\mathcal{P}_\pm |\psi_1, \ldots, \psi_N\rangle$. The operator adds a particle "in the first entry," while maintaining the (anti-)symmetry of the state.

The action of $a_j^\dagger$ in the occupation number representation is therefore given by

$$a_j^\dagger |n_1, \dots, n_j, \dots\rangle = \zeta^{S_j} \sqrt{n_j + 1}\, |n_1, \dots, n_j + 1, \dots\rangle, \qquad S_j = \sum_{i=1}^{j-1} n_i\,. \tag{4.14}$$

The prefactor $\sqrt{n_j+1} = \sqrt{(n_j+1)!/n_j!}$ follows from the normalization (4.8) of the occupation number states. The sign factor ζ^{S_j} arises because the added state φ_j stands at the beginning of $|\dots\}$ and must be brought to the correct position in the occupation number representation. To do so, it must be permuted past the n_1 wavefunctions φ_1, the n_2 wavefunctions φ_2, ..., and the n_{j-1} wavefunctions φ_{j-1}, cf. (4.4).

From equation (4.14) it follows that every state in the Fock space can be generated by repeated application of creation operators to the vacuum state $|0\rangle$ [6]

$$|n_1, n_2, \dots\rangle = \frac{(a_1^\dagger)^{n_1}}{\sqrt{n_1!}} \frac{(a_2^\dagger)^{n_2}}{\sqrt{n_2!}} \dots\ |0\rangle\,. \tag{4.15}$$

Equation (4.15) is a strong statement: one can generate the complicated (anti-)symmetric many-particle state (left) by successive application of the linear creation operators (right) to a single reference state (see exercise).

It may turn out that constraints (in particular that $n_j = 0, 1$ for fermions) lead to complicated relations between the various $a_j^\dagger$. We will now show that this is not the case. Consider the action of two creation operators $a_i^\dagger$ and $a_j^\dagger$ with $i \neq j$ on an arbitrary product state

$$\begin{aligned} a_i^\dagger a_j^\dagger |\psi_1, \psi_2, \dots\} &= |\varphi_i, \varphi_j, \psi_1, \psi_2, \dots\} = \zeta |\varphi_j, \varphi_i, \psi_1, \psi_2, \dots\} \\ &= \zeta a_j^\dagger a_i^\dagger |\psi_1, \psi_2, \dots\}\,. \end{aligned}$$

.

It follows that $(i \neq j)$

$$\begin{aligned} [a_i^\dagger, a_j^\dagger]_+ &= [a_i^\dagger, a_j^\dagger] = a_i^\dagger a_j^\dagger - a_j^\dagger a_i^\dagger = 0, \qquad \text{(bosons)}, \\ [a_i^\dagger, a_j^\dagger]_- &= \{a_i^\dagger, a_j^\dagger\} = a_i^\dagger a_j^\dagger + a_j^\dagger a_i^\dagger = 0, \qquad \text{(fermions)}. \end{aligned} \tag{4.16}$$

Thus only the question about $i = j$ remains. For bosons ($\zeta = 1$) one obviously has $[a_j^\dagger, a_j^\dagger] = 0$ (identical operators commute). For fermions the Pauli principle comes into play, which forbids occupying a state more than once. This is encoded in the operator identity $(a_j^\dagger)^2 = 0$ (nilpotency), from which $\{a_j^\dagger, a_j^\dagger\} = 0$ follows.

[6] The order of the operators is important for fermions! In the form given here, no sign factors arise because we first "fill" the last state with particles.

Altogether the creation operators thus satisfy the algebra

$$[a_i^\dagger, a_j^\dagger]_\zeta = 0. \tag{4.17}$$

Next we turn to the adjoint operator $a_j = (a_j^\dagger)^\dagger$. The general matrix elements of (4.14) have the form

$$\langle n_1, \dots, n_j, \dots | a_j^\dagger | n_1', \dots, n_j', \dots \rangle = \zeta^{S_j} \sqrt{n_j' + 1}\, \delta_{n_1, n_1'} \cdots \delta_{n_j, n_j'+1} \cdots . \tag{4.18}$$

Complex conjugation of the above equation leads to the relation

$$\langle n_1', \dots, n_j', \dots | a_j | n_1, \dots, n_j, \dots \rangle = \zeta^{S_j} \sqrt{n_j}\, \delta_{n_1', n_1} \cdots \delta_{n_j', n_j - 1} \cdots . \tag{4.19}$$

Thus for the adjoint operator a_i

$$a_j | n_1, \dots, n_j, \dots \rangle = \zeta^{S_j} \sqrt{n_j}\, | n_1, \dots, n_j - 1, \dots \rangle. \tag{4.20}$$

Since a_i lowers the particle number, it is called the annihilation operator. Adjoining (4.17) immediately yields $[a_i, a_j]_\zeta = 0$. From (4.20) one obtains the relation [7]

$$a_j |0\rangle = 0, \tag{4.21}$$

which (uniquely) determines the vacuum state.

To complete the algebra we still need the (anti-)commutation relations between a_i and $a_j^\dagger$. For this we consider the action of a_i and $a_j^\dagger$ on an arbitrary state in the occupation-number representation.

For $i < j$ one has

$$\begin{aligned} a_i a_j^\dagger | \dots, n_i, \dots, n_j, \dots \rangle =& a_i \zeta^{S_j} \sqrt{n_j + 1}\, | \dots, n_i, \dots, n_j + 1, \dots \rangle \\ =& \zeta^{S_j + S_i} \sqrt{(n_j + 1) n_i}\, | \dots, n_i - 1, \dots, n_j + 1, \dots \rangle \end{aligned}$$

and analogously

$$\begin{aligned} a_j^\dagger a_i | \dots, n_i, \dots, n_j, \dots \rangle =& a_j^\dagger \zeta^{S_i} \sqrt{n_i}\, | \dots, n_i - 1, \dots, n_j, \dots \rangle \\ =& \zeta^{(S_j - 1) + S_i} \sqrt{(n_j + 1) n_i}\, | \dots, n_i - 1, \dots, n_j + 1, \dots \rangle. \end{aligned}$$

Combining these two equations yields $(a_i a_j^\dagger - \zeta a_j^\dagger a_i) | n_1, n_2, \dots \rangle = 0$ and hence $[a_i, a_j^\dagger]_\zeta = 0$ for $i < j$. The proof of this relation for $i > j$ follows analogously.

Lastly we consider the case $i = j$: For bosons we have

$$a_j a_j^\dagger | \dots, n_j, \dots \rangle = (n_j + 1) | \dots, n_j, \dots \rangle$$

[7] It is important not to confuse the basis state $|0\rangle$ in the Hilbert space without particles (vacuum) with the number 0.

while

$$a_j^\dagger a_j |\dots, n_j, \dots\rangle = n_j |\dots, n_j, \dots\rangle.$$

From the difference of the two equations $(a_j a_j^\dagger - a_j^\dagger a_j)|n_1, n_2, \dots\rangle = |n_1, n_2, \dots\rangle$ follows and hence $[a_j, a_j^\dagger] = 1$. For fermions one must note that $n_j = 0, 1$.

Thus $a_j a_j^\dagger |\dots, n_j, \dots\rangle = (1 - n_j)|\dots, n_j, \dots\rangle$, $a_j^\dagger a_j |\dots, n_j, \dots\rangle = n_j |\dots, n_j, \dots\rangle$, from which $\{a_i^\dagger, a_i\} = 1$ follows.

In summary the creation and annihilation operators satisfy the canonical commutation relations

Bosons:	Fermions:
$[a_i, a_j^\dagger] = \delta_{ij},$	$\{a_i, a_j^\dagger\} = \delta_{ij},$
$[a_i, a_j] = 0,$	$\{a_i, a_j\} = 0,$
$[a_i^\dagger, a_j^\dagger] = 0,$	$\{a_i^\dagger, a_j^\dagger\} = 0,$

(4.22)

each of which forms a closed algebra. Noting that the full complexity of the Fock space is generated by repeatedly applying $a_j^\dagger$, the simplicity of the commutation relations is remarkable.

4.4 Basis change

The creation and annihilation operators directly refer to a basis. It is therefore important to understand how the formalism behaves under a basis change. Let the single-particle states $\{|\varphi_i\rangle\}$ and $\{|\psi_j\rangle\}$ be two orthonormal bases. A change between the two basis systems is described in QM by the unitary basis transformation

$$\begin{aligned} |\varphi_i\rangle &= \sum_j \langle\psi_j|\varphi_i\rangle\, |\psi_j\rangle, \\ |\psi_j\rangle &= \sum_i \langle\varphi_i|\psi_j\rangle\, |\varphi_i\rangle \end{aligned} \tag{4.23}$$

In the language of second quantization this basis change corresponds to a transformation of the annihilation operators a_i (to φ_i) into b_j (to ψ_j). From the relation $|\varphi_i\rangle = a_i^\dagger|0\rangle$ we obtain

$$a_i^\dagger|0\rangle = |\varphi_i\rangle \overset{(4.23)}{=} \sum_j \langle\psi_j|\varphi_i\rangle\, |\psi_j\rangle = \sum_j \langle\psi_j|\varphi_i\rangle\, b_j^\dagger\, |0\rangle. \tag{4.24}$$

A comparison of coefficients then leads to the transformation rules

$$\begin{aligned} a_i^\dagger &= \sum_j \langle\psi_j|\varphi_i\rangle b_j^\dagger, & b_j^\dagger &= \sum_i \langle\varphi_i|\psi_j\rangle a_i^\dagger, \\ a_i &= \sum_j \langle\varphi_i|\psi_j\rangle b_j, & b_j &= \sum_i \langle\psi_j|\varphi_i\rangle a_i. \end{aligned} \tag{4.25}$$

The commutation relations of the new operators $b_j, b_j^\dagger$ follow directly from those of a_i, $a_i^\dagger$. In particular one has

$$[b_i, b_j]_\zeta \overset{(4.25)}{=} \sum_{k,k'} \langle\psi_i|\varphi_k\rangle\langle\psi_j|\varphi_{k'}\rangle \underbrace{[a_k, a_{k'}]_\zeta}_{=0} = 0$$

Furthermore, with the completeness of $\{|\varphi_k\rangle\}$

$$\begin{aligned}[b_i, b_j^\dagger]_\zeta &\overset{(4.25)}{=} \sum_{k,k'} \langle\psi_i|\varphi_k\rangle\langle\varphi_{k'}|\psi_j\rangle \overbrace{[a_k, a_{k'}^\dagger]_\zeta}^{=\delta_{kk'}} \\ &= \sum_k \langle\psi_i|\varphi_k\rangle\langle\varphi_k|\psi_j\rangle = \langle\psi_i|\psi_j\rangle = \delta_{ij}\,. \end{aligned} \tag{4.26}$$

Thus $b_j, b_j^\dagger$ also satisfy the canonical commutation relations.

As an application of basis transformations we consider fermions on the lattice. The operators $c_i^\dagger$, c_i create or annihilate, respectively, a fermion at site i. On the other hand the operators $c_{\boldsymbol{k}}^\dagger$, and $c_{\boldsymbol{k}}$ create/annihilate plane waves with wave vector $\boldsymbol{k}$. Then

$$\begin{aligned} c_{\boldsymbol{k}}^\dagger &= \sum_i \langle i|\boldsymbol{k}\rangle\, c_i^\dagger = \sum_i \frac{e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}}{\sqrt{\Omega}}\, c_i^\dagger, \\ c_i^\dagger &= \sum_{\boldsymbol{k}} \frac{e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_i}}{\sqrt{\Omega}}\, c_{\boldsymbol{k}}^\dagger. \end{aligned} \tag{4.27}$$

Next we are interested in the two-particle state $|\boldsymbol{kq}\rangle$, in particular its representation in the position basis on the lattice, which is given by the matrix elements $\langle ij|\boldsymbol{kq}\rangle$. To this end we use the definitions

$$|ij\rangle = c_i^\dagger c_j^\dagger\,|0\rangle, \qquad |\boldsymbol{kq}\rangle = c_{\boldsymbol{k}}^\dagger c_{\boldsymbol{q}}^\dagger|0\rangle, \tag{4.28}$$

$$\begin{aligned} \Rightarrow \langle ij|\boldsymbol{kq}\rangle &= \langle 0|c_j c_i c_{\boldsymbol{k}}^\dagger c_{\boldsymbol{q}}^\dagger|0\rangle \\ &= \frac{1}{\Omega}\sum_{lm}\langle 0|c_j c_i e^{i\boldsymbol{k}\cdot\boldsymbol{r}_l} c_l^\dagger e^{i\boldsymbol{q}\cdot\boldsymbol{r}_m} c_m^\dagger|0\rangle \\ &= \frac{1}{\Omega}\sum_{lm} e^{i\boldsymbol{k}\cdot\boldsymbol{r}_l} e^{i\boldsymbol{q}\cdot\boldsymbol{r}_m}\langle 0|c_j c_i c_l^\dagger c_m^\dagger|0\rangle. \end{aligned} \tag{4.29}$$

The calculation of the matrix elements $\langle 0|c_j c_i c_l^\dagger c_m^\dagger|0\rangle$ shows how to proceed in the second-quantized formalism: one uses the commutation relations (4.22) and the

definition of the vacuum (4.21),

$$\begin{aligned}
\langle 0|c_j c_i c_l^\dagger c_m^\dagger|0\rangle &= \langle 0|c_j(\{c_i, c_l^\dagger\} - c_l^\dagger c_i)c_m^\dagger|0\rangle \\
&= \langle 0|c_j(\delta_{il} - c_l^\dagger c_i)c_m^\dagger|0\rangle \\
&= \delta_{il}\underbrace{\langle 0|c_j c_m^\dagger|0\rangle}_{\langle 0|\underbrace{\{c_j,c_m^\dagger\}}_{\delta_{jm}} - c_m^\dagger \underbrace{c_j|0\rangle}_{0}} - \langle 0|c_j c_l^\dagger \underbrace{c_i c_m^\dagger}_{\underbrace{\{c_i,c_m^\dagger\}}_{\delta_{im}} - c_m^\dagger c_i}|0\rangle \\
&= \delta_{il}\delta_{jm} - \langle 0|c_j c_l^\dagger(\delta_{im} - c_m^\dagger \underbrace{c_i}_{0})|0\rangle \\
&= \delta_{il}\delta_{jm} - \delta_{im}\langle 0|\underbrace{\{c_j, c_l^\dagger\}}_{\delta_{jl}} - c_l^\dagger \underbrace{c_j|0\rangle}_{0} \\
&= \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}.
\end{aligned} \tag{4.30}$$

We insert the result into (4.29) and obtain the two-particle matrix element,

$$\begin{aligned}
\langle ij|\boldsymbol{kq}\rangle &= \frac{1}{\Omega}\sum_{lm}(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})e^{i\boldsymbol{k}\cdot\boldsymbol{r}_l}e^{i\boldsymbol{q}\cdot\boldsymbol{r}_m} \\
&= \frac{1}{\Omega}\begin{vmatrix} e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i} & e^{i\boldsymbol{k}\cdot\boldsymbol{r}_j} \\ e^{i\boldsymbol{q}\cdot\boldsymbol{r}_i} & e^{i\boldsymbol{q}\cdot\boldsymbol{r}_j} \end{vmatrix}.
\end{aligned} \tag{4.31}$$

Hence $\langle ij|\boldsymbol{kq}\rangle$ is precisely the Slater determinant of the plane-wave product function $\varphi_{\boldsymbol{k}}(\boldsymbol{r}_i)\,\varphi_{\boldsymbol{q}}(\boldsymbol{r}_j)$. Correspondingly $\langle ij|\boldsymbol{kq}\rangle$ is already correctly antisymmetrized.

4.5 Operators in second quantization

In quantum mechanics one needs, besides the states, also operators. The operators too can be compactly represented in the formalism of second quantization. As an introduction we introduce an alternative representation of the annihilation operator a_j (to φ_j), which is analogous to (4.13).

For every single-particle state φ_j we define the creation operator $a_j\colon \mathcal{H}^{\otimes N} \to \mathcal{H}^{\otimes(N-1)}$ by its action

$$a_j|\psi_1, \ldots, \psi_N\} = \sum_{i=1}^{N} \zeta^{i-1}\langle\varphi_j|\psi_i\rangle\ |\psi_1, \ldots, \psi_{i-1}, \psi_{i+1}, \ldots, \psi_N\}. \tag{4.32}$$

on an arbitrary (anti-)symmetric product state. Here $|\psi_1, \ldots, \psi_{i-1}, \psi_{i+1}, \ldots, \psi_N\}$ denotes the state in $\mathcal{H}^{\otimes(N-1)}$ obtained by omitting the i-th state.

We now show that the so-defined a_j agrees with (4.20). By linearity it suffices to verify this on the states $|n_1, \ldots, n_j, \ldots\rangle$. One finds (with $\langle\varphi_i|\varphi_j\rangle = \delta_{ij}$ and the definition

(4.32))

$$a_j|\ldots,\overbrace{\varphi_j,\ldots,\varphi_j}^{n_j\text{ times}},\ldots\} = \zeta^{S_j}\, n_j|\ldots,\overbrace{\varphi_j,\ldots,\varphi_j}^{(n_j-1)\text{ times}},\ldots\}\,,$$

where we have used that n_j terms give the same contribution.[8] Hence from (4.32) we obtain the action on the occupation-number states

$$\begin{aligned} a_j|n_1,\ldots,n_j,\ldots\rangle &= \frac{1}{\sqrt{\prod_i n_i!}}\, a_j|\ldots,\overbrace{\varphi_j,\ldots,\varphi_j}^{n_j\text{ times}},\ldots\} \\ &= \zeta^{S_j}\frac{1}{\sqrt{\prod_i n_i!}}\, n_j|\ldots,\overbrace{\varphi_j,\ldots,\varphi_j}^{(n_j-1)\text{ times}},\ldots\} \\ &= \zeta^{S_j}\sqrt{n_j}\,|n_1,\ldots,n_j-1,\ldots\rangle, \end{aligned} \tag{4.33}$$

which agrees with (4.20).

4.5.1 Single-particle operators

Let $b = \sum_{\alpha,\beta} b_{\alpha\beta}|\varphi_\alpha\rangle\langle\varphi_\beta|$ be a general operator on $\mathcal{H}$ with matrix elements $b_{\alpha\beta} = \langle\varphi_\alpha|b|\varphi_\beta\rangle$.[9] Then the associated single-particle operator B on the Fock space $\mathcal{F}$ is defined as

$$B = \bigoplus_{N=0}^{\infty} B_N = 0 \oplus B_1 \oplus B_2 \oplus \ldots \tag{4.34}$$

with

$$B_1 = b, \qquad B_2 = b_1 + b_2 = b\otimes I + I\otimes b, \qquad B_N = \sum_{i=1}^{N} b_i \tag{4.35}$$

cf. (2.4). This operator is expressed in second quantization by

$$B = \sum_{\alpha,\beta} b_{\alpha\beta}\, a_\alpha^\dagger a_\beta = \sum_{\alpha,\beta}\langle\varphi_\alpha|b|\varphi_\beta\rangle\, a_\alpha^\dagger a_\beta \tag{4.36}$$

.

To prove this we show the representation $B = a_\alpha^\dagger a_\beta$ for the operator $b = |\varphi_\alpha\rangle\langle\varphi_\beta|$. By the linearity of (4.36) in b this yields the general relation. The single-particle operator B acts on a general product state $|\psi_1,\ldots,\psi_N\}$ in the N-particle subspace

[8] For bosons the factor ζ^k vanishes. For fermions $n_j \in \{0,1\}$ holds.

[9] In this section we use Greek indices α,β for sums over basis states to distinguish them from i, the sum over particles.

by[10]

$$B|\psi_1,\dots,\psi_N\} = \sqrt{N!}\mathcal{P}_\pm B_N|\psi_1,\dots,\psi_N\rangle = \sqrt{N!}\mathcal{P}_\pm \sum_{i=1}^{N}\Big(|\varphi_\alpha\rangle\langle\varphi_\beta|\Big)_i|\psi_1,\dots,\psi_N\rangle$$
$$= \sum_{i=1}^{N}\langle\varphi_\beta|\psi_i\rangle\,|\psi_1,\dots,\psi_{i-1},\varphi_\alpha,\psi_{i+1},\dots,\psi_N\}\,. \tag{4.37}$$

This result is equivalent to the second-quantized expression

$$\begin{aligned} a_\alpha^\dagger a_\beta|\psi_1,\dots,\psi_N\} &\overset{(4.32)}{=} \sum_{i=1}^{N}\zeta^{i-1}\langle\varphi_\beta|\psi_i\rangle\, a_\alpha^\dagger|\psi_1,\dots,\psi_{i-1},\psi_{i+1},\dots,\psi_N\} \\ &\overset{(4.13)}{=} \sum_{i=1}^{N}\zeta^{i-1}\langle\varphi_\beta|\psi_i\rangle\,|\varphi_\alpha,\psi_1,\dots,\psi_{i-1},\psi_{i+1},\dots,\psi_N\} \\ &\overset{(4.4)}{=} \sum_{i=1}^{N}\langle\varphi_\beta|\psi_i\rangle\,|\psi_1,\dots,\psi_{i-1},\varphi_\alpha,\psi_{i+1},\dots,\psi_N\} \end{aligned} \tag{4.38}$$

for arbitrary N, thereby establishing the relation (4.36) on the entire Fock space.

An important single-particle operator is the particle number. It corresponds to the operator $b = I$ with $N = B = \sum_\alpha a_\alpha^\dagger a_\alpha$. The operator N acts in the Fock space by simply returning N, i.e.

$$N = 0 \oplus 1 \oplus 2 \oplus 3 \oplus \cdots\,. \tag{4.39}$$

If a Hamiltonian is a single-particle operator, i.e. $H = \sum_i h_i$, it is called quadratic. All single-particle problems treated in nonrelativistic quantum mechanics fall into this category. In many-body theory one calls quadratic Hamiltonians solvable, because once the (complete) solution of the single-particle problem

$$h|\psi_\alpha\rangle = \varepsilon_\alpha|\psi_\alpha\rangle \tag{4.40}$$

is known, all many-body eigenstates can be constructed immediately. Defining the creation and annihilation operators b_α with respect to the energy eigenbasis (see Section 4.4), one obtains

$$H = \sum_\alpha \varepsilon_\alpha b_\alpha^\dagger b_\alpha\,. \tag{4.41}$$

The many-body eigenstates with $H|\boldsymbol{n}\rangle = E_{\boldsymbol{n}}|\boldsymbol{n}\rangle$ are simply occupation-number states $|\boldsymbol{n}\rangle = |n_1, n_2, \dots\rangle$ built from the single-particle eigenstates $|\varphi_\alpha\rangle$. The energy of the many-body states is given by

$$E_{\boldsymbol{n}} = \sum_\alpha n_\alpha\varepsilon_\alpha. \tag{4.42}$$

[10] Note that the operator B_N commutes with $\mathcal{P}_\pm$ because it acts identically on every state.

The statistics of the species enters only through the occupation numbers $n_\alpha = 0, 1, \ldots$ for bosons and $n_\alpha = 0, 1$ for fermions.

Many-body Hamiltonians, however, are usually given by a sum of a single-particle and a two-particle part. Then solving the single-particle problem is not sufficient to solve the full many-body problem. Rather, one typically uses the single-particle solution as a basis to treat the many-body problem perturbatively with the two-particle operator as the perturbation. An exact solution of the many-body problem is then usually not possible, and the computational effort, e.g. on a computer, grows with the particle number N even exponentially.

4.5.2 Two-particle operators

The matrix elements of a general two-particle operator $B = \frac{1}{2}\sum_{i\neq j} b_{ij}$ are given by

$$b_{\alpha_1\alpha_2,\beta_1\beta_2} = \langle \varphi_{\alpha_1}\varphi_{\alpha_2}|b|\varphi_{\beta_1}\varphi_{\beta_2}\rangle\,. \tag{4.43}$$

Its Fock representation reads

$$B = \frac{1}{2}\sum_{\substack{\alpha_1,\alpha_2\\ \beta_1,\beta_2}} \langle \varphi_{\alpha_1}\varphi_{\alpha_2}|b|\varphi_{\beta_1}\varphi_{\beta_2}\rangle\, a^\dagger_{\alpha_1} a^\dagger_{\alpha_2} a_{\beta_2} a_{\beta_1}\,. \tag{4.44}$$

We prove the relation for a product operator $b = d \otimes d$ of a single-particle operator d. The general case again follows by linearity in b. We note that $B = \frac{1}{2}\sum_{i\neq j}^{N}(d\otimes d)_{ij} = \frac{1}{2}\sum_{i,j}(d\otimes d)_{ij} - \frac{1}{2}\sum_i (d^2)_i$. Moreover,

$$\sum_{i,j}(d\otimes d)_{ij} = \left(\sum_i d_i\right)\otimes\left(\sum_j d_j\right) = \left(\sum_i d_i\right)^2. \tag{4.45}$$

From (4.36) we already know the representation of single-particle operators

$$\sum_i d_i = \sum_{\alpha,\beta} d_{\alpha\beta} a^\dagger_\alpha a_\beta, \qquad \text{and} \qquad \sum_i (d^2)_i = \sum_{\alpha,\beta}(d^2)_{\alpha\beta} a^\dagger_\alpha a_\beta \tag{4.46}$$

. Altogether one has

$$\begin{aligned} B &= \frac{1}{2}\left(\sum_{\alpha,\beta} d_{\alpha\beta} a^\dagger_\alpha a_\beta\right) - \frac{1}{2}\sum_{\alpha,\beta}(d^2)_{\alpha\beta} a^\dagger_\alpha a_\beta \\ &= \frac{1}{2}\sum_{\substack{\alpha_1,\alpha_2\\ \beta_1,\beta_2}} d_{\alpha_1\beta_1} d_{\alpha_2\beta_2} a^\dagger_{\alpha_1} a_{\beta_1} a^\dagger_{\alpha_2} a_{\beta_2} - \frac{1}{2}\sum_{\alpha,\beta}(d^2)_{\alpha\beta} a^\dagger_\alpha a_\beta \\ &= \frac{1}{2}\sum_{\substack{\alpha_1,\alpha_2\\ \beta_1,\beta_2}} d_{\alpha_1\beta_1} d_{\alpha_2\beta_2}\, a^\dagger_{\alpha_1} a^\dagger_{\alpha_2} a_{\beta_2} a_{\beta_1}\,, \end{aligned} \tag{4.47}$$

where we have used

$$a^\dagger_{\alpha_1}\overbrace{a_{\beta_1}a^\dagger_{\alpha_2}}^{=\zeta a^\dagger_{\alpha_2}a_{\beta_1}+\delta_{\alpha_2,\beta_1}}a_{\beta_2} = \zeta a^\dagger_{\alpha_1}a^\dagger_{\alpha_2}a_{\beta_1}a_{\beta_2}+\delta_{\alpha_2,\beta_1}a^\dagger_{\alpha_1}a_{\beta_2} = a^\dagger_{\alpha_1}a^\dagger_{\alpha_2}a_{\beta_2}a_{\beta_1}+\delta_{\alpha_2,\beta_1}a^\dagger_{\alpha_1}a_{\beta_2} \quad (4.48)$$

. The result follows with $b_{\alpha_1\alpha_2,\beta_1\beta_2} = d_{\alpha_1\beta_1}d_{\alpha_2\beta_2}$. Here one should note in particular the reversed order of the creation and annihilation operators. The particles are first "removed" and then reinserted in reverse order.

The derivation presented here extends straightforwardly to the case of a ν-particle operator. In general this has the form

$$B = \frac{1}{\nu!}\sum_{\substack{\alpha_1,\ldots,\alpha_\nu\\ \beta_1,\ldots,\beta_\nu}} \langle\varphi_{\alpha_1},\ldots,\varphi_{\alpha_\nu}|b|\varphi_{\beta_1},\ldots,\varphi_{\beta_\nu}\rangle\, a^\dagger_{\alpha_1}\ldots a^\dagger_{\alpha_\nu}a_{\beta_\nu}\ldots a_{\beta_1}. \quad (4.49)$$

4.6 Field Operators

By performing a basis transformation to the position basis $\{|\boldsymbol{r}s\rangle\}$, we obtain the field operators $\Psi_s(\boldsymbol{r})$ (where s denotes the spin quantum number). The field operators are convenient because they are independent of the system and thus allow for a unified description. Let $\{\varphi_\alpha\}$ be an arbitrary orthonormal basis in the single-particle Hilbert space. Completeness yields the expansion

$$|\boldsymbol{r}s\rangle = \sum_\alpha |\varphi_\alpha\rangle \underbrace{\langle\varphi_\alpha|\boldsymbol{r}s\rangle}_{\varphi^*_{\alpha,s}(\boldsymbol{r})}. \quad (4.50)$$

It is $|\varphi_\alpha\rangle = a^\dagger_\alpha|0\rangle$ and we analogously define $|\boldsymbol{r}s\rangle = \Psi^\dagger_s(\boldsymbol{r})|0\rangle$, i.e., the field operator $\Psi^\dagger_s(\boldsymbol{r})$ creates a particle with spin s at position $\boldsymbol{r}$. The relation (4.50) can then be written as

$$\Psi^\dagger_s(\boldsymbol{r})|0\rangle = \sum_\alpha \varphi^*_{\alpha,s}(\boldsymbol{r})\, a^\dagger_\alpha|0\rangle \quad (4.51)$$

and we obtain the transformation to the field operators (cf. (4.25))

$$\Psi^\dagger_s(\boldsymbol{r}) = \sum_\alpha \varphi^*_{\alpha,s}(\boldsymbol{r})\, a^\dagger_\alpha, \qquad \Psi_s(\boldsymbol{r}) = \sum_\alpha \varphi_{\alpha,s}(\boldsymbol{r})\, a_\alpha. \quad (4.52)$$

The notation for the so-called field operators $\Psi^\dagger_s(\boldsymbol{r})$, $\Psi_s(\boldsymbol{r})$ is historically motivated and is meant to recall the wave function $\psi_s(\boldsymbol{r})$.

The commutation relations of the field operators follow from completeness and from the canonical relations of the a_α and $a^\dagger_\alpha$ operators. The only nontrivial case is

$$\begin{aligned}[\Psi_s(\boldsymbol{r}),\Psi^\dagger_{s'}(\boldsymbol{r}')]_\zeta &= \sum_{\alpha,\alpha'}\varphi_{\alpha,s}(\boldsymbol{r})\varphi^*_{\alpha',s'}(\boldsymbol{r}')\overbrace{[a_\alpha,a^\dagger_{\alpha'}]_\zeta}^{\delta_{\alpha,\alpha'}} = \sum_\alpha \langle\boldsymbol{r},s|\varphi_\alpha\rangle\langle\varphi_\alpha|\boldsymbol{r}',s'\rangle\\ &= \delta^{(3)}(\boldsymbol{r}-\boldsymbol{r}')\delta_{ss'} \end{aligned} \quad (4.53)$$

For bosons this yields

$$
\begin{aligned}
&[\Psi_s(\boldsymbol{r}), \Psi_{s'}(\boldsymbol{r}')] = 0, \\
&[\Psi_s^\dagger(\boldsymbol{r}), \Psi_{s'}^\dagger(\boldsymbol{r}')] = 0, \\
&[\Psi_s(\boldsymbol{r}), \Psi_{s'}^\dagger(\boldsymbol{r}')] = \delta^{(3)}(\boldsymbol{r}-\boldsymbol{r}')\delta_{ss'}.
\end{aligned}
$$

For fermions this yields

$$
\begin{aligned}
&\{\Psi_s(\boldsymbol{r}), \Psi_{s'}(\boldsymbol{r}')\} = 0, \\
&\{\Psi_s^\dagger(\boldsymbol{r}), \Psi_{s'}^\dagger(\boldsymbol{r}')\} = 0, \\
&\{\Psi_s(\boldsymbol{r}), \Psi_{s'}^\dagger(\boldsymbol{r}')\} = \delta^{(3)}(\boldsymbol{r}-\boldsymbol{r}')\delta_{ss'}.
\end{aligned}
\tag{4.54}
$$

4.6.1 The Many-Particle Position Basis

In the following, we suppress the spin index s to keep the notation simple. We define the states of the many-particle position basis by

$$
|\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N\} = \Psi^\dagger(\boldsymbol{r}_1)\Psi^\dagger(\boldsymbol{r}_2)\ldots\Psi^\dagger(\boldsymbol{r}_N)|0\rangle \tag{4.55}
$$

By construction, the states are (anti-)symmetric under exchange of two particles. $\Psi^\dagger(\boldsymbol{r})$ adds a particle at $\boldsymbol{r}$ while $\Psi(\boldsymbol{r})$ annihilates a particle at $\boldsymbol{r}$. Acting with the field operators on an (anti-)symmetric position state thus yields

$$
\Psi^\dagger(\boldsymbol{r})|\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N\} = |\boldsymbol{r}, \boldsymbol{r}_1, \ldots, \boldsymbol{r}_N\} \tag{4.56}
$$

$$
\begin{aligned}
\Psi(\boldsymbol{r})\,|\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N\} &= \Psi(\boldsymbol{r})\Psi^\dagger(\boldsymbol{r}_1)\ldots\Psi^\dagger(\boldsymbol{r}_N)|0\rangle \\
&= \sum_{i=1}^{N} \zeta^{i-1}\delta(\boldsymbol{r}-\boldsymbol{r}_i)|\boldsymbol{r}_1, \ldots, \boldsymbol{r}_{i-1}, \boldsymbol{r}_{i+1}, \ldots, \boldsymbol{r}_N\}.
\end{aligned}
\tag{4.57}
$$

The equivalence in the last line follows by permuting $\Psi(\boldsymbol{r})$ through all $\Psi^\dagger(\boldsymbol{r}_i)$. The physical interpretation is that a particle at $\boldsymbol{r}$ can only be removed if one of the N coordinates $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N$ equals $\boldsymbol{r}$.

The overlap of two states in the many-particle position basis is determined analogously to equation (4.7) as

$$
\begin{aligned}
\{\boldsymbol{r}'_1, \ldots, \boldsymbol{r}'_{N'}|\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N\} = \delta_{NN'} \sum_{\pi\in S_n} \zeta^\pi \delta(\boldsymbol{r}_1 - \boldsymbol{r}'_{\pi(1)}) \\
\times \delta(\boldsymbol{r}_2 - \boldsymbol{r}'_{\pi(2)})\cdots\delta(\boldsymbol{r}_N - \boldsymbol{r}'_{\pi(N)}).
\end{aligned}
\tag{4.58}
$$

The identity on the Fock space can also be written in the many-particle position basis. To this end, one starts with the identity on the Hilbert space of distinguishable particles and applies the (anti-)symmetrization as an operator transformation

$$
\begin{aligned}
I_{\mathcal{F}} &= \sum_N \int d^3r_1 \cdots d^3r_N \, \mathcal{P}_\pm |\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N\rangle\langle \boldsymbol{r}_1, \ldots, \boldsymbol{r}_N| \mathcal{P}_\pm \\
&= \sum_N \frac{1}{N!} \int d^3r_1 \cdots d^3r_N |\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N\}\{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N|
\end{aligned}
\tag{4.59}
$$

With this representation of the identity, arbitrary states in the many-particle position basis can be expanded as usual:

$$
|\Phi\rangle = \sum_N \frac{1}{N!} \int d^3r_1 \cdots d^3r_N \, \varphi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)|\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N\} \tag{4.60}
$$

$|\Phi\rangle$ has the correct symmetry for arbitrary wave functions $\varphi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N)$, even those that are not (anti-)symmetrized. This is seen most easily by considering equation (4.60) for φ with fixed N

$$\begin{aligned}\{\boldsymbol{r}'_1,\ldots,\boldsymbol{r}'_N|\Phi\rangle &\overset{(4.58)}{=} \frac{1}{N!}\int d^3r_1\cdots d^3r_N\,\varphi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N)\{\boldsymbol{r}'_1,\ldots,\boldsymbol{r}'_N|\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N\}\\ &= \frac{1}{N!}\sum_\pi \zeta^\pi\varphi(\boldsymbol{r}'_{\pi(1)},\ldots,\boldsymbol{r}'_{\pi(N)}) = \varphi(\boldsymbol{r}'_1,\ldots,\boldsymbol{r}'_N),\end{aligned}$$

Here, the last equation holds if φ is already (anti-)symmetrized. If φ is also normalized, i.e., $\int d^3r_1\cdots d^3r_N\,|\varphi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N)|^2 = 1$, then $|\Phi\rangle$ is also normalized, $\langle\Phi|\Phi\rangle = 1$.

4.6.2 Operators Expressed via Field Operators

As already mentioned, the field operators are particularly well suited for deriving the representation of various operators in second quantization. In this section, we demonstrate this by means of a number of important examples. We begin with the single-particle operators.

4.6.3 Single-Particle Operators

The representation of a single-particle operator B in terms of field operators is obtained from equation (4.25) by expressing the operators a_α with (4.52) in terms of the field operators. Starting from a single-particle operator $b = \iint d^3r'd^3r\,B(\boldsymbol{r}',\boldsymbol{r})|\boldsymbol{r}'\rangle\langle\boldsymbol{r}|$ in the position basis, we obtain the general expression for the single-particle operator

$$\begin{aligned}B &= \sum_{\alpha,\beta}\langle\varphi_\alpha|b|\varphi_\beta\rangle\,a^\dagger_\alpha a_\beta = \sum_{\alpha,\beta}\iint d^3r'd^3r\,\varphi^*_\alpha(\boldsymbol{r}')B(\boldsymbol{r}',\boldsymbol{r})\varphi_\beta(\boldsymbol{r})a^\dagger_\alpha a_\beta\\ &= \iint d^3r'd^3r\,\Psi^\dagger(\boldsymbol{r}')B(\boldsymbol{r}',\boldsymbol{r})\Psi(\boldsymbol{r}).\end{aligned}\tag{4.61}$$

It is instructive to compare this second-quantized form with the first-quantized form of the expectation value in the state $\varphi(\boldsymbol{r})$,

$$\langle b\rangle_\varphi = \iint d^3r'd^3r\,\varphi^*(\boldsymbol{r}')B(\boldsymbol{r}',\boldsymbol{r})\varphi(\boldsymbol{r}).\tag{4.62}$$

If one replaces the wave function $\varphi(\boldsymbol{r})$ by the field operator $\Psi(\boldsymbol{r})$ in the expression for $\langle b\rangle_\varphi$, one obtains the operator in second-quantized form! This formal similarity is the reason why the formalism with creation and annihilation operators is called "second quantization"; single-particle wave functions appear to have become operators that create and annihilate particles. However, this is only a formal connection. We have derived the second-quantized formalism from first-quantized many-particle quantum mechanics without additional assumptions. The formal connection is therefore rather an expression of the "naturalness" of second quantization.

Example Density Operator: The expectation value of the single-particle density operator

$$\begin{aligned}\langle\rho(\boldsymbol{r})\rangle_\varphi &= \int d^3r'\,\varphi^*(\boldsymbol{r}')\delta^{(3)}(\boldsymbol{r}-\boldsymbol{r}')\varphi(\boldsymbol{r}') \\ &= \varphi^*(\boldsymbol{r})\varphi(\boldsymbol{r}),\end{aligned} \tag{4.63}$$

becomes, upon rewriting in second-quantized form, the density operator of the many-particle system

$$\rho(\boldsymbol{r}) = \Psi^\dagger(\boldsymbol{r})\Psi(\boldsymbol{r}). \tag{4.64}$$

The particle-number operator, which in first-quantized form is obtained by integrating over the density $\rho(\boldsymbol{r})$, has in second quantization the form

$$N = \int d^3r\,\Psi^\dagger(\boldsymbol{r})\Psi(\boldsymbol{r}). \tag{4.65}$$

For translationally invariant systems it is often convenient to work with quantities in momentum space. To this end, we define the Fourier transform of the density as $\rho(\boldsymbol{q}) = \int d^3r\, e^{-i\boldsymbol{q}\cdot\boldsymbol{r}}\rho(\boldsymbol{r})$ and use the representation $\Psi(\boldsymbol{r}) = \Omega^{-1/2}\sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}}c_{\boldsymbol{k}}$. This yields

$$\begin{aligned}\rho(\boldsymbol{q}) &= \int d^3r\, e^{-i\boldsymbol{q}\cdot\boldsymbol{r}}\rho(\boldsymbol{r}) = \sum_{\boldsymbol{k}\boldsymbol{k}'}\frac{1}{\Omega}\int d^3r\, e^{-i\boldsymbol{q}\cdot\boldsymbol{r}}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}e^{i\boldsymbol{k}'\cdot\boldsymbol{r}}c^\dagger_{\boldsymbol{k}}\,c_{\boldsymbol{k}'} \\ &= \sum_{\boldsymbol{k}} c^\dagger_{\boldsymbol{k}}c_{\boldsymbol{k}+\boldsymbol{q}}.\end{aligned} \tag{4.66}$$

Example Current Density Operator: The expression for the current density in first quantization

$$\langle\boldsymbol{j}(\boldsymbol{r})\rangle = \frac{\hbar}{2mi}\left[\varphi^*(\boldsymbol{r})\boldsymbol{\nabla}\varphi(\boldsymbol{r}) - \varphi(\boldsymbol{r})\boldsymbol{\nabla}\varphi^*(\boldsymbol{r})\right] \tag{4.67}$$

becomes in the Fock space the current density operator of the many-particle system

$$\boldsymbol{j}(\boldsymbol{r}) = \frac{\hbar}{2mi}\left[\Psi^\dagger(\boldsymbol{r})\boldsymbol{\nabla}\Psi(\boldsymbol{r}) - \big(\boldsymbol{\nabla}\Psi^\dagger(\boldsymbol{r})\big)\Psi(\boldsymbol{r})\right]. \tag{4.68}$$

Note the order of the operators: First, $\Psi(\boldsymbol{r})$ must "remove" a particle from the state $|\Phi\rangle$, then the operator acts at that position, and finally the particle is reinserted by $\Psi^\dagger(\boldsymbol{r})$.

The Fourier representation of the current density operator is determined in complete analogy to that of the density operator, with the result

$$\boldsymbol{j}(\boldsymbol{q}) = \frac{\hbar}{2m}\sum_{\boldsymbol{k}}(2\boldsymbol{k}+\boldsymbol{q})\,c^\dagger_{\boldsymbol{k}}c_{\boldsymbol{k}+\boldsymbol{q}}. \tag{4.69}$$

Example Kinetic Energy: The expectation value of the kinetic energy in first quantization

$$\langle T\rangle_\varphi = \int d^3r\, \varphi^*(\boldsymbol{r})\frac{-\hbar^2\nabla^2}{2m}\varphi(\boldsymbol{r}) = \frac{\hbar^2}{2m}\int d^3r\, \boldsymbol{\nabla}\varphi^*(\boldsymbol{r})\cdot\boldsymbol{\nabla}\varphi(\boldsymbol{r}),$$

leads in second quantization to the operator of the kinetic energy

$$T = \frac{\hbar^2}{2m}\int d^3r\, \boldsymbol{\nabla}\Psi^\dagger(\boldsymbol{r})\cdot\boldsymbol{\nabla}\Psi(\boldsymbol{r}). \tag{4.70}$$

In momentum space one obtains the result

$$T = \sum_{\boldsymbol{k}} \frac{\hbar^2k^2}{2m} c^\dagger_{\boldsymbol{k}} c_{\boldsymbol{k}} \tag{4.71}$$

Example Spin Density: The spin density of a spin-1/2 particle in first quantization

$$\langle \boldsymbol{s}(\boldsymbol{r})\rangle_\varphi = \frac{\hbar}{2}\sum_{s,s'} \varphi^*_s(\boldsymbol{r})\boldsymbol{\sigma}_{ss'}\varphi_{s'}(\boldsymbol{r}) \tag{4.72}$$

defines the many-particle spin density operator

$$\boldsymbol{s}(\boldsymbol{r}) = \frac{\hbar}{2}\sum_{ss'} \Psi^\dagger_s(\boldsymbol{r})\,\boldsymbol{\sigma}_{ss'}\,\Psi_{s'}(\boldsymbol{r}). \tag{4.73}$$

Two-Particle Operators

The interaction $V(\boldsymbol{r}_1,\boldsymbol{r}_2)$ is a two-particle operator. In second quantization it has the form

$$\begin{aligned} H_{\text{ww}} &= \frac{1}{2}\sum_{\alpha,\beta}\langle\varphi_{\alpha_1}\varphi_{\alpha_2}|V(\boldsymbol{r}_1,\boldsymbol{r}_2)|\varphi_{\beta_1}\varphi_{\beta_2}\rangle a^\dagger_{\alpha_1}a^\dagger_{\alpha_2}a_{\beta_2}a_{\beta_1} \\ &= \frac{1}{2}\int d^3r\, d^3r'\, \Psi^\dagger(\boldsymbol{r})\Psi^\dagger(\boldsymbol{r}')V(\boldsymbol{r},\boldsymbol{r}')\Psi(\boldsymbol{r}')\Psi(\boldsymbol{r}). \end{aligned} \tag{4.74}$$

Note the correct ordering $\Psi(\boldsymbol{r})^\dagger\Psi(\boldsymbol{r}')^\dagger\Psi(\boldsymbol{r}')\Psi(\boldsymbol{r})$ of the operators. Also here, the connection between the expectation value

$$\langle\varphi_\alpha\varphi_\beta|V(\boldsymbol{r}_1,\boldsymbol{r}_2)|\varphi_\alpha\varphi_\beta\rangle = \int d^3r d^3r'\, V(\boldsymbol{r},\boldsymbol{r}')|\varphi_\alpha(\boldsymbol{r})|^2\,|\varphi_\beta(\boldsymbol{r}')|^2 \tag{4.75}$$

and the operator under the replacement $\varphi \mapsto \Psi$ holds, modulo "correct" ordering of the operators.

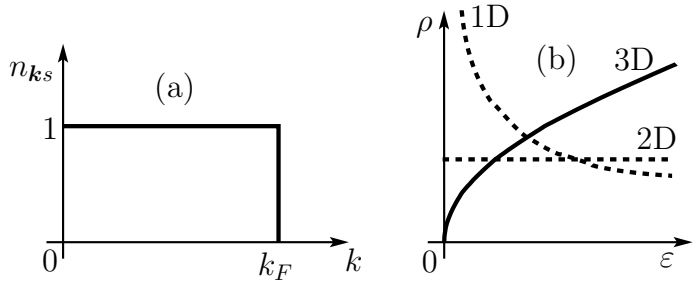


Figure 4.1: (a) Occupation numbers $n_{\boldsymbol{k}s}$ of the electron states with wave vector $\boldsymbol{k}$ and spin s for the Fermi sea. (b) Densities of states $\rho(\varepsilon)$ for free fermions in 3D (solid line) and in 2D and 1D.

Example Hamiltonian: The typical Hamiltonian of interacting particles with spin has the form (with spin-independent interaction)

$$H = \frac{\hbar^2}{2m} \int d^3r \sum_s \boldsymbol{\nabla} \Psi_s^\dagger(\boldsymbol{r}) \cdot \boldsymbol{\nabla} \Psi_s(\boldsymbol{r}) \tag{4.76}$$
$$+ \frac{1}{2} \int d^3r\, d^3r' \sum_{s,s'} \Psi_s^\dagger(\boldsymbol{r}) \Psi_{s'}^\dagger(\boldsymbol{r}') V(\boldsymbol{r}-\boldsymbol{r}') \Psi_{s'}(\boldsymbol{r}') \Psi_s(\boldsymbol{r}).$$

For an electron gas, $s = \pm 1/2$, $m = m_e$, $V(r) = e^2/r$ and $\Psi_s^\dagger(\boldsymbol{r})$, $\Psi_s(\boldsymbol{r})$ are fermionic field operators. In momentum space, the Hamiltonian has the form

$$H = \sum_{\boldsymbol{k}s} \frac{\hbar^2 k^2}{2m} c_{\boldsymbol{k}s}^\dagger c_{\boldsymbol{k}s} + \frac{1}{2} \sum_{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{q}} \sum_{s's} V_{\boldsymbol{q}}\, c_{\boldsymbol{k}+\boldsymbol{q}s}^\dagger c_{\boldsymbol{k}'-\boldsymbol{q}s'}^\dagger c_{\boldsymbol{k}'s'} c_{\boldsymbol{k}s}. \tag{4.77}$$

where $V_{\boldsymbol{q}} = 4\pi e^2/q^2$ is the Fourier transform of the Coulomb potential. As in the first-quantized picture, the Hamiltonian generates the dynamics of the states, see below.

4.7 Correlations in the Fermi Gas

We briefly summarize the description of the Fermi gas (cf. Section 2.5.7). On account of the translational invariance of the system, we define creation operators

$$\langle x | c_{\boldsymbol{k}\sigma}^\dagger | 0 \rangle = \langle \boldsymbol{r}, s | c_{\boldsymbol{k}\sigma}^\dagger | 0 \rangle = \varphi_{\boldsymbol{k}}(\boldsymbol{r}) \chi_\sigma(s) \tag{4.78}$$

for the plane waves

$$\varphi_{\boldsymbol{k}}(\boldsymbol{r}) \chi_\sigma(s) = \frac{1}{\sqrt{\Omega}}\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\, \delta_{\sigma s}. \tag{4.79}$$

The field operator then has the form

$$\Psi_s(\boldsymbol{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\, c_{\boldsymbol{k}s}\,. \tag{4.80}$$

The Fermi sea (in which all states up to the Fermi wave vector k_F are filled)

$$|\Phi_0\rangle = \prod_{\substack{k\le k_F \\ s\in\{\uparrow,\downarrow\}}} c^\dagger_{\boldsymbol{k}s}|0\rangle, \tag{4.81}$$

is the ground state of the Hamiltonian

$$H = \sum_{\boldsymbol{k},s} \varepsilon_{\boldsymbol{k}} c^\dagger_{\boldsymbol{k}s} c_{\boldsymbol{k}s} = \sum_{\boldsymbol{k},s} \frac{\hbar^2 k^2}{2m} c^\dagger_{\boldsymbol{k}s} c_{\boldsymbol{k}s}. \tag{4.82}$$

at fixed particle density $n = N/\Omega$. [11] The Fermi sea is defined by the occupation numbers (Fermi statistics, Pauli exclusion principle, cf. Figure 4.1)

$$n_{\boldsymbol{k}s} = \langle\Phi_0|c^\dagger_{\boldsymbol{k}s} c_{\boldsymbol{k}s}|\Phi_0\rangle = \begin{cases} 1, & k \le k_F, \\ 0, & \text{otherwise.} \end{cases} \tag{4.83}$$

The parameters k_F and ε_F (Fermi wave vector and -energy) are determined by the density n of the Fermi gas according to

$$k_F = (3\pi^2 n)^{1/3}, \qquad \varepsilon_F = \frac{\hbar^2 k_F^2}{2m}. \tag{4.84}$$

As discussed earlier, typical values for $k_F \simeq 1\,\text{Å}^{-1}$ (at $n = N/\Omega \simeq 10^{23}\text{cm}^{-3}$ in metals) and correspondingly $\varepsilon_F \simeq$ eV. The density of states (per spin) $\rho(\varepsilon) = \Omega^{-1}\sum_{\boldsymbol{k}} \delta(\varepsilon - \varepsilon_{\boldsymbol{k}})$ is given by (cf. Figure 4.1(b))

$$\rho(\varepsilon) = \frac{4\pi k^2}{(2\pi)^3 (d\epsilon/dk)} = \frac{m}{2\pi^2\hbar^3}\sqrt{2m\varepsilon} = \frac{3}{4}\frac{n}{\varepsilon_F}\left(\frac{\varepsilon}{\varepsilon_F}\right)^{1/2} \tag{4.85}$$

and its value at the Fermi edge is

$$\rho_0 = \rho(\varepsilon_F) = \frac{3}{4}\frac{n}{\varepsilon_F} \quad \text{(density of states/spin)}, \qquad 2\rho_0 = \text{density of states}. \tag{4.86}$$

As an exercise, one can determine the density of states for the Fermi gas in one dimension ($\rho_{1\text{D}} \propto 1/\sqrt{\varepsilon}$) and in two dimensions ($\rho_{2\text{D}} \propto$ constant.) In the following we calculate the density matrix $\langle\Phi_0|\Psi^\dagger_s(\boldsymbol{r})\Psi_s(\boldsymbol{r}')|\Phi_0\rangle$ and the pair correlations $\langle\Phi_0|\Psi^\dagger_s(\boldsymbol{r})\Psi^\dagger_{s'}(\boldsymbol{r}')\Psi_{s'}(\boldsymbol{r}')\Psi_s(\boldsymbol{r})|\Phi_0\rangle$ in the Fermi gas.

[11] Alternatively, one can introduce the grand-canonical Hamiltonian $H_{\text{gk}} = H - \mu\sum_{\boldsymbol{k}\sigma} c^\dagger_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma}$, so that Φ_0 with $\epsilon_F = \mu$ is the ground state in the entire Fock space.

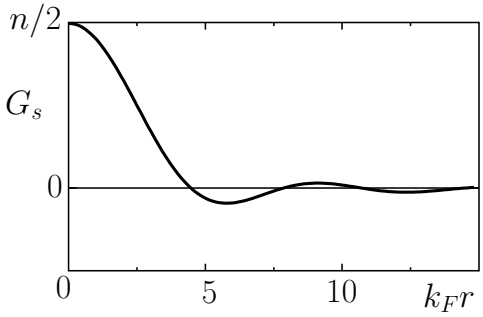


Figure 4.2: Density matrix $G_s(k_F r)$.

4.7.1 Density Matrix

The local density operator for electrons with spin s is given by $\rho_s(\boldsymbol{r}) = \Psi_s^\dagger(\boldsymbol{r})\Psi_s(\boldsymbol{r})$. The density matrix of the Fermi sea involves the generalization to different positions $\boldsymbol{r}'$ and $\boldsymbol{r}'$ (non-diagonal density operator),

$$G_s(\boldsymbol{r}-\boldsymbol{r}') \equiv \langle\Phi_0|\Psi_s^\dagger(\boldsymbol{r})\Psi_s(\boldsymbol{r}')|\Phi_0\rangle \tag{4.87}$$

$$= \sum_{\boldsymbol{k},\boldsymbol{k}'} \frac{\exp(-i\boldsymbol{k}\cdot\boldsymbol{r})}{\sqrt{\Omega}} \frac{\exp(i\boldsymbol{k}\cdot\boldsymbol{r}')}{\sqrt{\Omega}} \underbrace{\langle\Phi_0|c_{\boldsymbol{k}s}^\dagger c_{\boldsymbol{k}'s}|\Phi_0\rangle}_{\delta_{\boldsymbol{k}\boldsymbol{k}'}n_{\boldsymbol{k}s}} . \tag{4.88}$$

The quantity $G_s(\boldsymbol{r}-\boldsymbol{r}')$ is the probability amplitude to remove an electron at $\boldsymbol{r}$ and to reinsert it at $\boldsymbol{r}'$. Evaluating the Kronecker $\delta_{\boldsymbol{k}\boldsymbol{k}'}$ symbol and going to the continuum leads to

$$\begin{aligned} G_s(\boldsymbol{r}-\boldsymbol{r}') &= \frac{1}{\Omega}\sum_{\boldsymbol{k}} e^{-i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')} n_{\boldsymbol{k}s} \\ &= \int \frac{d^3k}{(2\pi)^3}\, e^{-i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')}\Theta(k_F-k) \\ &= \frac{1}{4\pi^2}\int_0^{k_F} dk\, k^2 \int_{-1}^{1} dz\, e^{-ik|\boldsymbol{r}-\boldsymbol{r}'|z} \\ &= \frac{3n}{2}\frac{\sin x - x\cos x}{x^3}\bigg|_{x=k_F|\boldsymbol{r}-\boldsymbol{r}'|} = \frac{3n}{2}\frac{j_1(k_F|\boldsymbol{r}-\boldsymbol{r}'|)}{k_F|\boldsymbol{r}-\boldsymbol{r}'|}, \end{aligned} \tag{4.89}$$

with the spherical Bessel function j_1. The result is shown in Figure 4.2; the limiting values $x = k_F r \to 0, \infty$ are given by

$$\begin{aligned} &x \to 0 : (x - x^3/6 - x + x^3/2)/x^3|_{x\to 0} = 1/3 \Rightarrow G_s(0) = n/2 \\ &x \to \infty : G_s(x) \sim 3n(\cos x)/2x^2. \end{aligned} \tag{4.90}$$

With $G_s(0) = n/2$ we find a homogeneous density $n = G_\uparrow(0) + G_\downarrow(0)$ of the particles in the Fermi gas.

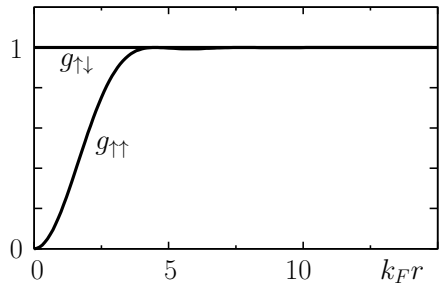


Figure 4.3: Paircorrelators $g_{\uparrow\uparrow}(k_F r)$ (with exchange hole due to the Pauli principle) and $g_{\uparrow\downarrow}(k_F r)$ (uncorrelated, since spin $\uparrow$ and spin $\downarrow$ are distinguishable).

4.8 Pair Correlations

The pair correlation

$$\left(\frac{n}{2}\right)^2 g_{ss'}(\boldsymbol{r}-\boldsymbol{r}') = \langle\Phi_0|\Psi_s^\dagger(\boldsymbol{r})\Psi_{s'}^\dagger(\boldsymbol{r}')\Psi_{s'}(\boldsymbol{r}')\Psi_s(\boldsymbol{r})|\Phi_0\rangle. \tag{4.91}$$

gives us the relative probability to find an electron at $\boldsymbol{r}', s'$ if there is already an electron at $\boldsymbol{r}, s$. The calculation of the pair correlator is a nice exercise in the use of the second-quantized formalism. We first express the field operators $\Psi_s(\boldsymbol{r})$ in terms of the annihilation operators $c_{\boldsymbol{k}s}$,

$$\left(\frac{n}{2}\right)^2 g_{ss'}(\boldsymbol{r}-\boldsymbol{r}') = \frac{1}{\Omega^2}\sum_{\boldsymbol{p},\boldsymbol{p}',\boldsymbol{q},\boldsymbol{q}'} e^{-i(\boldsymbol{p}-\boldsymbol{p}')\cdot\boldsymbol{r}} e^{-i(\boldsymbol{q}-\boldsymbol{q}')\cdot\boldsymbol{r}'} \langle\Phi_0|c_{\boldsymbol{p}s}^\dagger c_{\boldsymbol{q}s'}^\dagger c_{\boldsymbol{q}'s'} c_{\boldsymbol{p}'s}|\Phi_0\rangle.$$

We begin with the case $s \neq s'$: Then $\boldsymbol{q}' = \boldsymbol{q}$ and $\boldsymbol{p}' = \boldsymbol{p}$ must hold, since otherwise the expectation value $\langle\dots\rangle = 0$ vanishes.[12] The expectation value $\langle\dots\rangle$ leads to

$$\begin{aligned}\langle\Phi_0|c_{\boldsymbol{p}s}^\dagger c_{\boldsymbol{q}s'}^\dagger c_{\boldsymbol{q}'s'} c_{\boldsymbol{p}'s}|\Phi_0\rangle &= \delta_{\boldsymbol{p}\boldsymbol{p}'}\delta_{\boldsymbol{q}\boldsymbol{q}'}\langle\Phi_0|c_{\boldsymbol{p}s}^\dagger c_{\boldsymbol{q}s'}^\dagger c_{\boldsymbol{q}s'} c_{\boldsymbol{p}s}|\Phi_0\rangle = \delta_{\boldsymbol{p}\boldsymbol{p}'}\delta_{\boldsymbol{q}\boldsymbol{q}'}\langle\Phi_0|c_{\boldsymbol{p}s}^\dagger n_{\boldsymbol{q}s'} c_{\boldsymbol{p}s}|\Phi_0\rangle \\ &= \delta_{\boldsymbol{p}\boldsymbol{p}'}\delta_{\boldsymbol{q}\boldsymbol{q}'} n_{\boldsymbol{p}s} n_{\boldsymbol{q}s'}, \end{aligned} \tag{4.92}$$

$$\left(\frac{n}{2}\right)^2 g_{ss'}(\boldsymbol{r}) = \frac{1}{\Omega^2}\sum_{\boldsymbol{p}\boldsymbol{q}} n_{\boldsymbol{p}s} n_{\boldsymbol{q}s'} = n_s n_{s'} \overset{n_s=\frac{n}{2}}{=} \left(\frac{n}{2}\right)^2 \cdot 1, \tag{4.93}$$

$$\Rightarrow g_{s\neq s'}(\boldsymbol{r}) = 1; \tag{4.94}$$

we thus find for $s \neq s'$ a constant value $g_{s\neq s'}(k_F|\boldsymbol{r}-\boldsymbol{r}'|) = 1$. Next, we calculate the case $s = s'$. Here we obtain contributions for $\boldsymbol{p}' = \boldsymbol{p}$ and $\boldsymbol{q}' = \boldsymbol{q}$ or for $\boldsymbol{p}' = \boldsymbol{q}$ and $\boldsymbol{q}' = \boldsymbol{p}$,

$$\begin{aligned}\langle\Phi_0|c_{\boldsymbol{p}s}^\dagger c_{\boldsymbol{q}s}^\dagger c_{\boldsymbol{q}'s} c_{\boldsymbol{p}'s}|\Phi_0\rangle &= \delta_{\boldsymbol{p}\boldsymbol{p}'}\delta_{\boldsymbol{q}\boldsymbol{q}'}\langle\Phi_0|c_{\boldsymbol{p}s}^\dagger c_{\boldsymbol{q}s}^\dagger c_{\boldsymbol{q}s} c_{\boldsymbol{p}s}|\Phi_0\rangle + \delta_{\boldsymbol{q}\boldsymbol{p}'}\delta_{\boldsymbol{p}\boldsymbol{q}'}\langle\Phi_0|c_{\boldsymbol{p}s}^\dagger c_{\boldsymbol{q}s}^\dagger \overbrace{c_{\boldsymbol{p}s} c_{\boldsymbol{q}s}}^{-c_{\boldsymbol{q}s}c_{\boldsymbol{p}s}}|\Phi_0\rangle \\ &= \left(\delta_{\boldsymbol{p}\boldsymbol{p}'}\delta_{\boldsymbol{q}\boldsymbol{q}'} - \delta_{\boldsymbol{p}'\boldsymbol{q}}\delta_{\boldsymbol{q}'\boldsymbol{p}}\right) n_{\boldsymbol{p}s} n_{\boldsymbol{q}s}, \end{aligned} \tag{4.95}$$

[12] The expectation value $\langle\dots\rangle$ vanishes for $\boldsymbol{q}' \neq \boldsymbol{q}$ or $\boldsymbol{p}' \neq \boldsymbol{p}$ because different states in $|\Phi_0\rangle$ and in $c_{\boldsymbol{p}s}^\dagger c_{\boldsymbol{q}s'}^\dagger c_{\boldsymbol{q}'s'} c_{\boldsymbol{p}'s}|\Phi_0\rangle$ are occupied.

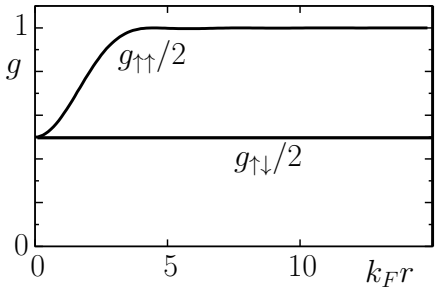


Figure 4.4: Pair correlator $g = [g_{\uparrow\uparrow} + g_{\uparrow\downarrow}]/2$ of the entire electronic system.

$$\begin{aligned}\left(\frac{n}{2}\right)^2 g_{ss}(\boldsymbol{r}) &= \frac{1}{\Omega^2}\sum_{\boldsymbol{pq}}\left(1-e^{-i(\boldsymbol{p}-\boldsymbol{q})\cdot\boldsymbol{r}}\right)n_{\boldsymbol{ps}}n_{\boldsymbol{qs}} = \left(\frac{n}{2}\right)^2 - \left(\frac{1}{\Omega}\sum_{\boldsymbol{p}} e^{-i\boldsymbol{p}\cdot\boldsymbol{r}} n_{\boldsymbol{ps}}\right)^2 \\ &= \left(\frac{n}{2}\right)^2 - G_s^2(\boldsymbol{r}) \\ &\overset{x=k_F r}{=} \left(\frac{n}{2}\right)^2\left[1-\frac{9}{x^6}\left(\sin x - x\cos x\right)^2\right],\end{aligned} \tag{4.96}$$

The correlator $g_{ss}(\boldsymbol{r}-\boldsymbol{r}')$ exhibits an exchange hole. Two fermions with identical spin avoid each other according to the Pauli exclusion principle; the exchange hole displaces exactly one electron with spin s from the vicinity of an electron with spin s, since

$$\frac{n}{2}\int d^3r\left[g_{ss}(\boldsymbol{r})-1\right] = -\underbrace{\frac{2\pi n}{k_F^3}}_{2/3\pi}\underbrace{\int_0^\infty dx\, x^2\,\frac{9}{x^6}\left(\sin x - x\cos x\right)^2}_{3\pi/2} = -1 \tag{4.97}$$

The two results are shown graphically in Figure 4.3. The pair correlator $g(\boldsymbol{r}) = \frac{1}{2}\left(g_{\uparrow\uparrow}(\boldsymbol{r}) + g_{\uparrow\downarrow}(\boldsymbol{r})\right)$ of the entire system is shown in Figure 4.4.

The above results were computed at the Hartree-Fock level: optimal Slater determinants for the homogeneous Fermi gas are built from translationally invariant single-particle states. The state $|\Phi_0\rangle$ is thus precisely a Slater determinant. If one additionally includes correlations, one finds corrections as schematically shown in Figure 4.5.

One can also compute the result simply with Wick's theorem The correlator $g_{ss'}(\boldsymbol{r}-\boldsymbol{r}')$ is determined by "contracting" the field operators (the contraction is indicated by the corresponding brackets). For the first term in (4.95)

$$\langle\Phi_0|\Psi_s^\dagger(\boldsymbol{r})\overbrace{\Psi_{s'}^\dagger(\boldsymbol{r}')\Psi_{s'}(\boldsymbol{r}')}\Psi_s(\boldsymbol{r})|\Phi_0\rangle$$

and for the second term

$$\langle\Phi_0|\Psi_s^\dagger(\boldsymbol{r})\overbrace{\Psi_{s'}^\dagger(\boldsymbol{r}')\Psi_{s'}(\boldsymbol{r}')\Psi_s(\boldsymbol{r})}|\Phi_0\rangle.$$

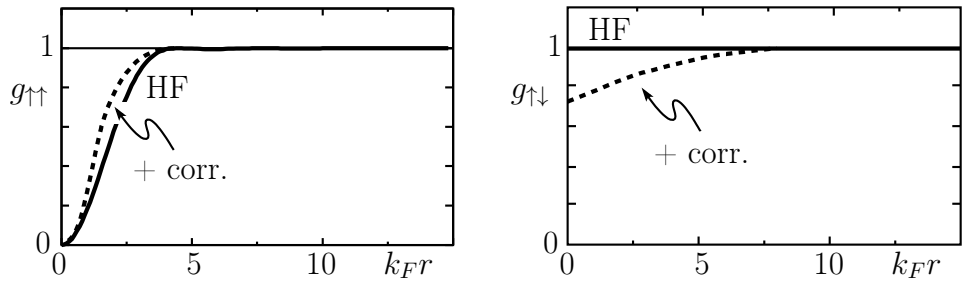


Figure 4.5: Corrections (schematic) to the pair correlators $g_{\uparrow\uparrow}$ and $g_{\uparrow\downarrow}$ upon inclusion of correlation effects.

The contractions correspond to expectation values in the ground state. Thus one obtains the two terms[13]

$$\left(\frac{n}{2}\right)^2 g_{ss'}(\boldsymbol{r}-\boldsymbol{r}') = \langle\Phi_0|\Psi_s^\dagger(\boldsymbol{r})\Psi_{s'}^\dagger(\boldsymbol{r}')\Psi_{s'}(\boldsymbol{r}')\Psi_s(\boldsymbol{r})|\Phi_0\rangle \tag{4.98}$$
$$= \langle\Phi_0|\Psi_s^\dagger(\boldsymbol{r})\Psi_s(\boldsymbol{r})|\Phi_0\rangle\langle\Phi_0|\Psi_{s'}^\dagger(\boldsymbol{r}')\Psi_{s'}(\boldsymbol{r}')|\Phi_0\rangle$$
$$- \langle\Phi_0|\Psi_s^\dagger(\boldsymbol{r})\Psi_{s'}(\boldsymbol{r}')|\Phi_0\rangle\langle\Phi_0|\Psi_{s'}^\dagger(\boldsymbol{r}')\Psi_s(\boldsymbol{r})|\Phi_0\rangle.$$

With this, one obtains the result directly from $\langle\Phi_0|\Psi_s^\dagger(\boldsymbol{r})\ \Psi_{s'}(\boldsymbol{r}')|\Phi_0\rangle = \delta_{ss'}G_s(\boldsymbol{r}-\boldsymbol{r}')$. One finds $(n/2)^2 g_{s\neq s'} = (n/2)^2 - 0$ (for $s \neq s'$) and $(n/2)^2 g_{ss} = (n/2)^2 - G_s^2$ (for $s = s'$).

The above factorization of $\langle\Phi_0|\Psi_s^\dagger\Psi_{s'}^\dagger\Psi_{s'}\Psi_s|\Phi_0\rangle$ is exact for the ground state $|\Phi_0\rangle$ in Hartree-Fock form (i.e., for a Slater determinant). Correspondingly, the factorization is valid for the noninteracting system. If one includes the interaction between the particles, the Hartree-Fock ground state is only an approximation. Correspondingly, the expectation value $\langle\Phi|\Psi_s^\dagger\Psi_{s'}^\dagger\Psi_{s'}^\dagger\Psi_s|\Phi\rangle$ in the true ground state $|\Phi\rangle$ is only approximately given by

$$\langle\Phi|\Psi_s^\dagger\Psi_{s'}^\dagger\Psi_{s'}^\dagger\Psi_s|\Phi\rangle \overset{\text{in HF}}{\approx} \langle\Phi|\Psi_s^\dagger\Psi_s|\Phi\rangle\langle\Phi|\Psi_{s'}^\dagger\Psi_{s'}|\Phi\rangle$$
$$- \langle\Phi|\Psi_s^\dagger\Psi_{s'}|\Phi\rangle\langle\Phi|\Psi_{s'}^\dagger\Psi_s|\Phi\rangle. \tag{4.99}$$

4.9 Equation of Motion

As usual in the first-quantized form of quantum mechanics, in second quantization we can also choose between the Schrödinger, Heisenberg, and interaction picture. In the Heisenberg picture the operators (and thus also the field operators $\Psi_s^\dagger(\boldsymbol{r})$ and $\Psi_s(\boldsymbol{r})$) become time dependent. For a time-independent Hamiltonian H the Heisenberg picture of an arbitrary operator $A(t)$ is given by

$$A_H(t) = U^\dagger A(t) U = e^{iHt/\hbar} A(t)\, e^{-iHt/\hbar} \qquad (\text{note } H_H(t) = H). \tag{4.100}$$

[13]The minus sign in the 2nd term arises from the exchange of the operators $\Psi_{s'}^\dagger(\boldsymbol{r}')\Psi_{s'}(\boldsymbol{r}')$ so that the two contracted operators stand next to each other.

The equation of motion of the operator $A_H(t)$ is

$$\begin{aligned} i\hbar\frac{dA_H(t)}{dt} &= \big[A_H(t), H\big] + i\hbar(\partial_t A)_H(t) \\ &= \big[A, H\big]_H(t) + i\hbar(\partial_t A)_H(t) \\ &\overset{\partial_t A=0}{=} \big[A, H\big]_H(t). \end{aligned} \tag{4.101}$$

Here, $(\partial_t A)_H(t) = U^\dagger \dot{A}(t) U$ and $[A, H]_H(t) = U^\dagger [A(t), H] U$.

We calculate the time derivative of the field operator $\Psi(\boldsymbol{r}; t) \equiv \Psi(\boldsymbol{r})_H(t) = U^\dagger \Psi(\boldsymbol{r}) U$ for a Hamiltonian H of the form

$$\begin{aligned} H = &\int d^3r\, \Psi^\dagger(\boldsymbol{r}) \frac{-\hbar^2\nabla^2}{2m} \Psi(\boldsymbol{r}) \\ &+ \frac{1}{2}\int d^3r\, d^3r'\, \Psi^\dagger(\boldsymbol{r})\Psi^\dagger(\boldsymbol{r}')V(\boldsymbol{r}-\boldsymbol{r}')\Psi(\boldsymbol{r}')\Psi(\boldsymbol{r}) \end{aligned} \tag{4.102}$$

where we have suppressed the spin quantum numbers for the field operator Ψ. From (4.101) we obtain

$$i\hbar\,\frac{\partial\Psi(\boldsymbol{r};t)}{\partial t} = \Big[\Psi(\boldsymbol{r}), H\Big]_H(t), \tag{4.103}$$

$$\begin{aligned} \Big[\Psi(\boldsymbol{r}), H\Big] = &\int d^3r' \Big[\Psi(\boldsymbol{r}), \Psi^\dagger(\boldsymbol{r}')\frac{-\hbar^2\nabla'^2}{2m}\Psi(\boldsymbol{r}')\Big] \\ &+ \frac{1}{2}\int d^3r'\, d^3r'' V(\boldsymbol{r}'-\boldsymbol{r}'')\Big[\Psi(\boldsymbol{r}), \Psi^\dagger(\boldsymbol{r}')\Psi^\dagger(\boldsymbol{r}'')\Psi(\boldsymbol{r}'')\Psi(\boldsymbol{r}')\Big]. \end{aligned} \tag{4.104}$$

To compute the commutators $[A, BC]$ we use the relations

$$[A, BC] = \begin{cases} [A,B]C + B[A,C], & \text{for bosons,} \\ \{A,B\}C - B\{A,C\}, & \text{for fermions.} \end{cases} \tag{4.105}$$

With this, for fermions one finds

$$\begin{aligned} [\Psi(\boldsymbol{r}), \Psi^\dagger(\boldsymbol{r}')\Psi(\boldsymbol{r}')] &= \underbrace{\{\Psi(\boldsymbol{r}), \Psi^\dagger(\boldsymbol{r}')\}}_{\delta^{(3)}(\boldsymbol{r}-\boldsymbol{r}')}\Psi(\boldsymbol{r}') - \Psi^\dagger(\boldsymbol{r}')\underbrace{\{\Psi(\boldsymbol{r}), \Psi(\boldsymbol{r}')\}}_{0} \\ [\Psi(\boldsymbol{r}), \Psi^\dagger(\boldsymbol{r}')\Psi^\dagger(\boldsymbol{r}'')\Psi(\boldsymbol{r}'')\Psi(\boldsymbol{r}')] &= \underbrace{\{\Psi(\boldsymbol{r}), \Psi^\dagger(\boldsymbol{r}')\}}_{\delta^{(3)}(\boldsymbol{r}-\boldsymbol{r}')}\Psi^\dagger(\boldsymbol{r}'')\Psi(\boldsymbol{r}'')\Psi(\boldsymbol{r}') \\ &\quad - \Psi^\dagger(\boldsymbol{r}')\underbrace{\{\Psi(\boldsymbol{r}), \Psi^\dagger(\boldsymbol{r}'')\Psi(\boldsymbol{r}'')\Psi(\boldsymbol{r}')\}}_{\underbrace{\{\Psi(\boldsymbol{r}), \Psi^\dagger(\boldsymbol{r}'')\}}_{\delta^{(3)}(\boldsymbol{r}-\boldsymbol{r}'')}\Psi(\boldsymbol{r}'')\Psi(\boldsymbol{r}')} \\ &= \delta^{(3)}(\boldsymbol{r}-\boldsymbol{r}')\Psi^\dagger(\boldsymbol{r}'')\Psi(\boldsymbol{r}'')\Psi(\boldsymbol{r}') - \delta^{(3)}(\boldsymbol{r}-\boldsymbol{r}'')\Psi^\dagger(\boldsymbol{r}')\Psi(\boldsymbol{r}'')\Psi(\boldsymbol{r}'). \end{aligned} \tag{4.106}$$

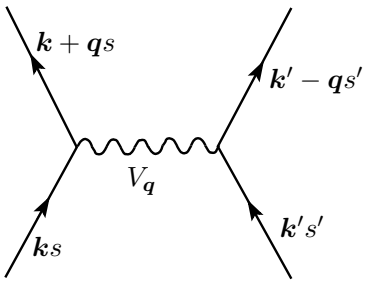


Figure 4.6: Interaction vertex for the Hamiltonian (4.110): The incoming vectors stand for the annihilation operators $c_{\boldsymbol{k}s}$, $c_{\boldsymbol{k}'s'}$, the outgoing vectors for the creation operators $c^\dagger_{\boldsymbol{k}+\boldsymbol{q}s}$, $c^\dagger_{\boldsymbol{k}'-\boldsymbol{q}s'}$. The potential $V_{\boldsymbol{q}}$ is represented by the wavy line.

Combining (4.104) to (4.106) yields

$$[\Psi(\boldsymbol{r}),H] = -\frac{\hbar^2}{2m}\nabla^2\Psi(\boldsymbol{r}) + \frac{1}{2}\int d^3r' \int d^3r''\, V(\boldsymbol{r}'-\boldsymbol{r}'') \times [\delta^{(3)}(\boldsymbol{r}-\boldsymbol{r}')\Psi^\dagger(\boldsymbol{r}'')\Psi(\boldsymbol{r}'')\Psi(\boldsymbol{r}') - \delta^{(3)}(\boldsymbol{r}-\boldsymbol{r}')\Psi^\dagger(\boldsymbol{r}'')\Psi(\boldsymbol{r}')\Psi(\boldsymbol{r}'')] \tag{4.107}$$

Exchanging $\Psi(\boldsymbol{r}')\Psi(\boldsymbol{r}'') = -\Psi(\boldsymbol{r}'')\Psi(\boldsymbol{r}')$ in the last term and using (4.103) gives the final result

$$i\hbar\frac{\partial\Psi(\boldsymbol{r};t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi(\boldsymbol{r};t) + \left(\int d^3r'\,\Psi^\dagger(\boldsymbol{r}';t)V(\boldsymbol{r}-\boldsymbol{r}')\Psi(\boldsymbol{r}';t)\right)\Psi(\boldsymbol{r};t);$$

An analogous calculation for bosons yields the same result.[14] Note that the calculation of the commutators at equal times $t = t'$ yields the results

$$\begin{aligned}[\Psi(\boldsymbol{r};t),\Psi^\dagger(\boldsymbol{r}';t)]_\pm &= \delta^{(3)}(\boldsymbol{r}-\boldsymbol{r}'),\\ [\Psi(\boldsymbol{r};t),\Psi(\boldsymbol{r}';t)]_\pm &= 0,\\ [\Psi^\dagger(\boldsymbol{r};t),\Psi^\dagger(\boldsymbol{r}';t)]_\pm &= 0,\end{aligned} \tag{4.108}$$

but that the same problem for different times t and t' is a difficult many-particle problem

$$\left.\begin{matrix}[\Psi(\boldsymbol{r};t),\Psi^\dagger(\boldsymbol{r}';t')]_\pm\\ [\Psi(\boldsymbol{r};t),\Psi(\boldsymbol{r}';t')]_\pm\\ [\Psi^\dagger(\boldsymbol{r};t),\Psi^\dagger(\boldsymbol{r}';t')]_\pm\end{matrix}\right\}\text{“many-particle problem”}, \tag{4.109}$$

which in general cannot be solved exactly, see e.g. Fetter and Walecka.

Of course, we can also work in momentum space. The Hamiltonian has the form

$$H = \sum_{\boldsymbol{k}s}\frac{\hbar^2k^2}{2m}\,c^\dagger_{\boldsymbol{k}s}c_{\boldsymbol{k}s} + \frac{1}{2}\sum_{\boldsymbol{k}\boldsymbol{k}'\boldsymbol{q}s's} V_{\boldsymbol{q}}\,c^\dagger_{\boldsymbol{k}+\boldsymbol{q}s}c^\dagger_{\boldsymbol{k}'-\boldsymbol{q}s'}c_{\boldsymbol{k}'s'}c_{\boldsymbol{k}s}, \tag{4.110}$$

[14] The reason for this is the equations (4.105), which relate a commutator between the operator A and a quadratic operator BC to a similar form in terms of the “elementary” (anti-)commutators.

see (4.77) and Figure 4.6. and the equation of motion for the operators $c_{\boldsymbol{k}s}(t) \equiv (c_{\boldsymbol{k}s})_H(t)$ is (with $V_{-\boldsymbol{q}} = V_{\boldsymbol{q}}, \varepsilon_{\boldsymbol{k}} = \hbar^2 k^2/2m$)

$$\begin{aligned} i\hbar \frac{dc_{\boldsymbol{k}s}(t)}{dt} &= \varepsilon_{\boldsymbol{k}} c_{\boldsymbol{k}s}(t) + \sum_{\boldsymbol{k}'\boldsymbol{q}s'} V_{\boldsymbol{q}}\, c^\dagger_{\boldsymbol{k}'+\boldsymbol{q}s'}(t)\, c_{\boldsymbol{k}'s'}(t) c_{\boldsymbol{k}+\boldsymbol{q}s}(t) \\ &= \varepsilon_{\boldsymbol{k}} c_{\boldsymbol{k}s}(t) + \sum_{\boldsymbol{q}} V_{\boldsymbol{q}}\, \rho(-\boldsymbol{q};t) c_{\boldsymbol{k}+\boldsymbol{q}s}(t), \end{aligned} \tag{4.111}$$

with the density operator [summed over both spin directions, cf. (4.66)]

$$\rho(-\boldsymbol{q}) = \sum_{\boldsymbol{k}s} c^\dagger_{\boldsymbol{k}+\boldsymbol{q}s} c_{\boldsymbol{k}s}\,. \tag{4.112}$$

Chapter 5

Interacting Electron Systems

In this chapter we study the influence of interactions on electron systems. We will see how the interaction between electrons in metals affects the ground-state energy of the Fermi gas within the Hartree-Fock approximation and will also compute the (dynamic) dielectric constant; this allows one to understand screening and plasmons in metals. Furthermore, we introduce the effect of superconductivity. The focus is on an application of the second quantization introduced in the previous chapter. We consider the states only at $T = 0$, so that only the (many-particle) ground state is occupied. However, because of the presence of the intrinsic energy scale E_g ($E_g = \varepsilon_F$ for the metal and $E_g = |\Delta|$ for the superconductor) the results remain valid for $T \ll E_g$.

5.1 Homogeneous Electron Gas

The homogeneous electron gas is the simplest model of a metal. The system is described by the Hamiltonian of interacting electrons, cf. (4.76),

$$H = \frac{\hbar^2}{2m} \int d^3r \sum_s \boldsymbol{\nabla} \Psi_s^\dagger(\boldsymbol{r}) \cdot \boldsymbol{\nabla} \Psi_s(\boldsymbol{r}) + \int d^3r \sum_s \Psi_s^\dagger(\boldsymbol{r}) V_{\text{ion}}(\boldsymbol{r}) \Psi_s(\boldsymbol{r}) \\ + \frac{1}{2} \int d^3r \, d^3r' \sum_{ss'} \Psi_s^\dagger(\boldsymbol{r}) \Psi_{s'}^\dagger(\boldsymbol{r}') V(\boldsymbol{r} - \boldsymbol{r}') \Psi_{s'}(\boldsymbol{r}') \Psi_s(\boldsymbol{r}). \tag{5.1}$$

The electrons interact via the Coulomb interaction $V(\boldsymbol{r}) = e^2/r$ and are embedded in a background of positively charged ions that guarantee the overall charge neutrality of the system and thereby confine the electron system. In reality, the ions form a crystal lattice that gives rise to electronic bands. For the homogeneous electron gas one makes the simplifying assumption that the positive background is smeared out homogeneously. Thus the electrons have a simple parabolic dispersion instead of a complicated band structure. As is well known from the solid-state lecture, the effects

of the ion cores can often be accounted for approximately by a material-dependent effective mass. This model is also called the Jellium model.

5.1.1 Ground-State Energy

We compute the ground-state energy E_{HF} in the Hartree-Fock approximation: for the (homogeneous) electron gas the optimal Hartree-Fock single-particle wave functions are plane waves $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\chi_\sigma(s)$ and the Fermi sea $|\Phi_0\rangle = \prod_{k<k_F,\sigma} c^\dagger_{\boldsymbol{k}\sigma}|0\rangle$ is the optimal Hartree-Fock ansatz with ground-state energy

$$E_{\mathrm{HF}} = \langle\Phi_0|H|\Phi_0\rangle = \langle T\rangle_{\mathrm{HF}} + \langle V_{\mathrm{ion}}\rangle_{\mathrm{HF}} + \langle V\rangle_{\mathrm{HF}}. \tag{5.2}$$

Here the average kinetic energy, see (4.83),

$$\begin{aligned}\langle T\rangle_{\mathrm{HF}} &= \langle\Phi_0|\sum_{\boldsymbol{k}\sigma}\frac{\hbar^2k^2}{2m}c^\dagger_{\boldsymbol{k}\sigma}c_{\boldsymbol{k}\sigma}|\Phi_0\rangle \\ &= \underbrace{2}_{\text{Spin}}\Omega\int_0^{k_F}\frac{4\pi k^2dk}{(2\pi)^3}\frac{\hbar^2k^2}{2m} \\ &= \frac{\Omega}{\pi^2}\frac{\hbar^2}{2m}\frac{1}{5}k_F^5 = \frac{\Omega}{\pi^2}\varepsilon_F\frac{k_F^3}{5} = \frac{3}{5}\varepsilon_F N,\end{aligned} \tag{5.3}$$

is precisely the energy of the Fermi gas with $N = n\Omega$ and $k_F^3 = 3\pi^2 n$. For the interaction energy $\langle V\rangle_{\mathrm{HF}}$ we obtain

$$\begin{aligned}\langle V\rangle_{\mathrm{HF}} &= \frac{1}{2}\int d^3r\,d^3r'\sum_{s,s'}V(\boldsymbol{r}-\boldsymbol{r}')\underbrace{\langle\Phi_0|\Psi^\dagger_s(\boldsymbol{r})\Psi^\dagger_{s'}(\boldsymbol{r}')\Psi_{s'}(\boldsymbol{r}')\Psi_s(\boldsymbol{r})|\Phi_0\rangle}_{(n/2)^2 g_{ss'}(\boldsymbol{r}-\boldsymbol{r}')} \\ &= \frac{1}{2}\int d^3r\,d^3r'\;V(\boldsymbol{r}-\boldsymbol{r}')\left[\left(\frac{n}{2}\right)^2\cdot\underbrace{2}_{s\neq s'} + \left(\frac{n}{2}\right)^2(1-g_s^2)\underbrace{2}_{s=s'}\right] \\ &= \frac{n^2}{2}\Omega\int d^3r\,V(\boldsymbol{r}) - \frac{n^2}{4}\Omega\int d^3r\,V(\boldsymbol{r})g_s(\boldsymbol{r})^2,\end{aligned} \tag{5.4}$$

with

$$g_s(\boldsymbol{r}) = \frac{2}{n}G_s(r) = \frac{3}{x^3}(\sin x - x\cos x)\Big|_{x=k_Fr}. \tag{5.5}$$

The first term in (5.4) is the Hartree energy; it accounts for the repulsion between the electrons. The second term originates in the symmetry properties of $|\Phi_0\rangle$ and is called the exchange energy. Both terms can be visualized by Feynman diagrams[1]—see

[1] Feynman diagrams arise from a perturbative expansion of a quantity (e.g. the full propagator (including interaction), the energy, ...) in the interaction term, where expectation values $\langle c^\dagger c^\dagger\cdots c\rangle$ are decomposed into products of free propagators $\langle c^\dagger c\rangle$ by means of Wick's theorem. Each free propagator then corresponds to a line in the Feynman diagram; curly lines describe the interaction V, fermionic loops (as in the Hartree term) carry additional signs determined by $\boldsymbol{k}$ of the internal lines, one integrates, etc. Once the rules for constructing diagrams have been set up (they follow from the problem-specific expansion in a small parameter), one can write down an analytic expression for each diagram.

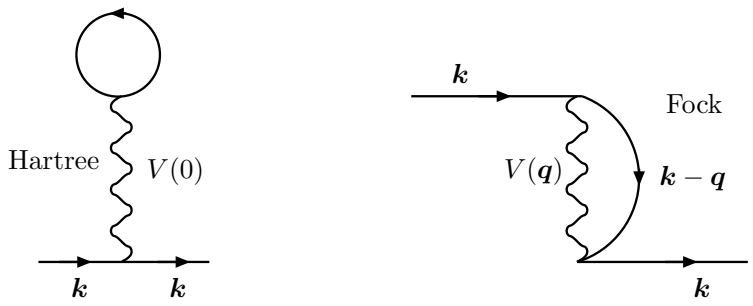


Figure 5.1: Feynman diagrams for the Hartree (left) and the Fock (exchange) term (right).

Figure 5.1.

For the Coulomb interaction $V(r) = e^2/r$ the interaction integral $\int d^3r\, V(r) = 4\pi \int dr\, r^2 e^2/r$ is divergent.[2] We have not yet included, however, the positively charged background produced by the ions, which has the same charge density as the electron gas. This background contributes additional terms to the Hartree energy. Altogether we have the Hartree energy

$$\langle V\rangle_{\rm H} + \langle V_{\rm ion}\rangle_{\rm H} = \frac{Nn}{2}\int d^3r\,\frac{e^2}{r} + \langle V_{\rm ion}\rangle_{\rm H} = \frac{Nn}{2}\int d^3r\,\left(\frac{e^2}{r} + \frac{e^2}{r} - 2\frac{e^2}{r}\right) = 0, \quad (5.6)$$

where the first term accounts for the repulsive electron-electron interaction, the second term the (also repulsive) ion-ion interaction, and the third term the interaction between electrons and ions. The vanishing of the Hartree potential means that the classical contribution of the Coulomb energy does not lead to metallic binding (energy lowering relative to the vacuum). It is the quantum-mechanical exchange effects (Fock term), which are rooted in the fundamental indistinguishability of the electrons, that yield a finite binding energy.

The Fock term contributes

$$\begin{aligned}\langle V\rangle_{\rm HF} &= -\frac{9}{4}nN\frac{e^2}{k_F^2}4\pi \underbrace{\int_0^\infty dx\,\frac{(\sin x - x\cos x)^2}{x^5}}_{1/4} \\ &= -\frac{3}{4\pi}Ne^2k_F \end{aligned} \quad (5.7)$$

to the ground-state energy.

[2]The divergence is a consequence of the long-range nature of the Coulomb interaction and implies that the electron system is incompressible; only if the compensating (here ionic) background follows the deformation does the overall system become compressible. The finite energy of the plasmons in the limit $\boldsymbol{k} \to 0$ is a consequence of this incompressibility. If, by contrast, the ions oscillate (in a phononic excitation), the energy of the mode vanishes for $\boldsymbol{k} \to 0$.

To describe an electron gas/liquid one usually introduces the Seitz radius r_s, which gives the volume available per electron. Measuring r_s in bohr radii a_B we obtain

$$\frac{4\pi}{3}(r_s a_B)^3 = \frac{1}{n}, \qquad r_s = \frac{5.44}{\sqrt[3]{n_{22}}}, \tag{5.8}$$

where $n = n_{22} \cdot 10^{22}\,\mathrm{cm}^{-3} = n_{22} \cdot 10\,\mathrm{nm}^{-3}$ denotes the electron density. For typical metals

$$n_{22} \approx \underbrace{0.9}_{\text{Caesium}} - \underbrace{2.5}_{\text{Beryllium}}, \tag{5.9}$$

and hence r_s takes values between $r_s \approx 2-6$, i.e. metals are neither at high nor at low density but at intermediate densities, which makes the problem difficult because one cannot apply perturbation theory.

For the energy (per particle) of the electron gas in the Hartree-Fock approximation we obtain

$$\frac{E_{\mathrm{HF}}}{N} = \frac{3}{5}\varepsilon_F - \frac{3}{4\pi}e^2 k_F \approx \left(\frac{2.21}{r_s^2} - \frac{0.916}{r_s}\right)\mathrm{Ry}. \tag{5.10}$$

Note that the exact ground-state energy $E_0/N < E_{\mathrm{HF}}/N$ satisfies $|\Phi_0\rangle$ is a Rayleigh-Ritz ansatz for the true ground state $|\Phi\rangle$.

The result (5.10) is correct for high densities and correspondingly small radii r_s. Not included in (5.10) are the correlation energies. They are computed by diagrammatic techniques and one finds for $r_s < 2.5$ the expansion

$$\frac{E_0}{N\,\mathrm{Ry}} \approx \underbrace{\frac{2.21}{r_s^2} - \frac{0.916}{r_s}}_{E_{\mathrm{HF}}/N\,\mathrm{Ry}} \underbrace{+0.0622\,\ln r_s - 0.094 + 0.018\,r_s \ln r_s + \cdots}_{E_{\mathrm{corr}}/N\,\mathrm{Ry}<0}. \tag{5.11}$$

Note that the result (5.11) $r_s \to 0$ is not analytic.

For small densities $r_s \to \infty$ one expects a transition to a crystal named after Wigner. For the correlation energy of the Wigner crystal one obtains

$$\frac{E_{\mathrm{korr}}}{N} \approx -\frac{0.88}{r_s}\mathrm{Ry}; \tag{5.12}$$

An interpolation[3] then yields the expression

$$\frac{E_{\mathrm{korr}}}{N} \approx -\frac{0.88}{r_s + 7.8}\mathrm{Ry} < 0. \tag{5.13}$$

Figure 5.2 finally shows the expected phases in a system with an interaction of the form $e^{-r/\lambda}/r$ interpolating between long ($\lambda \to \infty$) and short ($\lambda \to 0$) range.

[3]The interpolation between two phases—the Fermi gas and the Wigner crystal—should be taken with caution.

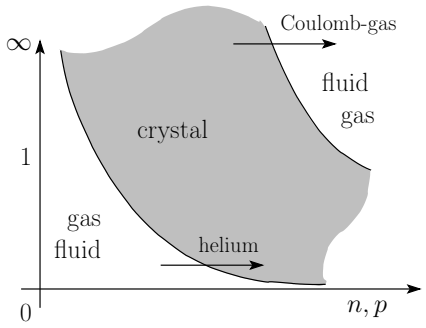


Figure 5.2: Phases in a system with (screened) Coulomb interaction; $\lambda =$ range, $n =$ density, $p =$ pressure, dimensionless parameter $= n\lambda^3$. Such a phase diagram with reentrance into the liquid/gas phase occurs, e.g. in helium.

5.1.2 Dielectric Constant $\varepsilon(\boldsymbol{q};\omega)$

It is not possible to determine the dielectric constant of the homogeneous electron gas exactly. However, there are good approximations that capture the essential properties of a metal—screening and plasmons. One such approximation we already encountered in the chapter on atoms as Thomas-Fermi screening. Although the approximation was developed for atoms, at no point in the derivation of the dielectric constant is

$$\varepsilon(\boldsymbol{q}) = 1 + \frac{q_{\mathrm{TF}}^2}{q^2} \tag{5.14}$$

with $q_{\mathrm{TF}}^2 = 4k_F/\pi a_B = 6\pi e^2 n/\varepsilon_F$ taken with reference to the atom. Rather, as already noted in 3.3, the result is valid for metals and describes their screening effect. With the Thomas-Fermi method, however, one cannot describe dynamic effects. To do so one needs the RPA (random phase approximation), which is introduced in the following.

We investigate a system of interacting electrons in an external potential $U(\boldsymbol{r};t)$, which is (partially) screened by the potential $V_s(\boldsymbol{r};t)$ as in Chap. 3.3. The potential $V_s(\boldsymbol{r};t)$ is generated internally by the electrons in the metal so that ultimately the total field $V = U + V_s$ is present. The dynamical dielectric function is then determined by, cf. (3.33),

$$\varepsilon(\boldsymbol{q};\omega) = \frac{U(\boldsymbol{q};\omega)}{V(\boldsymbol{q};\omega)} = 1 - \frac{V_s(\boldsymbol{q};\omega)}{V(\boldsymbol{q};\omega)}. \tag{5.15}$$

We wish to solve the problem self-consistently and therefore assume that the electrons are exposed to the total field V. The effective (time-dependent) Hamiltonian of the electrons is given by

$$\begin{aligned} H_{\mathrm{eff}} &= \int d^3r \sum_s \left[\frac{\hbar^2}{2m} \boldsymbol{\nabla}\Psi_s^\dagger(\boldsymbol{r}) \cdot \boldsymbol{\nabla}\Psi_s(\boldsymbol{r}) + \Psi_s^\dagger(\boldsymbol{r}) V(\boldsymbol{r};t) \Psi_s(\boldsymbol{r}) \right] \\ &= \sum_{\boldsymbol{k}s} \varepsilon_{\boldsymbol{k}} c_{\boldsymbol{k}s}^\dagger c_{\boldsymbol{k}s} + \frac{1}{\Omega} \sum_{\boldsymbol{q}} V(\boldsymbol{q};t)\rho(-\boldsymbol{q}) \end{aligned} \tag{5.16}$$

with the volume $\Omega = L^3$ and the (Fourier-transformed) density operator $\rho(-\boldsymbol{q}) = \sum_{\boldsymbol{k}s} c_{\boldsymbol{k}+\boldsymbol{q}s}^\dagger c_{\boldsymbol{k}s}$, cf. (4.112). Note that the electron-electron interaction does not appear

directly in the Hamiltonian H_{eff}. Its contribution is taken into account indirectly by the screening part V_s to V; i.e., we include screening effects but neglect correlation and exchange effects.

We assume that the external potential $U(\boldsymbol{r};t)$ generates a classically oscillating field. We seek the quantum-mechanical response of the electron system to this classical potential. We will assume U to be small and consider the problem only to first order in U; this is referred to as linear response. An oscillating potential $U(\boldsymbol{r};t) = U(\boldsymbol{r};\omega)e^{-i\omega t}$ with frequency ω then leads to a response $\langle\rho(\boldsymbol{r};t)\rangle = \rho(\boldsymbol{r};\omega)e^{-i\omega t}$ at the same frequency. Furthermore, the superposition principle holds, so that we need only compute the response to a single frequency and the general result follows by simple addition of the various contributions at different frequencies.

For the homogeneous electron gas the expectation value of $\rho(\boldsymbol{q})$ vanishes (except for $\boldsymbol{q} = 0$). Owing to the applied potential U this expectation value becomes finite. The screening charge $-e\langle\rho\rangle$ produces the potential

$$V_s(\boldsymbol{q};\omega) = \frac{4\pi e^2}{q^2}\langle\rho(\boldsymbol{q};\omega)\rangle. \tag{5.17}$$

To obtain an explicit expression for the screening potential V_s, we will solve the equation of motion for the charge density approximately,

$$\begin{aligned} i\hbar\frac{d}{dt}\rho(-\boldsymbol{q};t) &= [\rho(-\boldsymbol{q};t), H_{\text{eff}}] \\ \Rightarrow \quad \hbar\omega\rho(-\boldsymbol{q};\omega) &= [\rho(-\boldsymbol{q};\omega), H_{\text{eff}}], \end{aligned} \tag{5.18}$$

Instead of solving the equation of motion for ρ directly, we first consider the equation of motion of a term

$$\hbar\omega c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\sigma}c_{\boldsymbol{p}\sigma} = [c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\sigma}c_{\boldsymbol{p}\sigma}, H_{\text{eff}}]. \tag{5.19}$$

The result for $\rho(-\boldsymbol{q})$ can then be obtained by summation over $\boldsymbol{p}$ and σ. The commutator on the right-hand side of (5.18) is given by the sum of the two terms

$$\begin{aligned} &\sum_{\boldsymbol{k}s}\varepsilon_{\boldsymbol{k}}[c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\sigma}c_{\boldsymbol{p}\sigma}, c^\dagger_{\boldsymbol{k}s}c_{\boldsymbol{k}s}] = (\varepsilon_{\boldsymbol{p}} - \varepsilon_{\boldsymbol{p}+\boldsymbol{q}})c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\sigma}c_{\boldsymbol{p}\sigma}, \\ &\sum_{\boldsymbol{q}'\boldsymbol{k}s}V(\boldsymbol{q}')[c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\sigma}c_{\boldsymbol{p}\sigma}, c^\dagger_{\boldsymbol{k}+\boldsymbol{q}'s}c_{\boldsymbol{k}s}] = \sum_{\boldsymbol{q}'}V(\boldsymbol{q}')(c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\sigma}c_{\boldsymbol{p}-\boldsymbol{q}'\sigma} - c^\dagger_{\boldsymbol{p}+\boldsymbol{q}+\boldsymbol{q}'\sigma}c_{\boldsymbol{p}\sigma}). \end{aligned} \tag{5.20}$$

In summary, we obtain the equation of motion

$$\begin{aligned} (\varepsilon_{\boldsymbol{p}+\boldsymbol{q}} - \varepsilon_{\boldsymbol{p}} + \hbar\omega)c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\sigma}c_{\boldsymbol{p}\sigma} &= \frac{1}{\Omega}\sum_{\boldsymbol{q}'}V(\boldsymbol{q}';\omega)(c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\sigma}c_{\boldsymbol{p}-\boldsymbol{q}'\sigma} - c^\dagger_{\boldsymbol{p}+\boldsymbol{q}+\boldsymbol{q}'\sigma}c_{\boldsymbol{p}\sigma}) \\ &\overset{\text{RPA}}{\approx} \frac{V(-\boldsymbol{q};\omega)}{\Omega}(c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\sigma}c_{\boldsymbol{p}+\boldsymbol{q}\sigma} - c^\dagger_{\boldsymbol{p}\sigma}c_{\boldsymbol{p}\sigma}), \end{aligned} \tag{5.21}$$

where in the last step we have performed the RPA by retaining only the term with $\boldsymbol{q}' = -\boldsymbol{q}$ from the sum and neglecting the other terms. This assumption is justified by

the fact that terms of the form $c^\dagger_{\boldsymbol{k}} c_{\boldsymbol{k}'}$ with $e^{i(\boldsymbol{k}'-\boldsymbol{k})\cdot\boldsymbol{r}}$ oscillate so that their contribution (after integration over a small volume element) cancels out, provided that $\boldsymbol{k}' \neq \boldsymbol{k}$.

The approximate equation can now be solved for $c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\sigma} c_{\boldsymbol{p}\sigma}$. Summing the result over $\boldsymbol{p}\sigma$ yields

$$\begin{aligned}\rho(-\boldsymbol{q};\omega) &= \frac{V(-\boldsymbol{q};\omega)}{\Omega} \sum_{\boldsymbol{p}\sigma} \frac{c^\dagger_{\boldsymbol{p}+\boldsymbol{q}\sigma} c_{\boldsymbol{p}+\boldsymbol{q}\sigma} - c^\dagger_{\boldsymbol{p}\sigma} c_{\boldsymbol{p}\sigma}}{\varepsilon_{\boldsymbol{p}+\boldsymbol{q}} - \varepsilon_{\boldsymbol{p}} + \hbar\omega} \\ &\overset{(\boldsymbol{p}\mapsto\boldsymbol{p}-\boldsymbol{q})}{=} \frac{V(-\boldsymbol{q};\omega)}{\Omega} \sum_{\boldsymbol{p}\sigma} \frac{c^\dagger_{\boldsymbol{p}-\boldsymbol{q}\sigma} c_{\boldsymbol{p}-\boldsymbol{q}\sigma} - c^\dagger_{\boldsymbol{p}\sigma} c_{\boldsymbol{p}\sigma}}{\varepsilon_{\boldsymbol{p}-\boldsymbol{q}} - \varepsilon_{\boldsymbol{p}} - \hbar\omega}.\end{aligned} \tag{5.22}$$

From this we can now compute the expectation value in the ground state and obtain ($\Theta(x)$ is the step function)

$$\begin{aligned}\langle\rho(\boldsymbol{q};\omega)\rangle &= \frac{V(\boldsymbol{q};\omega)}{\Omega} \sum_{\boldsymbol{p}\sigma} \frac{\langle n_{\boldsymbol{p}+\boldsymbol{q}\sigma}\rangle - \langle n_{\boldsymbol{p}\sigma}\rangle}{\varepsilon_{\boldsymbol{p}+\boldsymbol{q}} - \varepsilon_{\boldsymbol{p}} - \hbar\omega} \\ &= V(\boldsymbol{q};\omega) \underbrace{2\int \frac{d^3p}{(2\pi)^3} \frac{\Theta(\varepsilon_F - \varepsilon_{\boldsymbol{p}+\boldsymbol{q}}) - \Theta(\varepsilon_F - \varepsilon_{\boldsymbol{p}})}{\varepsilon_{\boldsymbol{p}+\boldsymbol{q}} - \varepsilon_{\boldsymbol{p}} - \hbar\omega - i0^+}}_{\chi(\boldsymbol{q};\omega)}\end{aligned} \tag{5.23}$$

where the small imaginary part in the denominator guarantees causality. With (5.15) and (5.17) one obtains the relation

$$\varepsilon_{\mathrm{RPA}}(\boldsymbol{q};\omega) = 1 - \frac{4\pi e^2}{q^2}\chi(\boldsymbol{q};\omega). \tag{5.24}$$

5.1.3 Plasma resonance

First we wish to examine the dielectric function for small momenta $|\boldsymbol{q}| \ll k_F$. To this end we expand

$$\Theta(\varepsilon_F - \varepsilon_{\boldsymbol{p}+\boldsymbol{q}}) = \Theta(\varepsilon_F - \varepsilon_{\boldsymbol{p}}) - \delta(\varepsilon_{\boldsymbol{p}} - \varepsilon_F)\boldsymbol{q}\cdot\underbrace{\boldsymbol{\nabla}\varepsilon_{\boldsymbol{p}}}_{\hbar v_F\hat{\boldsymbol{p}}}$$

and obtain

$$\begin{aligned}\chi(\boldsymbol{q};\omega) &= -2\int \frac{d^3p}{(2\pi)^3} \frac{v_F\,\boldsymbol{q}\cdot\hat{\boldsymbol{p}}\,\delta(\varepsilon_{\boldsymbol{p}} - \varepsilon_F)}{v_F\,\boldsymbol{q}\cdot\hat{\boldsymbol{p}} - \omega - i0^+} \\ &\overset{\omega \gg qv_F}{\approx} 2\int \frac{d^3p}{(2\pi)^3} \frac{\boldsymbol{q}\cdot\hat{\boldsymbol{p}}\,\delta(p - k_F)}{\hbar\omega}\left(1 + \frac{v_F\boldsymbol{q}\cdot\hat{\boldsymbol{p}}}{\hbar\omega} + \cdots\right) \\ &= \frac{2}{(2\pi)^2}\int_{-1}^{1} dz \frac{k_F^2 qz}{\hbar\omega}\left(1 + \frac{v_F qz}{\omega}\right) \\ &= \frac{k_F^2 v_F q^2}{3\pi^2\hbar\omega^2} = \frac{nq^2}{m\omega^2}\end{aligned} \tag{5.25}$$

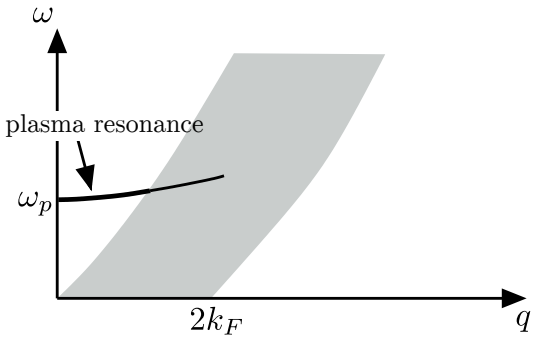


Figure 5.3: Excitation spectrum of the homogeneous electron gas in the ω-q plane. The grey shaded area corresponds to the electron-hole continuum and the solid line outside the continuum corresponds to the plasma resonance with dispersion determined by $\varepsilon(q;\omega) = 0$, which is damped (Landau damping) as soon as it enters the continuum.

where we have introduced the angle θ between $\boldsymbol{p}$ and $\boldsymbol{q}$ with $z = \cos\theta$. Thus the dielectric function $\varepsilon_{\mathrm{RPA}}(0;\omega) = 1 - (\omega_p/\omega)^2$ changes sign at $\omega_p = \sqrt{4\pi e^2 n/m}$. For $\omega < \omega_p$ ε is negative and the metal reflects all incident radiation perfectly. For $\omega > \omega_p$ the dielectric function becomes positive like in a normal dielectric and the metal becomes transparent. The frequency ω_p is called the plasma frequency and the associated quasiparticle is the plasmon, see Figure 5.3.[4] The plasmon frequency for typical metals is given by $\omega_p \simeq 10^{16}\,\mathrm{Hz}$ ($\hbar\omega_p \simeq 10eV$), where we have used $n \simeq 10^{23}\,\mathrm{cm}^{-3}$ and $m \simeq 10^{-27}\,\mathrm{g}$. Thus metals are opaque to visible light and become transparent only for X-rays.

At $\omega = \omega_p$ we have $\varepsilon = 0$. Thus there is the possibility that the system exhibits an excitation with $V, V_s, \rho \neq 0$ without an external field being applied ($U = 0$). This is the plasma oscillation, a collective excitation of the electron gas. The plasma oscillation is, however, a well-defined excitation only for small ω, q. If one computes $\chi(\boldsymbol{q},\omega)$ in general, one finds that the plasma oscillation survives as long as $\mathrm{Im}\,\chi = 0$. The imaginary part of χ leads to damping due to dissipation, see Figure 5.3. In the electron gas, electron-hole excitations $\hbar\omega = \varepsilon_{\boldsymbol{p}+\boldsymbol{q}} - \varepsilon_{\boldsymbol{p}}$ with energy ω and momentum $\boldsymbol{q}$ correspond to the denominator of 5.23. Since $\boldsymbol{p}$ is free, the electron-hole excitations

[4]Note that Planck's constant does not appear in the expression for the plasma frequency, so that this is a purely classical effect with a classical explanation (see, e.g., Feynman Lectures Vol. 2).

form a continuum without sharp energy-momentum relation. In fact, electron-hole excitations fill the region

$$0 \le \hbar\omega \le \hbar v_F q + \varepsilon_{\boldsymbol{q}}, \qquad (q \le 2k_F) \tag{5.26}$$

$$-\hbar v_F q + \varepsilon_{\boldsymbol{q}} \le \hbar\omega \le \hbar v_F q + \varepsilon_{\boldsymbol{q}}, \qquad (q \ge 2k_F)\,. \tag{5.27}$$

5.1.4 Screening

To describe screening effects we must evaluate the dielectric function at $\omega \to 0$. Setting $\omega = 0$ in (5.23), we can determine the polarization function simply,

$$\begin{aligned}
\operatorname{Re}\chi(\boldsymbol{q}) &= 2\operatorname{Re}\int \frac{d^3p}{(2\pi)^3}\frac{\Theta(\varepsilon_F - \varepsilon_{\boldsymbol{p}+\boldsymbol{q}}) - \Theta(\varepsilon_F - \varepsilon_{\boldsymbol{p}})}{\varepsilon_{\boldsymbol{p}+\boldsymbol{q}} - \varepsilon_{\boldsymbol{p}} - i0^+} \\
&= 2\operatorname{Re}\int \frac{d^3p}{(2\pi)^3}\Theta(\varepsilon_F - \varepsilon_{\boldsymbol{p}})\left(\frac{1}{\varepsilon_{\boldsymbol{p}} - \varepsilon_{\boldsymbol{p}-\boldsymbol{q}} - i0^+} - \frac{1}{\varepsilon_{\boldsymbol{p}+\boldsymbol{q}} - \varepsilon_{\boldsymbol{p}} - i0^+}\right) \\
&= -\frac{m}{\pi^2\hbar^2}\int_0^{k_F} dp\, p^2 \operatorname{Re}\int_{-1}^{1} dz \left(\frac{1}{q^2 - 2qpz + i0^+} + \frac{1}{q^2 + 2qpz - i0^+}\right) \\
&= -\frac{m}{\pi^2\hbar^2 q}\int_0^{k_F} dp\, p \log\left|\frac{q+2p}{q-2p}\right| \\
&= -\frac{m}{2\pi^2\hbar^2 q}\left(k_F^2 \log\left|\frac{2k_F+q}{2k_F-q}\right| + \int_0^{k_F} dp\,\frac{4qp^2}{4p^2-q^2}\right) \\
&= -\frac{m}{2\pi^2\hbar^2 q}\left[k_F q + (k_F^2 - \tfrac{1}{4}q^2)\log\left|\frac{2k_F+q}{2k_F-q}\right|\right],
\end{aligned} \tag{5.28}$$

where we have used the fact that $\varepsilon_{\boldsymbol{p}+\boldsymbol{q}} - \varepsilon_{\boldsymbol{p}} = \hbar^2(q^2 + 2\boldsymbol{q}\cdot\boldsymbol{p})/2m$. Inserting the result into the expression for ε yields the static dielectric function

$$\operatorname{Re}\varepsilon_{\mathrm{RPA}}(\boldsymbol{q}) = 1 + \frac{q_{\mathrm{TF}}^2}{q^2}\left(\frac{1}{2} + \frac{4k_F^2 - q^2}{8k_F q}\log\left|\frac{2k_F+q}{2k_F-q}\right|\right). \tag{5.29}$$

For $q \ll k_F$ the dielectric function agrees exactly with the Thomas-Fermi result (5.14). However, the RPA result has a divergent first derivative at $q = \pm 2k_F$. The reason for this is the sharp Fermi surface. Consider a point charge at the origin with $U(q) = 4\pi e^2/q^2$; this leads, in addition to the exponential screening of the charge, to oscillations in the charge distribution with wave vector $\simeq 2k_F$;

$$\langle\rho(\boldsymbol{r})\rangle = \int d^3q\,[\varepsilon_{\mathrm{RPA}}^{-1}(\boldsymbol{q}) - 1]e^{i\boldsymbol{q}\cdot\boldsymbol{r}} \sim \frac{\cos 2k_F r}{r^3} \tag{5.30}$$

These oscillations are called Friedel oscillations. As shown in Figure 5.4, the Friedel oscillations cause the test charge to be alternately over- and underscreened.

5.2 Superconductivity

Superconductivity was discovered in 1911 by Heike Kamerlingh Onnes. He cooled mercury with helium down to 4 K and observed that the resistance of the mercury

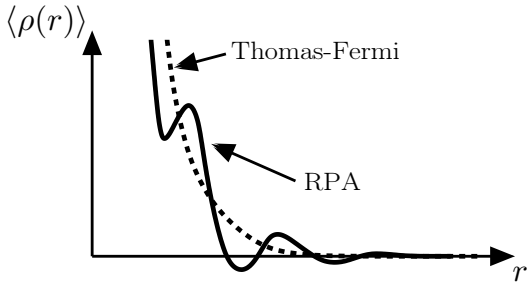


Figure 5.4: Friedel oscillations of the screening charge $\langle\rho\rangle \propto V_s$ as a function of the distance r from a point charge at the origin.

vanished abruptly at the lowest temperatures. Later experiments showed that a current in a superconductor can be maintained without an applied voltage for long times ($\gtrsim 1$ year). This confirms that the material is truly resistanceless. A further property of superconductors is their perfect diamagnetism, the Meißner-Ochsenfeld effect. In the section on the London equations we will show that the Meißner-Ochsenfeld effect is more characteristic of a superconductor than the vanishing of the resistance. The external magnetic field can penetrate only near the surface into the material; deeper layers are field-free. This "expulsion" of the magnetic field is independent of whether the sample was already superconducting before the magnetic field was switched on or whether it becomes superconducting only after the magnetic field is applied.

5.2.1 London Equations

For some time, superconductivity lacked a theoretically satisfactory explanation. In 1935, the brothers Fritz and Heinz London succeeded at least in formulating phenomenological equations. It took another 20 years until a microscopic theory was developed. The question that arises is how to describe a material that exhibits infinite conductivity. From the Newtonian equation of motion $m\partial_t \boldsymbol{v} = -e\boldsymbol{E}$ of an electron one directly obtains the first London equation for the current density $\boldsymbol{j} = -ne\boldsymbol{v}$,

$$\partial_t \boldsymbol{j} = \frac{ne^2}{m}\boldsymbol{E}. \tag{5.31}$$

This equation describes a "perfect conductor," since with $\boldsymbol{E} = 0$ it does not necessarily follow $\boldsymbol{j} = 0$, but it cannot predict the Meißner-Ochsenfeld effect. To show this, we derive from it the equation of motion for the magnetic field $\boldsymbol{H}$. For this we need the Maxwell equation $\boldsymbol{\nabla} \times \boldsymbol{H} = 4\pi\boldsymbol{j}/c$ (we consider the stationary case with $\partial_t \boldsymbol{E} = 0$). Taking the curl of this equation leads to (we use $\boldsymbol{\nabla} \cdot \boldsymbol{H} = 0$ and $\boldsymbol{\nabla} \times \boldsymbol{E} = -\partial_t \boldsymbol{H}$)

$$\nabla^2 \partial_t \boldsymbol{H} = -\overbrace{\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times}^{\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot) - \nabla^2} \partial_t \boldsymbol{H} = -\frac{4\pi}{c}\boldsymbol{\nabla} \times \partial_t \boldsymbol{j} \overset{(5.31)}{=} \frac{1}{\lambda_L^2}\partial_t \boldsymbol{H} \tag{5.32}$$

with the London penetration depth

$$\lambda_L = \sqrt{\frac{mc^2}{4\pi ne^2}}. \tag{5.33}$$

Equation (5.32) only shows that a magnetic field in a perfect conductor remains constant in time; the magnetic field lines are frozen in. This, however, is not in agreement with the experiments of Meißner and Ochsenfeld. Therefore the second London equation

$$\boldsymbol{\nabla} \times \boldsymbol{j} = \frac{ne^2}{mc}\boldsymbol{H} \tag{5.34}$$

requires the vanishing of the integration constant in (5.32). Thus (5.32) becomes

$$\nabla^2 \boldsymbol{H} = \frac{4\pi}{c}\boldsymbol{\nabla} \times \boldsymbol{j} = \frac{1}{\lambda_L^2}\boldsymbol{H} \tag{5.35}$$

and (all) magnetic fields are completely screened on the length scale λ_L; namely, if one looks for a homogeneous solution $\boldsymbol{H}(\boldsymbol{r}) = \boldsymbol{H}_0$ one immediately finds $\boldsymbol{H}_0 = 0$. The ordinary form of the Maxwell equation is recovered with $\lambda_L \to \infty$ ($n \to 0$). Physically, (5.35) means that the magnetic field (the photon) acquires a "mass" by coupling to the superconducting condensate. This is called the Anderson-Higgs mechanism and is analogous to the effect in the electroweak theory where the W and Z gauge bosons acquire mass by coupling to the Higgs boson.

Using the vector potential $\boldsymbol{A}$ in the London gauge ($\boldsymbol{\nabla} \cdot \boldsymbol{A} = 0$) one can combine the two London equations into a single equation,

$$\boldsymbol{j} = -\frac{ne^2}{mc}\boldsymbol{A} = -\frac{c}{4\pi\lambda_L^2}\boldsymbol{A}, \tag{5.36}$$

from which the first and second London equations can be derived by applying ∂_t and $\boldsymbol{\nabla} \times$, respectively. The London gauge guarantees that the continuity equation $\boldsymbol{\nabla} \cdot \boldsymbol{j} = 0$ holds. Note that the particle density n in the London equations refers only to the density of the superconducting species n_s, which in general is smaller than the electron density, and thus $\boldsymbol{j}$ denotes the corresponding supercurrent density $\boldsymbol{j}_s$, which in general differs from the total current density.

5.2.2 BCS Theory

The microscopic theory of superconductors was developed in 1957 by Bardeen, Cooper, and Schrieffer (BCS). It is based on a discovery by Leon Cooper that an arbitrarily weak attractive potential between two electrons at the Fermi level leads to a bound state of the two electrons. The pairing mechanism is important because a superfluid state is typical of bosons, which at low temperatures condense into a macroscopically occupied ground-state wavefunction. Fermions, by contrast, form a Fermi sea. For fermions to behave superfluidly (in a bosonic manner), they must first pair up and then condense.

The question remains why electrons attract each other at all, since their (Coulomb) interaction is known to be repulsive. Part of the answer was developed in the previous section: at high density the electron repulsion is not so strong because the interaction is screened on the length scale q_{TF}^{-1}. A clue to the solution came from the isotope effect: already in the early 1950s it was found for mercury that the transition temperature T_c at which superconductivity sets in depends on the isotope ($T_c \propto M_{\mathrm{Ion}}^{-1/2}$). This shows that superconductivity is not solely an electronic problem but that one must take into account the effect of phonons (lattice vibrations). A (unjustifiably) simplified calculation yields an effective electron-electron interaction $V_{\mathrm{eff}} = 4\pi e^2/q^2\varepsilon(q;\omega)$ with

$$\frac{1}{\varepsilon(q;\omega)} \approx \frac{q^2}{q^2+q_{\mathrm{TF}}^2}\left(1+\frac{\omega_q^2}{\omega^2-\omega_q^2}\right), \tag{5.37}$$

where the first term accounts for the polarization of the electrons and the second term that of the positively charged ions with dispersion $\omega_q = sq$; s is the speed of sound in the metal. Note that the time scale of the electron response is given by the short time ω_p^{-1}. We may therefore treat the interaction for the scattering processes of electrons near the Fermi surface as instantaneous, which is why we used the static expression for the electronic part of ε. The response of the (heavy) ions, by contrast, takes place on the slower time scale ω_D^{-1} (ω_D is the Debye frequency of the phonons) and is much slower, so that the ionic part must be treated dynamically. The phonon-polarization effect causes the potential for $|\omega| < \omega_q \simeq \omega_D$ to be attractive.

Although we account for the phonon effect only by an additional term in $\varepsilon(\boldsymbol{q};\omega)$, the problem is still too difficult to solve exactly. Since we want to describe superconductivity, the idea is to describe only the part of the interaction that leads to the pairing of electrons. The effect of the remaining interaction can in principle be taken into account afterwards within perturbation theory. We will see that the pairing interaction leads to an energy gap; this suppresses effects of the remaining interaction. We now consider the reduced (pairing) Hamiltonian[5]

$$H_{\mathrm{pair}} = \sum_{\boldsymbol{k}\sigma} \underbrace{(\varepsilon_{\boldsymbol{k}}-\mu)}_{\xi_{\boldsymbol{k}}} c_{\boldsymbol{k}\sigma}^\dagger c_{\boldsymbol{k}\sigma} + \sum_{\boldsymbol{k}\boldsymbol{l}} V_{\mathrm{eff}}(\boldsymbol{k}-\boldsymbol{l}) c_{\boldsymbol{k}\uparrow}^\dagger c_{-\boldsymbol{k}\downarrow}^\dagger c_{-\boldsymbol{l}\downarrow} c_{\boldsymbol{l}\uparrow}; \tag{5.38}$$

The pairing Hamiltonian describes only the scattering of Cooper pairs $|\boldsymbol{k}\uparrow,-\boldsymbol{k}\downarrow\rangle = c_{\boldsymbol{k}\uparrow}^\dagger c_{-\boldsymbol{k}\downarrow}^\dagger|0\rangle$, which leads to the pairing of the electrons.[6]

We expect that the BCS pairing Hamiltonian (5.38) leads to a phase-coherent superposition of Cooper pairs $|\boldsymbol{k}\uparrow,-\boldsymbol{k}\downarrow\rangle$. When the Cooper pairs condense, the operator

[5] The chemical potential μ is equal to $T=0$ at ε_F and is fixed by the (average) electron number.

[6] In the simplest case we consider here, a Cooper pair is a bound state of two electrons as a singlet with total momentum zero. This is called s-wave pairing. Other pairing states are conceivable. For example, the electrons can pair as a triplet (p-wave) or the electron pair can have a finite total momentum (Larkin-Ovchinnikov-Fulde-Ferrell in a magnetic field). For the vast majority of materials, however, the s-wave pairing channel with vanishing total momentum is favored.

$c_{-\boldsymbol{k}\downarrow}c_{\boldsymbol{k}\uparrow}$ acquires a finite expectation value $b_{\boldsymbol{k}}$ (coherent state). Because of the macroscopic occupation of the state, we expect fluctuations to be small. This motivates the mean-field approach

$$c_{-\boldsymbol{k}\downarrow}c_{\boldsymbol{k}\uparrow} = b_{\boldsymbol{k}} + \underbrace{(c_{-\boldsymbol{k}\downarrow}c_{\boldsymbol{k}\uparrow} - b_{\boldsymbol{k}})}_{\ll b_{\boldsymbol{k}}}; \tag{5.39}$$

By inserting into (5.38) and neglecting terms quadratic in the fluctuations, we obtain the BCS model Hamiltonian

$$H_{\mathrm{BCS}} = \sum_{\boldsymbol{k}\sigma} \xi_{\boldsymbol{k}} c^\dagger_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma} + \sum_{\boldsymbol{k}\boldsymbol{l}} V_{\mathrm{eff}}(\boldsymbol{k}-\boldsymbol{l})(c^\dagger_{\boldsymbol{k}\uparrow}c^\dagger_{-\boldsymbol{k}\downarrow}b_{\boldsymbol{l}} + b^*_{\boldsymbol{k}}c_{-\boldsymbol{l}\downarrow}c_{\boldsymbol{l}\uparrow} - b^*_{\boldsymbol{k}}b_{\boldsymbol{l}}) \tag{5.40}$$

with the self-consistency equation $b_{\boldsymbol{k}} = \langle c_{-\boldsymbol{k}\downarrow}c_{\boldsymbol{k}\uparrow}\rangle$, which can be solved exactly.

Note that, because of the mean-field approach, the Hamiltonian no longer conserves the particle number. For large particle numbers this is not serious. The chemical potential μ then simply fixes the average particle number and the relative fluctuations vanish as $N^{-1/2}$, similarly to the grand canonical ensemble.[7]

As a first step toward solving (5.40), we introduce the energy gap

$$\Delta_{\boldsymbol{k}} = -\sum_{\boldsymbol{l}} V_{\mathrm{eff}}(\boldsymbol{k}-\boldsymbol{l})b_{\boldsymbol{l}} = -\sum_{\boldsymbol{l}} V_{\mathrm{eff}}(\boldsymbol{k}-\boldsymbol{l})\langle c_{-\boldsymbol{l}\downarrow}c_{\boldsymbol{l}\uparrow}\rangle \tag{5.41}$$

with which the BCS Hamiltonian can be written as

$$H_{\mathrm{BCS}} = \sum_{\boldsymbol{k}\sigma} \xi_{\boldsymbol{k}} c^\dagger_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma} - \sum_{\boldsymbol{k}} \left(\Delta_{\boldsymbol{k}} c^\dagger_{\boldsymbol{k}\uparrow}c^\dagger_{-\boldsymbol{k}\downarrow} + \Delta^*_{\boldsymbol{k}} c_{-\boldsymbol{k}\downarrow}c_{\boldsymbol{k}\uparrow} - b^*_{\boldsymbol{k}}\Delta_{\boldsymbol{k}}\right). \tag{5.42}$$

Moreover, we wish to define new fermionic operators $\beta_{\boldsymbol{k}\sigma}$ that diagonalize the Hamiltonian. A transformation to new fermionic operators is achieved by the Bogoliubov-Valatin transformation (see Section 5.2.3)

$$c_{\boldsymbol{k}\uparrow} = u^*_{\boldsymbol{k}}\beta_{\boldsymbol{k}\uparrow} + v_{\boldsymbol{k}}\beta^\dagger_{-\boldsymbol{k}\downarrow}, \qquad c^\dagger_{-\boldsymbol{k}\downarrow} = -v^*_{\boldsymbol{k}}\beta_{\boldsymbol{k}\uparrow} + u_{\boldsymbol{k}}\beta^\dagger_{-\boldsymbol{k}\downarrow}, \tag{5.43}$$

with the (unitarity) condition $|u_{\boldsymbol{k}}|^2 + |v_{\boldsymbol{k}}|^2 = 1$. The choice[8]

$$|v_{\boldsymbol{k}}|^2 = 1 - |u_{\boldsymbol{k}}|^2 = \frac{1}{2}\left(1 - \frac{\xi_{\boldsymbol{k}}}{E_{\boldsymbol{k}}}\right),\ 2u^*_{\boldsymbol{k}}v_{\boldsymbol{k}} = \frac{\Delta_{\boldsymbol{k}}}{E_{\boldsymbol{k}}},\ E_{\boldsymbol{k}} = \sqrt{\xi^2_{\boldsymbol{k}} + |\Delta_{\boldsymbol{k}}|^2}, \tag{5.44}$$

leads, after insertion into (5.40) and simple rearrangements, to the diagonalized form (see the following section)

$$H_{\mathrm{BCS}} = \sum_{\boldsymbol{k}}(\xi_{\boldsymbol{k}} - E_{\boldsymbol{k}} + b^*_{\boldsymbol{k}}\Delta_{\boldsymbol{k}}) + \sum_{\boldsymbol{k}\sigma} E_{\boldsymbol{k}}\beta^\dagger_{\boldsymbol{k}\sigma}\beta_{\boldsymbol{k}\sigma}. \tag{5.45}$$

[7] Since one would like to assign a quantum-mechanical phase φ to the superconductor, the uncertainty in the particle number N is even indispensable, because phase and particle number are conjugate variables and therefore satisfy the Heisenberg uncertainty relation $\Delta N\,\Delta\varphi \geq 1$.

[8] The phases of $u_{\boldsymbol{k}}$, $v_{\boldsymbol{k}}$, and $\Delta_{\boldsymbol{k}}$ are related, since $\Delta^*_{\boldsymbol{k}}u_{\boldsymbol{k}}v_{\boldsymbol{k}} > 0$. Without loss of generality we may choose $u_{\boldsymbol{k}} > 0$ and then obtain that $v_{\boldsymbol{k}}$ and $\Delta_{\boldsymbol{k}}$ have the same phase.

The first term gives the ground-state energy and the second term describes the excitations as creation of Bogoliubov quasiparticles with excitation energy $E_{\boldsymbol{k}} \geq 0$. For $|\Delta_{\boldsymbol{k}}| > 0$ the quasiparticles, unlike in a metal, exhibit an energy gap.

In the following, we will first show that (5.43) and (5.44) indeed lead to the diagonal form (5.45). Next we will determine the self-consistent energy gap $\Delta_{\boldsymbol{k}}$.

5.2.3 Excursion: General quadratic Hamilton operators

Quadratic Hamilton operators that do not necessarily conserve particle number generally have the form

$$H = \sum_{i=1}^{n}\sum_{j=1}^{n}(h_{ij}c_i^\dagger c_j + \tfrac{1}{2}\Delta_{ij}c_i^\dagger c_j^\dagger + \tfrac{1}{2}\Delta_{ij}^* c_j c_i) = \boldsymbol{c}^\dagger h \boldsymbol{c} + \tfrac{1}{2}(\boldsymbol{c}^\dagger \Delta \boldsymbol{c}^\dagger + \boldsymbol{c}\Delta^\dagger \boldsymbol{c}) \tag{5.46}$$

with n the number of fermionic modes, $h = h^\dagger$ and $\Delta = -\Delta^T$.[9] One can diagonalize this Hamilton operator by means of a Bogoliubov-Valatin transformation. To this end one writes the Hamilton operator as a quadratic form

$$H = \tfrac{1}{2}\begin{pmatrix}\boldsymbol{c}^\dagger & \boldsymbol{c}\end{pmatrix}\underbrace{\begin{pmatrix} h & \Delta \\ -\Delta^* & -h^T\end{pmatrix}}_{H_{\mathrm{BdG}}}\begin{pmatrix}\boldsymbol{c} \\ \boldsymbol{c}^\dagger\end{pmatrix} + \tfrac{1}{2}\operatorname{tr} h; \tag{5.47}$$

so that every quadratic second-quantized Hamilton operator H is mapped to a Hermitian $2n \times 2n$ matrix H_{BdG}. The first-quantized Hamilton operator (matrix) H_{BdG} is called the "Bogoliubov-de Gennes Hamilton operator."

The Bogoliubov-de Gennes Hamilton operator H_{BdG} has twice as many degrees of freedom (namely $2n$) as the original fermionic system. The additional degrees of freedom arise because $\boldsymbol{c}$ appears both to the right and to the left of H_{BdG} in the quadratic form (5.47), and one therefore has the freedom to choose where to place the term $\boldsymbol{c}^\dagger h \boldsymbol{c}$ in the BdG matrix. In (5.47) we have chosen the "symmetric" variant, and hence the Bogoliubov-de Gennes Hamilton operator exhibits the ("particle-hole") symmetry $\{H_{\mathrm{BdG}}, C\} = 0$ with $C = \tau_x K$; here K denotes the operator of complex conjugation and τ_x acts on the block form in (5.47). Note that C is not a symmetry in the conventional sense, since it does not commute with the Hamilton operator. Moreover, the symmetry is always present by construction in a Bogoliubov-de Gennes Hamilton operator (and hence cannot be broken).

The task is to bring the second-quantized Hamilton operator H into the form $H = \sum_j \varepsilon_j \beta_j^\dagger \beta_j$ with fermionic operators β_j; the quadratic Hamilton operator is then solved, since all many-body eigenstates are occupation-number states with respect

[9] At first glance it appears as if equation (5.46) does not require any restriction on Δ. This is correct, but since the symmetric part $\Delta_S = (\Delta + \Delta^T)/2$ does not contribute due to the fermionic anticommutation relations, $\sum_{ij}(\Delta_S)_{ij}c_i c_j = -\sum_{ij}(\Delta_S)_{ij}c_j c_i = -\sum_{ij}(\Delta_S)_{ij}c_i c_j \Rightarrow \sum_{ij}(\Delta_S)_{ij}c_i c_j = 0$, one may choose Δ to be antisymmetric without loss of generality $\Delta = (\Delta - \Delta^T)/2$.

to $n_j = \beta_j^\dagger \beta_j$. We will see that one can additionally require that the single-particle energies $\varepsilon_j \geq 0$, so that the ground state corresponds to the vacuum (with respect to the β operators) and ε_j are precisely the (single-particle) excitation energies. This task can be accomplished by a so-called Bogoliubov transformation (from $\boldsymbol{c}$ to $\boldsymbol{\beta}$)

$$T \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\beta}^\dagger \end{pmatrix} = \begin{pmatrix} \boldsymbol{c} \\ \boldsymbol{c}^\dagger \end{pmatrix} \tag{5.48}$$

with $T \in \mathrm{Gl}(2n, \mathbb{C})$. The invariance of the quadratic Hamilton operator H under such a transformation requires that

$$H'_{\mathrm{BdG}} = T^\dagger H_{\mathrm{BdG}} T \tag{5.49}$$

where H'_{BdG} is the Bogoliubov-de Gennes Hamiltonian in the basis β. Hence we wish H'_{BdG} to be a diagonal matrix.

Note that, unlike the "ordinary" basis change from Section 4.4, here creation and annihilation operators are mixed. This is analogous to classical mechanics: in the Lagrangian formalism general basis transformations are allowed, whereas in the Hamiltonian formalism one may additionally mix position and momentum coordinates. At first glance this appears to introduce more degrees of freedom, yet only canonical transformations that leave the Poisson bracket invariant are allowed. Here the role of the Poisson bracket is taken over by the canonical anticommutation relations $\{c_i, c_j^\dagger\} = \delta_{ij}$ and $\{c_i, c_j\} = 0$. Indeed we require that the operators β_i, which diagonalize the Hamilton operator, are themselves fermionic operators so that they generate the usual fermionic Fock space.

We therefore demand invariance of[10]

$$\left\{ \begin{pmatrix} \boldsymbol{c} \\ \boldsymbol{c}^\dagger \end{pmatrix}, \begin{pmatrix} \boldsymbol{c} \\ \boldsymbol{c}^\dagger \end{pmatrix}^\dagger \right\} = I_{2n} \tag{5.50}$$

under the transformation (5.49). Substituting the transformation (5.49) and using the bilinearity of the anticommutator, while demanding that the anticommutator remains invariant, we obtain the constraint

$$I_{2n} = \left\{ T \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\beta}^\dagger \end{pmatrix}, \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\beta}^\dagger \end{pmatrix}^\dagger T^\dagger \right\} = T \underbrace{\left\{ \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\beta}^\dagger \end{pmatrix}, \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\beta}^\dagger \end{pmatrix}^\dagger \right\}}_{I_{2n}} T^\dagger = T T^\dagger; \tag{5.51}$$

i.e., T is unitary.

Diagonalizing a Hermitian matrix H_{BdG} by a unitary transformation T is achieved via the standard eigenvalue problem

$$H_{\mathrm{BdG}} \boldsymbol{t}_j = \varepsilon_j \boldsymbol{t}_j \tag{5.52}$$

[10] We regard the anticommutator of two vectors $\boldsymbol{v}$ and $\boldsymbol{v}$ as a matrix with matrix elements $(\{\boldsymbol{v}, \boldsymbol{w}\})_{ij} = \{v_i, w_j\}$. The anticommutator in (5.50) is a quadratic form that plays the role of a metric on the space of creation and annihilation operators.

where $\boldsymbol{t}_j$ are the columns of T. The particle-hole symmetry C guarantees that for every eigenvector $\boldsymbol{t}_+$ with eigenvalue $\varepsilon \geq 0$ there exists an additional eigenvector $\boldsymbol{t}_- = C\boldsymbol{t}_+$ with eigenvalue $-\varepsilon \leq 0$; namely

$$\underbrace{H_{\mathrm{BdG}}}_{-C^{-1}H_{\mathrm{BdG}}C}\boldsymbol{t}_+ = \varepsilon \boldsymbol{t}_+ \qquad \Rightarrow \qquad H_{\mathrm{BdG}}\boldsymbol{t}_- = -\varepsilon \boldsymbol{t}_-. \tag{5.53}$$

Ordering the eigenvalues so that positive eigenvalues appear first, followed by negative eigenvalues on the diagonal of H'_{BdG}, the particle-hole symmetry ensures that $\boldsymbol{\beta}^\dagger$ in the transformation (5.49) are precisely the adjoint operators of $\boldsymbol{\beta}$. In summary we obtain

$$\begin{aligned} H &= \tfrac{1}{2}\begin{pmatrix}\boldsymbol{\beta}^\dagger & \boldsymbol{\beta}\end{pmatrix}\begin{pmatrix} \varepsilon_1 & & & & & \\ & \ddots & & & & \\ & & \varepsilon_n & & & \\ & & & -\varepsilon_1 & & \\ & & & & \ddots & \\ & & & & & -\varepsilon_n \end{pmatrix}\begin{pmatrix}\boldsymbol{\beta} \\ \boldsymbol{\beta}^\dagger\end{pmatrix} + \tfrac{1}{2}\operatorname{tr} h \\ &= \sum_{j=1}^{n} \varepsilon_j \beta_j^\dagger \beta_j + \tfrac{1}{2}\Big(\operatorname{tr} h - \sum_{j=1}^{n}\varepsilon_j\Big). \end{aligned} \tag{5.54}$$

Consider as an example the model Hamiltonian (5.42); it can be written in Bogoliubov-de Gennes form

$$H = \sum_{\boldsymbol{k}}\left[\begin{pmatrix} c_{\boldsymbol{k}\uparrow}^\dagger & c_{-\boldsymbol{k}\downarrow}\end{pmatrix} H_{\boldsymbol{k}} \begin{pmatrix} c_{\boldsymbol{k}\uparrow} \\ c_{-\boldsymbol{k}\downarrow}^\dagger\end{pmatrix} + \xi_{\boldsymbol{k}} + b_{\boldsymbol{k}}^* \Delta_{\boldsymbol{k}}\right] \tag{5.55}$$

with

$$H_{\boldsymbol{k}} = \begin{pmatrix} \xi_{\boldsymbol{k}} & -\Delta_{\boldsymbol{k}} \\ -\Delta_{\boldsymbol{k}}^* & -\xi_{\boldsymbol{k}}\end{pmatrix}. \tag{5.56}$$

Due to translational symmetry the problem reduces to diagonalizing the 2×2 Bogoliubov-de Gennes Hamiltonian $H_{\boldsymbol{k}}$. The two eigenstates of the Hamiltonian are given by

$$H_{\boldsymbol{k}}\begin{pmatrix} u_{\boldsymbol{k}}^* \\ -v_{\boldsymbol{k}}^*\end{pmatrix} = E_{\boldsymbol{k}}\begin{pmatrix} u_{\boldsymbol{k}}^* \\ -v_{\boldsymbol{k}}^*\end{pmatrix}, \qquad H_{\boldsymbol{k}}\begin{pmatrix} v_{\boldsymbol{k}} \\ u_{\boldsymbol{k}}\end{pmatrix} = -E_{\boldsymbol{k}}\begin{pmatrix} v_{\boldsymbol{k}} \\ u_{\boldsymbol{k}}\end{pmatrix} \tag{5.57}$$

with $u_{\boldsymbol{k}}$ and $v_{\boldsymbol{k}}$ as in (5.44).[11] The transformation to the basis $\beta_{\boldsymbol{k}\uparrow}, \beta_{-\boldsymbol{k}\downarrow}^\dagger$ yields the

[11] We have thereby $T_{\boldsymbol{k}} = \begin{pmatrix} u_{\boldsymbol{k}}^* & v_{\boldsymbol{k}} \\ -v_{\boldsymbol{k}}^* & u_{\boldsymbol{k}}\end{pmatrix}$. The unitarity of T is guaranteed by (5.44).

result (5.45),

$$\begin{aligned} H &= \sum_{\boldsymbol{k}} [E_{\boldsymbol{k}} (\beta^\dagger_{\boldsymbol{k}\uparrow} \beta_{\boldsymbol{k}\uparrow} - \beta_{-\boldsymbol{k}\downarrow} \beta^\dagger_{-\boldsymbol{k}\downarrow}) + \xi_{\boldsymbol{k}} + b^*_{\boldsymbol{k}} \Delta_{\boldsymbol{k}}] \\ &= \underbrace{\sum_{\boldsymbol{k}} (\xi_{\boldsymbol{k}} - E_{\boldsymbol{k}} + b^*_{\boldsymbol{k}} \Delta_{\boldsymbol{k}})}_{E_g} + \sum_{\boldsymbol{k}\sigma} E_{\boldsymbol{k}} \beta^\dagger_{\boldsymbol{k}\sigma} \beta_{\boldsymbol{k}\sigma}, \end{aligned} \tag{5.58}$$

with ground-state energy E_g and excitation energies $E_{\boldsymbol{k}}$.

5.2.4 Self-consistency equation

Next we must solve the self-consistency equation (5.41) to see whether for a superconductor $|\Delta_{\boldsymbol{k}}| > 0$ actually arises. The trivial solution of the self-consistency equation with $\Delta_{\boldsymbol{k}} = 0$, so that $v_{\boldsymbol{k}} = 1$ for $\xi_{\boldsymbol{k}} < 0$ and $v_{\boldsymbol{k}} = 0$ for $\xi_{\boldsymbol{k}} > 0$, describes the Fermi sea in the normal state. The self-consistency equation reads

$$\begin{aligned} \Delta_{\boldsymbol{k}} &= -\sum_{\boldsymbol{l}} V_{\text{eff}}(\boldsymbol{k} - \boldsymbol{l}) \langle c_{-\boldsymbol{l}\downarrow} c_{\boldsymbol{l}\uparrow} \rangle \overset{(5.43)}{=} -\sum_{\boldsymbol{l}} V_{\text{eff}}(\boldsymbol{k} - \boldsymbol{l}) u^*_{\boldsymbol{l}} v_{\boldsymbol{l}} \\ &\overset{(5.44)}{=} -\frac{1}{2} \sum_{\boldsymbol{l}} V_{\text{eff}}(\boldsymbol{k} - \boldsymbol{l}) \frac{\Delta_{\boldsymbol{l}}}{E_{\boldsymbol{l}}}, \end{aligned} \tag{5.59}$$

where for the evaluation we have used that the expectation value of the pairing term in the ground state $\langle c_{-\boldsymbol{l}\downarrow} c_{\boldsymbol{l}\uparrow} \rangle$, namely that the ground state $|0\rangle$ is the vacuum for the operators $\beta_{\boldsymbol{k}}$ with $\beta_{\boldsymbol{k}}|0\rangle = 0$.

To describe superconductivity we need an attractive interaction. As we have already seen, the effective electron-electron potential V_{eff} becomes attractive for energy transfers with $|\xi_{\boldsymbol{k}} - \xi_{\boldsymbol{l}}| < \hbar\omega_D$. In order to solve the self-consistency equation we make the stronger assumption that V_{eff} is attractive whenever $|\xi_{\boldsymbol{k}}|$ and $|\xi_{\boldsymbol{l}}|$ are both smaller than $\hbar\omega_D$. Moreover we replace the interaction in this range by a constant, i.e., $V_{\text{eff}}(\boldsymbol{k} - \boldsymbol{l}) = -g\,\Theta(\hbar\omega_D - |\xi_{\boldsymbol{k}}|)\,\Theta(\hbar\omega_D - |\xi_{\boldsymbol{l}}|)/\Omega$, $g > 0$.[12] It is easy to see that for this model the self-consistency equation is solved by $\Delta_{\boldsymbol{k}} = \Delta\,\Theta(\hbar\omega_D - |\xi_{\boldsymbol{k}}|)$. Inserting the ansätze for $\Delta_{\boldsymbol{k}}$ and V_{eff} into the self-consistency equation and canceling the common factor $\Delta \neq 0$, we obtain[13]

$$\begin{aligned} 1 &= \frac{g}{2\Omega} \sum_{\substack{\boldsymbol{k} \\ |\xi_{\boldsymbol{k}}| < \hbar\omega_D}} \frac{1}{E_{\boldsymbol{k}}} = \frac{g}{2} \overbrace{\int_{|\xi| < \hbar\omega_D} d\xi \frac{\rho(\xi)}{\sqrt{|\Delta|^2 + \xi^2}}}^{\approx \rho_0 \int_{|\xi| < \hbar\omega_D} d\xi (\xi^2 + |\Delta|^2)^{-1/2}} \\ &\approx g\rho_0 \operatorname{arsinh}(\hbar\omega_D / |\Delta|); \end{aligned} \tag{5.60}$$

[12] These assumptions are based on the work of BCS; we expect that the solution of the self-consistency equation does not depend on the concrete form of the potential but only on the typical strength $-g$ and the cutoff energy $\hbar\omega_D$.

[13] The density of states is defined as $\rho(\xi) = \sum_{\boldsymbol{k}} \delta(\xi - \xi_{\boldsymbol{k}})$. For a general function $f(x)$ we therefore have $\sum_{\boldsymbol{k}} f(\xi_{\boldsymbol{k}}) = \int d\xi\, \rho(\xi) f(\xi)$.

here $\rho_0 = \rho(0)$ denotes the density of states at the Fermi level, which we have assumed constant over the small integration range $\simeq \omega_D \ll \varepsilon_F$. Solving for $|\Delta|$ yields the final result

$$|\Delta| = \frac{\hbar\omega_D}{\sinh(1/\rho_0 g)} \approx 2\hbar\omega_D e^{-1/\rho_0 g} \tag{5.61}$$

where in the last step we have used that the system is typically weakly coupled with $\rho_0 g \simeq 0.3$. Equation (5.61) confirms Cooper's result that any arbitrarily weak attractive potential pairs electrons and thereby produces a finite energy gap.

5.2.5 Condensation energy

The transition into the superconducting phase is accompanied by a lowering of the ground-state energy, which in analogy to the gas-liquid transition is called the condensation energy. Using the expression for E_g from (5.58) and $\sum_{\boldsymbol{k}} b_{\boldsymbol{k}} = |\Delta|\Omega/g$ one obtains the condensation energy density

$$\begin{aligned}
u_{\text{cond}} &= \frac{1}{\Omega}[E_g(\Delta) - E_g(\Delta = 0)] \qquad\qquad \rho_0 \int_0^{\hbar\omega_D} d\xi \, \frac{|\Delta|^2}{\sqrt{|\Delta|^2+\xi^2}} \\
&= \frac{1}{\Omega} \sum_{|\boldsymbol{k}|\geq k_F} (\xi_{\boldsymbol{k}} - E_{\boldsymbol{k}}) + \frac{1}{\Omega} \sum_{|\boldsymbol{k}|\leq k_F} (\overbrace{-\xi_{\boldsymbol{k}}}^{|\xi_{\boldsymbol{k}}|} - E_{\boldsymbol{k}}) + \overbrace{\frac{|\Delta|^2}{g}} \\
&\approx 2\rho_0 \int_0^{\hbar\omega_D} d\xi \left(\xi - \sqrt{|\Delta|^2+\xi^2} + \frac{|\Delta|^2}{2\sqrt{|\Delta|^2+\xi^2}} \right) \\
&= \rho_0 \int_0^{\hbar\omega_D} d\xi \, \frac{2\xi\sqrt{|\Delta|^2+\xi^2} - 2\xi^2 - 1}{\sqrt{|\Delta|^2+\xi^2}} \\
&= \rho_0 \xi \left(\xi - \sqrt{|\Delta|^2+\xi^2} \right) \Big|_{\xi=0}^{\hbar\omega_D} \\
&\overset{\hbar\omega_D \gg |\Delta|}{\approx} -\tfrac{1}{2}\rho_0 |\Delta|^2 .
\end{aligned} \tag{5.62}$$

5.2.6 Density of states

We have seen that the fermionic quasiparticles $\beta_{\boldsymbol{k}}^\dagger$, which describe the excitations above the superconducting ground state, stand in a one-to-one correspondence with the electron excitations $c_{\boldsymbol{k}}^\dagger$ in the metal. This insight allows us to relate the density of states of the Bogoliubov quasiparticles $\rho_s(E)$ to the density of states ρ_0 in the normal state, $\rho_s(E) = \sum_{\boldsymbol{k}} \delta(E - E_{\boldsymbol{k}}) = (d\xi/dE) \sum_{\boldsymbol{k}} \delta(\xi - \xi_{\boldsymbol{k}}) = (d\xi/dE)\rho_0$. This leads to the simple result (the quasiparticle excitations obey the spectrum $E = \sqrt{|\Delta|^2+\xi^2}$)

$$\frac{\rho_s(E)}{\rho_0} = \frac{d\xi}{dE} = \begin{cases} \frac{E}{\sqrt{E^2-|\Delta|^2}} & E > |\Delta|, \\ 0 & E < |\Delta|, \end{cases} \tag{5.63}$$

with a divergent density of states above the energy gap. Note that, because of the aforementioned relation between excitations in the superconducting and normal state,

the integrated density of states is the same. Upon entering the superconducting state, states with energy below the gap are simply transferred above the gap, which causes the divergence.

5.2.7 Derivation of the London equations

With the BCS equation and its solution, we have seen that an attractive electron-electron potential leads to an energy gap and thereby lowers the total energy of the system by the condensation energy. So far, however, it is not clear that the new state described by $|\Delta| > 0$ is indeed a superconducting state. We have already seen that the London equations provide a phenomenological description (Meißner-Ochsenfeld effect, perfect conductivity) of a superconductor. We will therefore now show that the response of the BCS quasiparticles to an external magnetic field leads precisely to the London equation, and hence that the state is indeed a superconductor.

We describe the constant magnetic field $\boldsymbol{H}$ by a vector potential $\boldsymbol{A}$ with $\boldsymbol{H} = \boldsymbol{\nabla} \times \boldsymbol{A}$. In doing so, we choose the London gauge $\boldsymbol{\nabla} \cdot \boldsymbol{A} = 0$. The magnetic field shifts the canonical momentum so that the kinetic term takes the form $(\boldsymbol{p} + e\boldsymbol{A}/c)^2/2m$; in this step we neglect effects arising from the Zeeman term. To first order in $\boldsymbol{A}$ this leads to the perturbation term[14]

$$\begin{aligned} H_A &= \frac{e}{mc} \int d^3r \sum_\sigma \Psi_\sigma^\dagger(\boldsymbol{r}) \boldsymbol{A} \cdot \boldsymbol{p} \Psi_\sigma(\boldsymbol{r}) = \frac{e\hbar}{mc} \sum_{\boldsymbol{k}\sigma} \boldsymbol{A} \cdot \boldsymbol{k}\, c_{\boldsymbol{k}\sigma}^\dagger c_{\boldsymbol{k}\sigma} \\ &= \frac{e\hbar}{mc} \sum_{\boldsymbol{k}\sigma} \boldsymbol{A} \cdot \boldsymbol{k}\, \beta_{\boldsymbol{k}\sigma}^\dagger \beta_{\boldsymbol{k}\sigma}. \end{aligned} \tag{5.64}$$

In the next step we wish to compute the current density $\boldsymbol{j} = -ne\boldsymbol{v} = \langle \boldsymbol{j}_p \rangle + \langle \boldsymbol{j}_A \rangle$. The two terms $\boldsymbol{j}_p$ and $\boldsymbol{j}_A$ denote the paramagnetic and diamagnetic current-density operators, respectively, corresponding to the expression $\boldsymbol{v} = \boldsymbol{p}/m + e\boldsymbol{A}/mc$ for the particle velocity. The diamagnetic current yields the simple result

$$\langle \boldsymbol{j}_A \rangle = -\frac{e^2 \boldsymbol{A}}{mc} \frac{1}{\Omega} \underbrace{\left\langle \sum_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma}^\dagger c_{\boldsymbol{k}\sigma} \right\rangle}_{N} = -\frac{ne^2}{mc} \boldsymbol{A}, \tag{5.65}$$

which is exactly the London equation provided we interpret n as the superfluid particle density n_s. However, in equation (5.65) n is always the total particle density.

To obtain the total current $\boldsymbol{j}$ we must still determine the paramagnetic contribution to the current density $\langle \boldsymbol{j}_p \rangle$ with the current operator

$$\boldsymbol{j}_p = -\frac{e\hbar}{m} \frac{1}{\Omega} \sum_{\boldsymbol{k}\sigma} \boldsymbol{k} c_{\boldsymbol{k}\sigma}^\dagger c_{\boldsymbol{k}\sigma} = -\frac{e\hbar}{m} \frac{1}{\Omega} \sum_{\boldsymbol{k}\sigma} \boldsymbol{k}\, \beta_{\boldsymbol{k}\sigma}^\dagger \beta_{\boldsymbol{k}\sigma}. \tag{5.66}$$

[14] We use that $c_{\boldsymbol{k}\uparrow}^\dagger c_{\boldsymbol{k}\uparrow} - c_{-\boldsymbol{k}\downarrow}^\dagger c_{-\boldsymbol{k}\downarrow} = (u_{\boldsymbol{k}} \beta_{\boldsymbol{k}\uparrow}^\dagger + v_{\boldsymbol{k}}^* \beta_{-\boldsymbol{k}\downarrow})(u_{\boldsymbol{k}}^* \beta_{\boldsymbol{k}\uparrow} + v_{\boldsymbol{k}} \beta_{-\boldsymbol{k}\downarrow}^\dagger) - (-v_{\boldsymbol{k}}^* \beta_{\boldsymbol{k}\uparrow} + u_{\boldsymbol{k}} \beta_{-\boldsymbol{k}\downarrow}^\dagger)(-v_{\boldsymbol{k}} \beta_{\boldsymbol{k}\uparrow}^\dagger + u_{\boldsymbol{k}}^* \beta_{-\boldsymbol{k}\downarrow}) = \beta_{\boldsymbol{k}\uparrow}^\dagger \beta_{\boldsymbol{k}\uparrow} - \beta_{-\boldsymbol{k}\downarrow}^\dagger \beta_{-\boldsymbol{k}\downarrow}$.

The perturbation term H_A shifts the energy of a particle with momentum $\boldsymbol{k}$ by $e\hbar\boldsymbol{A}\cdot\boldsymbol{k}/mc$. At $T=0$ all states of negative energy are occupied, and we obtain

$$\begin{aligned}\langle\beta^\dagger_{\boldsymbol{k}\uparrow}\beta_{\boldsymbol{k}\uparrow}\rangle - \langle\beta^\dagger_{-\boldsymbol{k}\downarrow}\beta_{-\boldsymbol{k}\downarrow}\rangle &= \Theta(-E_{\boldsymbol{k}} - e\hbar\boldsymbol{A}\cdot\boldsymbol{k}/mc) - \Theta(-E_{\boldsymbol{k}} + e\hbar\boldsymbol{A}\cdot\boldsymbol{k}/mc)\\ &= -\frac{2e\hbar\boldsymbol{A}\cdot\boldsymbol{k}}{mc}\delta(E_{\boldsymbol{k}})\end{aligned} \tag{5.67}$$

to first order in $\boldsymbol{A}$. Inserting this result into equation (5.64) gives

$$\begin{aligned}\langle\boldsymbol{j}_p\rangle &= \frac{2e^2\hbar^2}{m^2c}\overbrace{\int\frac{d^3k}{(2\pi)^3}\delta(E_{\boldsymbol{k}})}^{\rho_s(0)\langle\cdot\rangle_{S^2}}\overbrace{(\boldsymbol{A}\cdot\boldsymbol{k})}^{\propto\cos\theta}\overbrace{\boldsymbol{k}}^{\propto\cos\theta}\\ &= \frac{2e^2\hbar^2}{m^2c}\rho_s(0)k_F^2\underbrace{\langle\cos^2\theta\rangle_{S^2}}_{1/3}\boldsymbol{A} = \frac{ne^2}{mc}\underbrace{\frac{4\varepsilon_F\rho_s(0)}{3n}}_{\rho_s(0)/\rho_0}\boldsymbol{A}\end{aligned} \tag{5.68}$$

where θ is the angle between $\boldsymbol{k}$ and $\boldsymbol{A}$ and $\langle\cdot\rangle_{S^2}$ denotes the average over the unit sphere. The sum of the paramagnetic and diamagnetic terms yields the final result

$$\boldsymbol{j} = -\frac{n_s e^2}{mc}\boldsymbol{A} \tag{5.69}$$

with the superfluid density

$$n_s = n[1-\rho_s(0)/\rho_0] = \begin{cases} n & |\Delta| > 0,\\ 0 & \Delta = 0.\end{cases} \tag{5.70}$$

At finite temperatures the tail of the Fermi-Dirac distribution causes $n_s = n(1 - e^{-|\Delta|/k_BT})$ as long as $|\Delta| \gg k_BT$. The phase transition occurs precisely at the point where $n_s = 0$.

Note that in the normal state there is no Meißner-Ochsenfeld effect at all, since the paramagnetic and diamagnetic currents cancel exactly. The origin of the paramagnetic response in the normal conductor lies in the fact that a vector potential $\boldsymbol{A}$ shifts the Fermi sphere in momentum space, which leads to a non-vanishing $\langle\boldsymbol{p}\rangle$. For the superconductor, the particles are condensed in the state with $\langle\boldsymbol{p}\rangle = 0$. Moreover, at the Fermi surface there are no quasiparticles present that could respond to an applied magnetic field, so that the pure diamagnetic response—which is always present—remains.

Chapter 6

Relativistic Particles: Dirac Equation

Quantum mechanics, as we have treated it so far, is based on non-relativistic mechanics and can be motivated by the correspondence principle. With this, we were able to successfully explain an important part of quantum physics. Interesting and fundamentally new perspectives open up when we incorporate special relativity into our formulation of quantum mechanics. The relativistic extension of quantum mechanics is the foundation of high-energy physics.

6.1 Klein-Gordon Equation

Non-relativistic mechanics and quantum mechanics are invariant under Galilei transformations between inertial systems,

$$\boldsymbol{x} \mapsto \boldsymbol{x}' = \boldsymbol{x} - \boldsymbol{v}t \qquad \text{and} \qquad t \mapsto t' = t, \tag{6.1}$$

where $\boldsymbol{v}$ is the relative velocity. Energy and momentum are related by the relationship $E = p^2/2m$, so that the Galilei transformation leads to

$$\boldsymbol{p} \mapsto \boldsymbol{p}' = \boldsymbol{p} - m\boldsymbol{v} \qquad \text{and} \qquad E \mapsto E' = E - \boldsymbol{p}\cdot\boldsymbol{v} + \tfrac{1}{2}mv^2 = \frac{p'^2}{2m} \tag{6.2}$$

i.e., the energy is form-invariant under Galilei transformations. The correspondence principle states that

$$E \mapsto i\hbar\partial_t \qquad \text{and} \qquad \boldsymbol{p} = -i\hbar\boldsymbol{\nabla}. \tag{6.3}$$

The Galilei-invariant relationship between momentum and energy yields the Schrödinger equation

$$i\hbar\partial_t\psi = -\frac{\hbar^2}{2m}\nabla^2\psi, \tag{6.4}$$

which has been our fundamental equation up to now.

Relativistic mechanics is no longer Galilei- but Lorentz-invariant. A Lorentz transformation from $(t, \boldsymbol{x})$ to $(t', \boldsymbol{x}')$ leaves the length[1] $x_\mu x^\mu = c^2t^2 - x^2 = x'_\mu x'^\mu$ of the four-vector $x^\mu = (ct, \boldsymbol{x})$ invariant. The covariant vector x_μ is obtained from the covariant vector x^μ and the Lorentz metric

$$(g_{\mu\nu}) = (g^{\mu\nu}) = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}. \tag{6.5}$$

Energy and momentum are also combined into a four-vector in special relativity $p^\mu = (E/c, \boldsymbol{p})$, with the invariant length

$$p_\mu p^\mu = p'_\mu p'^\mu = m^2c^2. \tag{6.6}$$

This is the invariant energy-momentum relation, which we can use to derive relativistic quantum mechanics with the correspondence principle. First, we replace p^μ by a differential operator

$$p_\mu = i\hbar\partial_\mu. \tag{6.7}$$

The first attempt yields the Hamiltonian $H = \sqrt{m^2c^4 + p^2c^2}$, with the wave equation

$$i\hbar\partial_t\psi = mc^2(1 - \lambda_C^2\nabla)^{1/2}\psi \tag{6.8}$$

with the Compton wavelength $\lambda_C = \hbar/mc = \alpha a_B \approx 3.8 \times 10^{-11}\,\mathrm{cm}$. This approach, however, raises serious problems. First, space and time are treated differently, and both do not appear manifestly as a covariant four-vector. A further problem is the square root, since in the derivation $\boldsymbol{\nabla}$ we must expand to infinite order. This means that the equation describes a non-local theory. Thus, causality is violated, since the right-hand side of the equation instantaneously connects the point $\boldsymbol{x}$ with points outside the causal light cone through the derivative. For these reasons, the above approach must be discarded. The problem of covariance and nonlocality is solved at once if we start from the square of the energy, i.e., $H^2 = m^2c^4 + p^2c^2$ with the Klein-Gordon equation

$$-\hbar\partial_t^2\psi = (-\hbar^2c^2\nabla^2 + m^2c^4)\psi \quad \mapsto \quad (\partial_\mu\partial^\mu + \lambda_C^{-2})\psi = 0. \tag{6.9}$$

A more detailed analysis of the Klein-Gordon equation shows that although causality is guaranteed, the equation still has "problems" because the Hamiltonian is not positive definite and there is no positive definite conserved quantity (particle density). However, this is not really a problem, since one can interpret this situation as the density corresponding to a charge density $e\rho$, and therefore particles of positive and negative energy have different charges. Thus the continuity equation holds for the

[1] In this chapter we use the Einstein summation convention, which states that summation is performed over repeated indices.

charge density, but not for the particle density. The Klein-Gordon equation introduces a new aspect into quantum mechanics, which first manifests itself in the breakdown of the single-particle theory. In quantum-mechanical processes, the particle number can change, while the total charge is conserved. This foreshadows the concept that for every particle there exists a corresponding antiparticle, which has the charge and energy with opposite sign. In fact, the Klein-Gordon equation is suitable for describing bosonic particles, such as the π mesons, in which particles and antiparticles are equivalent.

6.2 Dirac Equation

The simplest relativistic equation we have constructed so far, the Klein-Gordon equation, has the following properties:

- scalar,
- covariant,
- 2^{ter} order in time,
- and has a non-positive-definite norm, so that the Hamiltonian H is not Hermitian.

In the following, we raise our ambitions and attempt to construct a theory based on a Hamiltonian with the following properties: the Hamiltonian

- may be tensorial,[2]
- should be covariant,
- should be 1^{te} order in time,
- should lead to a positive-definite norm so that we obtain a Hermitian Hamiltonian H.

The most general ansatz for a Hamiltonian with these properties (linear in ∂_t and covariant implies linearity in space, i.e., in $\boldsymbol{\nabla}$) has the form

$$i\hbar\partial_t\Psi \;=\; H\Psi = \left(c\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta mc^2\right)\Psi, \tag{6.10}$$

with Ψ a (Dirac) spinor with n components, $\boldsymbol{\alpha}$, β are $n\times n$ matrices, Hermitian and dimensionless; the momentum is defined as usual, $\boldsymbol{p} = -i\hbar\boldsymbol{\nabla}/\hbar$. Our task is to find n, $\boldsymbol{\alpha}$, and β such that (6.10) reproduces the usual relativistic dispersion

$$E^2 = p^2c^2 + m^2c^4 \tag{6.11}$$

[2]The restriction to a scalar theory does not yield a solution.

. To this end, we square the equation according to

$$\begin{aligned} E^2\Psi \mapsto -\hbar^2\partial_t^2\Psi &= i\hbar\partial_t\left(c\boldsymbol{\alpha}\cdot\boldsymbol{p}+\beta mc^2\right)\Psi \\ &= \left(c\boldsymbol{\alpha}\cdot\boldsymbol{p}+\beta mc^2\right) i\hbar\partial_t\Psi \\ &= \left(c\boldsymbol{\alpha}\cdot\boldsymbol{p}+\beta mc^2\right)^2\Psi \\ &\stackrel{!}{=} \left(p^2c^2+m^2c^4\right)\Psi. \end{aligned} \tag{6.12}$$

Hence, it must hold that

$$\begin{aligned} \left(c\alpha_i p^i+\beta mc^2\right)\left(c\alpha_j p^j+\beta mc^2\right) &= \underbrace{c^2\alpha_i^2(p^i)^2}_{c^2p^2} \\ &+ \underbrace{\beta^2m^2c^4}_{m^2c^4} + \underbrace{\left(\alpha_i\beta+\beta\alpha_i\right)p^imc^3}_{0} + \underbrace{\frac{1}{2}\left(\alpha_i\alpha_j+\alpha_j\alpha_i\right)_{i\neq j}p^ip^jc^2}_{0} \end{aligned} \tag{6.13}$$

Thus, we obtain the conditions

$$\alpha_i^2 = I_n, \qquad \beta^2 = I_n, \tag{6.14}$$

$$\{\alpha_i,\beta\} = 0, \qquad \{\alpha_i,\alpha_j\}_{i\neq j} = 0, \tag{6.15}$$

where $\{\cdot,\cdot\}$ denotes the anticommutator as usual. We therefore need four Hermitian, anticommuting matrices whose squares yield I_n. It turns out that these conditions can only be satisfied if $n \geq 4$. To prove this, we note that

1. The anticommutator $\{\alpha_i,\beta\} = 0$ implies that $\beta\alpha_i\beta = -\alpha_i$ and hence $\operatorname{tr}\alpha_i = -\operatorname{tr}[\beta\alpha_i\beta] = -\operatorname{tr}[\beta^2\alpha_i] = -\operatorname{tr}\alpha_i$ and hence $\operatorname{tr}\alpha_i = 0$. Likewise, from $\alpha_i\beta\alpha_i = -\beta$ it follows that $\operatorname{tr}\beta = 0$, so that all four matrices α_i,β are traceless.
2. The four matrices α_i and β have eigenvalues ±1, since α_i,β are Hermitian and thus also diagonalizable; since $\alpha_i^2 = \beta^2 = I_n$, the squares of the eigenvalues must be equal to 1.
3. n must be even, $n = 2m$, with $m \in \mathbb{N}$, since $\operatorname{tr}\alpha_i = 0, \operatorname{tr}\beta = 0$ the eigenvalues 1 and -1 must occur pairwise.
4. Finally, $n > 2$ must hold, since a Hermitian matrix has a real main diagonal and complex conjugate elements symmetric about the main diagonal, so that we find the following number of degrees of freedom

$$\underbrace{n}_{\text{reelle HD}} + \underbrace{2}_{\text{komplex}} \cdot \underbrace{(n^2-n)}_{\text{nichtdiag.}} \underbrace{/\ 2}_{\text{symm.}} = n^2; \tag{6.16}$$

additionally $\operatorname{tr} = 0$ and there remain n^2-1 degrees of freedom. For $n = 2$ we can find three independent Hermitian matrices, for example the Pauli matrices σ_i,

$$\left[\begin{pmatrix}0&1\\1&0\end{pmatrix},\begin{pmatrix}0&-i\\i&0\end{pmatrix},\begin{pmatrix}1&0\\0&-1\end{pmatrix}\right] = [\sigma_1,\sigma_2,\sigma_3] \ = \ \boldsymbol{\sigma}. \tag{6.17}$$

with $\sigma_i^2 = I_2$ and the commutation rules $\sigma_i\sigma_j + \sigma_j\sigma_i = 0$.

For $n = 4$ we find 15 independent matrices with vanishing trace; of these we choose the matrices

$$\begin{aligned}
\beta &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}; \\
\alpha_1 &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{pmatrix}; \\
\alpha_2 &= \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \sigma_2 \\ \sigma_2 & 0 \end{pmatrix}; \\
\alpha_3 &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \sigma_3 \\ \sigma_3 & 0 \end{pmatrix};
\end{aligned} \tag{6.18}$$

with the required properties $\beta^2 = \alpha_i^2 = I_4$ and $\{\alpha_i, \alpha_j\} = 0$, $\{\alpha_i, \beta\} = 0$.

Thus, we obtain the Dirac equation in the form

$$i\hbar\partial_t\Psi = \left(-i\hbar c\alpha_i\partial_{x_i} + \beta mc^2\right)\Psi, \tag{6.19}$$

with the Dirac 4$^{\text{er}}$ spinor

$$\Psi = \begin{pmatrix} \Psi_i \\ \vdots \\ \Psi_4 \end{pmatrix} = \begin{pmatrix} \chi \\ \eta \end{pmatrix}, \tag{6.20}$$

and the 2$^{\text{er}}$ spinors χ and η.

The covariant form of the Dirac equation is obtained by multiplying (6.19) with β and defining the Dirac matrices

$$\begin{aligned}
\gamma^\mu &= \left(\beta, \beta\alpha^i\right), \\
\gamma^0 &= \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}, \\
\gamma^i &= \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix},
\end{aligned} \tag{6.21}$$

; the Dirac matrices are characterized by the following anticommutators [3],

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} I_n, \tag{6.22}$$

with $g^{\mu\nu}$ the metric tensor. From (6.19) one then obtains (again we set $\hbar = c = 1$ for covariant formulas; it is $\partial_\mu = \partial_{x^\mu}$)

$$\left(i\gamma^\mu \partial_\mu - m\right)\, \Psi(x) = 0. \tag{6.23}$$

We use the Einstein summation convention and introduce the notation of the "slash" operator, $\not{a} = \gamma^\mu a_\mu$; this further simplifies (6.23) to

$$\left(i\not{\partial}_x - m\right)\, \Psi(x) = 0. \tag{6.24}$$

(6.24) is covariant if Ψ transforms under Lorentz transformations such that in the new system the identical equation holds. This condition defines the behavior of $\Psi \in \mathcal{H}$ (the Hilbert space of states) under Lorentz transformations and thereby defines a four-dimensional representation of the Lorentz group in $\mathcal{H}$: For $\Lambda \in \mathcal{L}$ a Lorentz transformation there exists a representation matrix $S(\Lambda)$ such that

$$\begin{aligned} \Psi'(x') &= S(\Lambda)\Psi(x), \\ x' &= \Lambda x, \end{aligned} \tag{6.25}$$

is the wave function in the new (primed) coordinate system. Covariance then requires that

$$\left(i\gamma^\mu \partial'_\mu - m\right) \Psi'(x') = 0 \tag{6.26}$$

holds. (6.26) fixes the representation $S(\Lambda)$ of the Lorentz group: With

$$\begin{aligned} x'^\mu &= \Lambda^\mu{}_\nu\, x^\nu \\ \partial'_\mu &= \left(\Lambda^{-1}\right)^\nu{}_\mu\, \partial_\nu \end{aligned} \tag{6.27}$$

and

$$S^{-1}(\Lambda)\left(i\gamma^\mu \left(\Lambda^{-1}\right)^\nu{}_\mu \partial_\nu - m\right\} S(\Lambda)\, \Psi(x) \stackrel{!}{=} \left(i\gamma^\mu \partial_\mu - m\right) \Psi = 0.$$

$S(\Lambda)$ must satisfy the following conditions,

$$\begin{aligned} \left(\Lambda^{-1}\right)^\nu{}_\mu S^{-1}(\Lambda)\, \gamma^\mu S(\Lambda) &= \gamma^\nu, \\ \Rightarrow \quad S^{-1}(\Lambda)\, \gamma^\mu S(\Lambda) &= \Lambda^\mu{}_\nu \gamma^\nu. \end{aligned} \tag{6.28}$$

Using (6.28) we can determine the representation $S(\Lambda)$. Here we state the final result without proof. An infinitesimal Lorentz transformation is characterized by $\epsilon^\mu{}_\nu$ with $\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \epsilon^\mu{}_\nu$. Multiplying this relation with $g^{\mu\lambda}$ we obtain

$$\Lambda^{\mu\lambda} = g^{\mu\lambda} + \epsilon^{\mu\lambda}. \tag{6.29}$$

[3] A set γ^ν of Dirac matrices is consistent if (6.19) is satisfied.

A Lorentz transformation is characterized by leaving the metric invariant, i.e. $g_{\mu\nu} = \Lambda^{\sigma}{}_{\mu} g_{\sigma\tau} \Lambda^{\tau}{}_{\nu}$ or in matrix notation $g = \Lambda^T g \Lambda$. This leads to the following condition for the infinitesimal transformation,

$$g_{\mu\nu} = g_{\mu\nu} + \epsilon^{\sigma}{}_{\mu} g_{\sigma\nu} + g_{\mu\tau} \epsilon^{\tau}{}_{\nu} + O(\epsilon^2) \quad \Rightarrow \quad 0 = \epsilon^{\sigma}{}_{\mu} g_{\sigma\nu} + g_{\mu\tau} \epsilon^{\tau}{}_{\nu}. \tag{6.30}$$

Raising the indices μ and ν yields the relation $0 = \epsilon^{\nu\mu} + \epsilon^{\mu\nu}$, i.e. ϵ is an antisymmetric tensor in μ and ν. The Lorentz transformation corresponding to Λ is given by

$$S(\epsilon^{\mu\nu}) = I_n - \frac{i}{4} \sigma_{\mu\nu} \epsilon^{\mu\nu} \tag{6.31}$$

with

$$\sigma_{\mu\nu} = \frac{i}{2} (\gamma_\mu \gamma_\nu - \gamma_\nu \gamma_\mu) = \frac{i}{2} [\gamma_\mu, \gamma_\nu]. \tag{6.32}$$

Insertion into (6.28) gives[4]

$$S^{-1}(\epsilon^{\nu\lambda}) \gamma^\mu S(\epsilon^{\nu\lambda}) = \gamma^\mu + \frac{i}{4} [\sigma_{\nu\lambda}, \gamma^\mu] \epsilon^{\nu\lambda} + O(\epsilon^2). \tag{6.33}$$

To demonstrate the covariance of the Dirac equation we need

$$\begin{aligned} \frac{2}{i} [\sigma^{\nu\lambda}, \gamma^\mu] &= [\overbrace{[\gamma^\nu, \gamma^\lambda]}^{-2g^{\nu\lambda} + 2\gamma^\nu\gamma^\lambda}, \gamma^\mu] = 2[\gamma^\nu \gamma^\lambda, \gamma^\mu] = 2\gamma^\nu \{\gamma^\lambda, \gamma^\mu\} - 2\{\gamma^\nu, \gamma^\mu\} \gamma^\lambda \\ &= 4 g^{\lambda\mu} \gamma^\nu - 4 g^{\nu\mu} \gamma^\lambda. \end{aligned} \tag{6.34}$$

We use this result in (6.33) and obtain

$$\begin{aligned} S^{-1}(\epsilon^{\nu\lambda}) \gamma^\mu S(\epsilon^{\nu\lambda}) &= \gamma^\mu + \frac{i}{4} [\sigma_{\nu\lambda}, \gamma^\mu] \epsilon^{\nu\lambda} = \gamma^\mu - \frac{1}{2} (\gamma_\nu \epsilon^{\nu\mu} - \gamma_\lambda \epsilon^{\mu\lambda}) \\ &= (\delta^{\mu}{}_{\nu} + \epsilon^{\mu}{}_{\nu}) \gamma^\nu = \Lambda^{\mu}{}_{\nu} \gamma^\nu. \end{aligned} \tag{6.35}$$

This shows that the S in (6.31) satisfies the condition (6.28) and the Dirac equation with this choice is Lorentz invariant.

It is important to realize that the transformation matrix $S(\epsilon^{\mu\nu})$ is not unitary. From the fact that γ^μ is Hermitian for $\mu = 0$ and anti-Hermitian otherwise, it follows that $\gamma^0 (\gamma^\mu)^\dagger \gamma^0 = \gamma^\mu$ and hence

$$\gamma^0 (\sigma^{\mu\nu})^\dagger \gamma^0 = -\frac{i}{2} \gamma^0 [(\gamma^\nu)^\dagger, (\gamma^\mu)^\dagger] \gamma^0 = \frac{i}{2} [\gamma^0 (\gamma^\mu)^\dagger \gamma^0, \gamma^0 (\gamma^\nu)^\dagger \gamma^0] = \sigma^{\mu\nu}.$$

With this relation we obtain the generalized unitarity relation (to first order in ϵ)

$$\begin{aligned} S^\dagger(\epsilon^{\mu\nu}) \gamma^0 S(\epsilon^{\mu\nu}) &= (I_n + \frac{i}{4} \sigma^\dagger_{\mu\nu} \epsilon^{\mu\nu}) \gamma^0 (I_n - \frac{i}{4} \sigma_{\mu\nu} \epsilon^{\mu\nu}) \\ &= \gamma^0 + \frac{i}{4} \underbrace{\sigma^\dagger_{\mu\nu} \gamma^0}_{\gamma^0 \sigma^{\mu\nu}} \epsilon^{\mu\nu} - \frac{i}{4} \gamma^0 \sigma_{\mu\nu} \epsilon^{\mu\nu} = \gamma^0 \end{aligned} \tag{6.36}$$

[4] Note that for an infinitesimal transformation $S^{-1}(\epsilon^{\mu\nu}) = S(-\epsilon^{\mu\nu})$, since $S^{-1}(\epsilon^{\mu\nu}) S(\epsilon^{\mu\nu}) = (I_n + \frac{i}{4} \sigma_{\mu\nu} \epsilon^{\mu\nu})(I_n - \frac{i}{4} \sigma_{\mu\nu} \epsilon^{\mu\nu}) = I_n + O(\epsilon^2)$.

or, written differently, $\gamma^0 S^\dagger \gamma^0 = S^{-1}$.

Finally, let us include the electromagnetic interaction, which we achieve by the substitution

$$-i\hbar\,\partial^\mu = p^\mu \to p^\mu - eA^\mu \tag{6.37}$$

$$i\hbar\partial_t\Psi \;=\; H\Psi = \left[c\alpha_i\left(-i\hbar\partial_{x_i} - eA_i/c\right) + e\phi + \beta mc^2\right]\Psi, \tag{6.38}$$

or in covariant notation with $\hbar = c = 1$,

$$\left[\gamma^\mu\left(i\partial_\mu - eA_\mu\right) - m\right]\,\Psi(x) = 0. \tag{6.39}$$

6.3 Conserved Norm

Let Ψ_a and Ψ_b be solutions of the Dirac equation, which we multiply with the corresponding spinor as follows,

$$\begin{aligned}\Psi_a^\dagger\cdot\left(i\hbar\partial_t\Psi_b = H\Psi_b\right),\\ \left(-i\hbar\partial_t\Psi_a^\dagger = (H\Psi_a)^\dagger\right)\cdot\Psi_b.\end{aligned} \tag{6.40}$$

Subtracting the two equations yields

$$\begin{aligned}i\hbar\partial_t(\Psi_a^\dagger\Psi_b) &= \Psi_a^\dagger H\Psi_b - (H\Psi_a)^\dagger\Psi_b\\ &= \frac{\hbar}{i}\,\alpha_i\Psi_a^\dagger(\overrightarrow{\partial_{x_i}} + \overleftarrow{\partial_{x_i}})\Psi_b\\ &= \frac{\hbar}{i}\,\partial_{x_i}(\Psi_a^\dagger\,\alpha_i\Psi_b)\end{aligned} \tag{6.41}$$

and we obtain

$$\partial_t(\Psi_a^\dagger\Psi_b) + \partial_{x_i}(\Psi_a^\dagger\,\alpha_i\Psi_b) = 0, \tag{6.42}$$

or, inserting ones with regard to $I_n = \beta^2 = (\gamma^0)^2$ and $\alpha_i = \beta\,\beta\alpha_i = \gamma^0\gamma^i$,

$$\partial_\mu\left(\overline{\Psi_a}\gamma^\mu\Psi_b\right) = 0; \tag{6.43}$$

where we have defined the Dirac adjoint

$$\overline{\Psi_a} \equiv \Psi_a^\dagger\gamma^0 \tag{6.44}$$

Integration over the spatial coordinates, taking into account the boundary conditions ($\int d^3r\,\boldsymbol{\nabla}\cdot\boldsymbol{j} = \int d\sigma\,\boldsymbol{j}\cdot\boldsymbol{n} = 0$), gives us the conserved quantity

$$\begin{aligned}\text{konst.} &= \int d^3r\;\overline{\Psi_a}\gamma^0\Psi_b\\ &= \int d^3r\;\Psi_a^\dagger\Psi_b \;=\; \langle\Psi_a,\Psi_b\rangle.\end{aligned} \tag{6.45}$$

(6.45) defines a norm, $\|\Psi\| = \langle\Psi,\Psi\rangle \geq 0$.

Note: The quantity $\Psi^\dagger\Psi = \overline{\Psi}\gamma^0\Psi$ is not a scalar under Lorentz transformations, but rather the 0^{te} component of the 4^{er} vector

$$\rho^\mu = \overline{\Psi}\gamma^\mu\Psi. \tag{6.46}$$

The density ρ^μ transforms like a 4^{er} vector:

$$\begin{aligned}\rho'^\mu(x') &= \overline{\Psi'}(x')\gamma^\mu\Psi'(x') \\ &= \overline{\Psi}(x)\underbrace{S^{-1}(\Lambda)\gamma^\mu S(\Lambda)}_{(6.28)\to\Lambda^\mu_\nu\gamma^\nu}\Psi(x) \\ &= \Lambda^\mu_\nu\rho^\nu(x),\end{aligned} \tag{6.47}$$

where we have used that

$$\begin{aligned}\overline{\Psi'}(x') &= \Psi^{\dagger\prime}(x')\gamma^0 = \Psi^\dagger(x)S^\dagger\gamma^0 \\ &= \underbrace{\Psi^\dagger(x)\gamma^0}\underbrace{\gamma^0 S^\dagger\gamma^0} \\ &= \quad\overline{\Psi}(x) \quad S^{-1}.\end{aligned} \tag{6.48}$$

6.4 Free Particles

We make the (block-) ansatz[5],

$$\Psi^{(\pm)}_{\boldsymbol{p}}(\boldsymbol{r};t) = N_p\, e^{i(\boldsymbol{p}\cdot\boldsymbol{r}\mp E_p t)/\hbar}\begin{pmatrix}\chi\\ \eta\end{pmatrix}, \tag{6.49}$$

with χ,η two 2^{er} spinors, the momentum $\boldsymbol{p} = \hbar\boldsymbol{k}$, $\boldsymbol{k} = 2\pi\boldsymbol{n}/L$ and $E_p > 0$. Insertion into the Dirac equation (6.19) yields

$$\pm E_p\begin{pmatrix}\chi\\ \eta\end{pmatrix} = \begin{pmatrix}mc^2 & c\boldsymbol{\sigma}\cdot\boldsymbol{p}\\ c\boldsymbol{\sigma}\cdot\boldsymbol{p} & -mc^2\end{pmatrix}\begin{pmatrix}\chi\\ \eta\end{pmatrix}. \tag{6.50}$$

The secular determinant is

$$(mc^2 \mp E_p)(-mc^2 \mp E_p) - c^2(\boldsymbol{\sigma}\cdot\boldsymbol{p})^2, \tag{6.51}$$

and we obtain solutions if the dispersion relation

$$E_p^2 = c^2p^2 + m^2c^4 \tag{6.52}$$

is satisfied, where we have used that $(\boldsymbol{\sigma}\cdot\boldsymbol{p})^2 = \sigma_i^2p_i^2 + \{\sigma_i,\sigma_j\}p_ip_j/2 = \sigma_i^2\,p_i^2 = I_2\,p_i^2 = p^2$. For the eigenvectors we obtain

$$\begin{aligned}+ &: \eta = \left(\frac{c\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p + mc^2}\right)\chi, \\ - &: \chi = -\left(\frac{c\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p + mc^2}\right)\eta,\end{aligned} \tag{6.53}$$

[5]Despite the positive semidefinite norm $\|\cdot\| \geq 0$ solutions with negative energies will be found.

and with $\chi^\dagger\chi \equiv 1$, as well as $\eta^\dagger\eta \equiv 1$ we obtain for the normalization[6]

$$\begin{aligned}\langle \Psi^{(+)}_{\boldsymbol{p}} \Psi^{(+)}_{\boldsymbol{p}} \rangle \;&=\; 1 = N_p^2 \Omega \chi^\dagger \left(1, \frac{c\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p + mc^2}\right) \begin{pmatrix} 1 \\ \frac{c\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+mc^2} \end{pmatrix} \chi \\ &= N_p^2 \Omega \left(1 + \frac{c^2 p^2}{(E_p + mc^2)^2}\right) \\ &= N_p^2 \Omega \left(1 + \frac{E_p^2 - m^2c^4}{(E_p + mc^2)^2}\right) \\ &= N_p^2 \Omega \frac{2E_p}{E_p + mc^2} \end{aligned} \tag{6.54}$$

and the same result holds for $\Psi^{(-)}_{\boldsymbol{p}}$.

To simplify this cumbersome notation one passes to the Dirac spinors.

6.4.1 Dirac Spinor

The free solutions of Dirac theory can be simply expressed by defining the Dirac spinors

$$\begin{aligned} u(\boldsymbol{p}, s) &\equiv \sqrt{E_p + mc^2} \begin{pmatrix} 1 \\ \frac{c\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p + mc^2} \end{pmatrix} \chi^{(s)}, \\ v(\boldsymbol{p}, s) &\equiv \sqrt{E_p + mc^2} \begin{pmatrix} \frac{c\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p + mc^2} \\ 1 \end{pmatrix} (-i\sigma_2 \chi^{(s)}), \end{aligned} \tag{6.55}$$

$$\begin{aligned} \Psi^{(+)}_{\boldsymbol{p},s}(\boldsymbol{r}; t) &= \frac{1}{\sqrt{2E_p\Omega}}\, u(\boldsymbol{p}, s)\, e^{i(\boldsymbol{p}\cdot\boldsymbol{r} - E_p t)/\hbar}, \\ \Psi^{(-)}_{-\boldsymbol{p},-s}(\boldsymbol{r}; t) &= \frac{1}{\sqrt{2E_p\Omega}}\, v(\boldsymbol{p}, s)\, e^{-i(\boldsymbol{p}\cdot\boldsymbol{r} - E_p t)/\hbar}. \end{aligned} \tag{6.56}$$

The 2^{er} spinor $\chi^{(s)}$ describes the spin of a spin-1/2 particle,

$$\begin{aligned} \chi^{1/2} &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ \chi^{-1/2} &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}; \end{aligned} \tag{6.57}$$

[6]The result is not Lorentz covariant, since $\|\cdot\|$ is not a scalar but the 0 component of a 4 vector.

likewise $-i\sigma_2\chi^{(s)}$ describes a spin, although

$$\begin{aligned} -i\sigma_2\chi^{(1/2)} = \eta^{(1/2)} &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \\ -i\sigma_2\chi^{(-1/2)} = \eta^{(-1/2)} &= \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \end{aligned} \tag{6.58}$$

and $\eta^{(s)}$ describes a down (up) spin for $s = 1/2$ $(= -1/2)$, exactly opposite to $\chi^{(s)}$,

$$\chi^{(s)} \propto \eta^{(-s)}. \tag{6.59}$$

6.5 Interpretation

The Dirac equation again possesses solutions with negative energy, see $\Psi^{(-)}_{-\boldsymbol{p},-s}$ in (6.56). In addition to the component structure $\binom{\chi}{\eta}$, which we also found for the Klein-Gordon equation and interpreted as particle-antiparticle pairs, the Dirac equation has a 2^{er} spinor structure in χ and η. This spinor structure is interpreted as the spin state of a spin-1/2 particle. Thus $\Psi^{(+)}_{\boldsymbol{p},s}$ describes a particle with momentum $\boldsymbol{p}$ and spin s, while the state $\Psi^{(-)}_{-\boldsymbol{p},-s}$ with momentum $-\boldsymbol{p}$ and spin $-s$ describes an antiparticle with momentum $\boldsymbol{p}$ and spin s. The states $\Psi^{(+)}$ and $\Psi^{(-)}$ are orthogonal, since

$$v^\dagger(-\boldsymbol{p}, s') \cdot u(\boldsymbol{p}, s) = u^\dagger(\boldsymbol{p}, s) \cdot v(-\boldsymbol{p}, s') \;=\; 0 \tag{6.60}$$

and hence

$$\begin{aligned} &\int d^3r\ \Psi^{(-)\dagger}_{\boldsymbol{p},-s'} \Psi^{(+)}_{\boldsymbol{p},s'} = \int \frac{d^3r}{2E_p\Omega}\, e^{-i(\boldsymbol{p}\cdot\boldsymbol{r}+E_pt)/\hbar} v^\dagger(-\boldsymbol{p}, s') \cdot u(\boldsymbol{p}, s)\, e^{i(\boldsymbol{p}\cdot\boldsymbol{r}-E_pt)/\hbar} \\ &= \frac{1}{2E_p}\, e^{-2iE_pt/\hbar}\, v^\dagger(-\boldsymbol{p}, s') \cdot u(\boldsymbol{p}, s) \\ &= \frac{E_p + mc^2}{2E_p}\, e^{-2iE_pt/\hbar}\, \eta^{(s')\dagger} \underbrace{\left(\frac{-c\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+mc^2}, 1\right) \begin{pmatrix} 1 \\ \dfrac{c\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+mc^2} \end{pmatrix}}_{0} \chi^{(s)}. \\ &= 0. \end{aligned} \tag{6.61}$$

Note that other combinations with $\boldsymbol{p} \neq \boldsymbol{p}'$ vanish because of $\int d^3r \cdots = 0$, but for $\Psi^{(-)\dagger}\Psi^{(+)}$ one obtains $\int d^3r \cdots = \Omega$.

Finally, the free solutions form a complete system,

$$\sum_{\boldsymbol{p},s} \left(\Psi^{(+)}_{\boldsymbol{p},s}(\boldsymbol{r}) \Psi^{(+)\dagger}_{\boldsymbol{p},s}(\boldsymbol{r}\,') + \Psi^{(-)}_{\boldsymbol{p},s}(\boldsymbol{r}) \Psi^{(-)\dagger}_{\boldsymbol{p},s}(\boldsymbol{r}\,') \right) = \begin{pmatrix} -I_2 & 0 \\ 0 & -I_2 \end{pmatrix} \delta^{(3)}(\boldsymbol{r} - \boldsymbol{r}'). \tag{6.62}$$

The antiparticles are again obtained by charge conjugation.

6.5.1 Charge Conjugation C

The charge conjugation

$$C: \Psi \to \Psi^C = C\beta\Psi^* \tag{6.63}$$

with

$$C = -i\alpha_2 = \begin{pmatrix} 0 & -i\sigma_2 \\ -i\sigma_2 & 0 \end{pmatrix} \tag{6.64}$$

transforms states of negative energy into those of positive energy, which obey dynamics with reversed charge: With

$$i\hbar\Psi = \left[c\,\alpha_i\,(-i\hbar\partial_{x_i} - eA_i/c) + e\phi + \beta mc^2\right]\Psi \tag{6.65}$$

it follows by complex conjugation and multiplication with $C\,\beta$:

$$\begin{aligned} -i\hbar\partial_t\Psi^C &= C\beta\left[c\,\alpha_i^*\,(i\hbar\partial_{x_i} - eA_i/c) + e\phi + \beta mc^2\right]\Psi^* \\ \downarrow C\beta\alpha_i^*\Psi^* &= \underbrace{C\beta C^{-1}}_{-\beta}\underbrace{C\alpha_i^* C^{-1}}_{-\alpha_i}\underbrace{C\beta C^{-1}}_{-\beta}\underbrace{C\beta\Psi^*}_{\Psi^C} \\ &= -\beta\alpha_i\beta\Psi^C = \alpha_i\Psi^C, \end{aligned}$$

$$\begin{aligned} \downarrow C\beta^2\Psi^* &= \underbrace{C\beta C^{-1}}_{-\beta}\underbrace{C\beta\Psi}_{\Psi^C} = -\beta\Psi^C, \\ &= \left[c\,\alpha_i\,(i\hbar\partial_{x_i} - eA_i/c) + e\phi - \beta mc^2\right]\Psi^C, \end{aligned} \tag{6.66}$$

and hence Ψ^C obeys the dynamics

$$\begin{aligned} i\hbar\partial_t\Psi^C &= \left[c\,\alpha_i\,(-i\hbar\partial_{x_i} + eA_i/c) - e\phi + \beta mc^2\right]\Psi^C, \\ &= H(-e)\Psi^C. \end{aligned} \tag{6.67}$$

Charge conjugation C also maps the states $\Psi^{(-)}_{-\boldsymbol{p},-s}$ to $\Psi^{(+)}_{\boldsymbol{p},s}$, since

$$\begin{aligned} C\beta v^*(\boldsymbol{p},s) &= \sqrt{E_p + mc^2}\begin{pmatrix} 0 & -i\sigma_2 \\ -i\sigma_2 & 0 \end{pmatrix}\begin{pmatrix} \dfrac{c\boldsymbol{\sigma}^*\cdot\boldsymbol{p}}{E_p + mc^2} \\ -1 \end{pmatrix}\left(-i\sigma_2\chi^{(s)}\right) \\ &= \sqrt{E_p + mc^2}\begin{pmatrix} 1 \\ \dfrac{c\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p + mc^2} \end{pmatrix}\chi^{(s)} = u(\boldsymbol{p},s). \end{aligned} \tag{6.68}$$

Here we have used that $\sigma_2\boldsymbol{\sigma}^*\sigma_2 = -\boldsymbol{\sigma}$. Analogous to (6.68) is

$$C\beta u^*(\boldsymbol{p},s) = v(\boldsymbol{p},s). \tag{6.69}$$

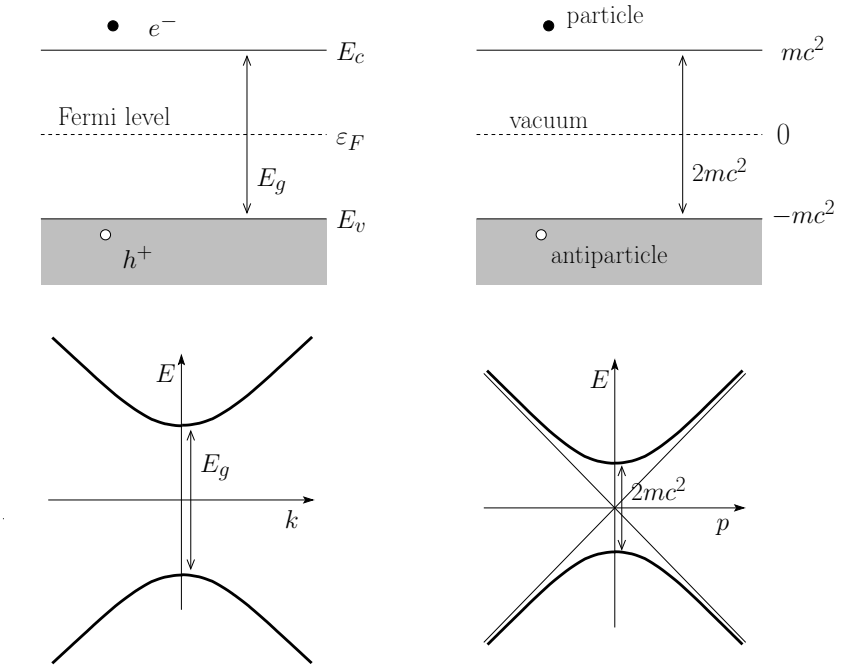


Figure 6.1: Left: Conduction and valence band (top) and dispersion (bottom) of electrons in a semiconductor with a direct band gap (e.g. GaAs, ZnO, ...). The excitations (missing electrons) in the valence band are interpreted as holes. Right: Hole theory of the Dirac equation. The states of negative energy are filled and block the decay of states of positive energy. The excitations with negative energy are interpreted as antiparticles.

Correspondingly we find that

$$\Psi^{(-)C}_{-\boldsymbol{p},-s} = C\beta\Psi^{(-)*}_{-\boldsymbol{p},-s} \;=\; \Psi^{(+)}_{\boldsymbol{p},s}, \tag{6.70}$$

and we identify the (+) energy state $\Psi^{(-)C}_{-\boldsymbol{p},-s}$, which obeys the $-e$ dynamics, as the antiparticle with identical mass and spin and with momentum $\boldsymbol{p}$ and spin component s:

$\Psi^{(+)}_{\boldsymbol{p},s}$ = particle with $E_p > 0$, momentum $\boldsymbol{p}$ and spin s,

$\Psi^{(-)}_{-\boldsymbol{p},-s}$ = state with $E_p < 0$, momentum $-\boldsymbol{p}$ and spin $-s$, interpreted as antiparticle with $E_p > 0$, momentum $\boldsymbol{p}$ and spin s.

Since these particles are fermions the Pauli exclusion principle applies and we can give an interpretation that is not possible for the bosonic Klein-Gordon particles:

If we interpret the physical vacuum as filled with negative-energy states, then the $E_p > 0$ states are stable due to Pauli blocking. Hence the lowest state has energy mc^2 and not $-\infty$. The state $\Psi^{(-)}_{-\boldsymbol{p},-s}$ describes the absence of the particle with equal

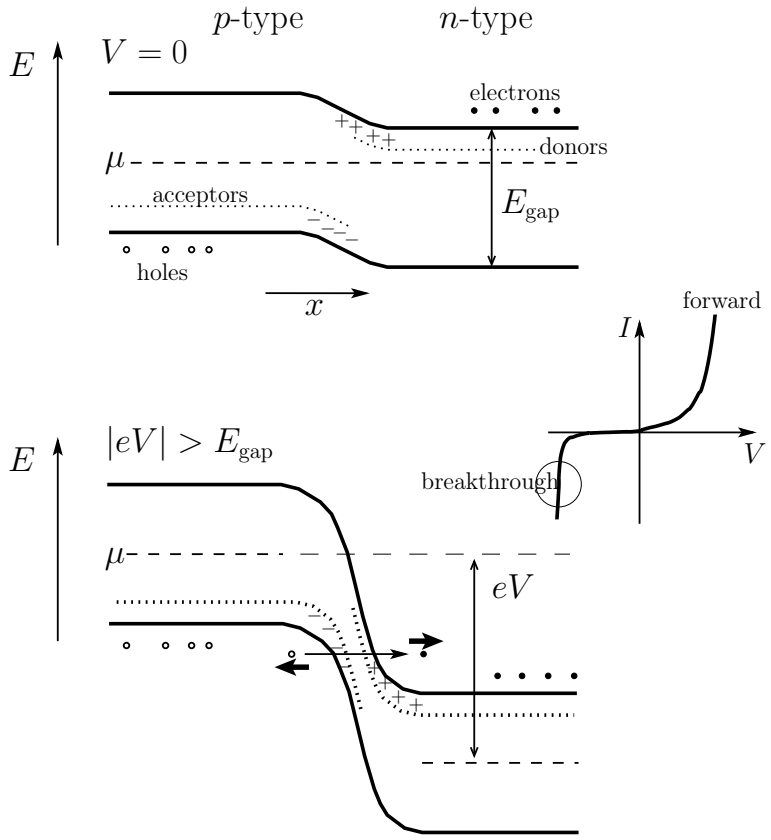


Figure 6.2: Electron-hole pair creation in the electric field of a p-n semiconductor junction. Top: Equilibrium situation without applied voltage, $V = 0$. Bottom: Situation in the reverse-biased p-n junction in breakdown regime where particle-hole pairs are generated via a tunneling process. Cf. with the current-voltage characteristic of the p-n diode.

mass and equal spin, but, being a hole, with opposite charge. Its energy is

$$0 - (-E) = \underbrace{0}_{\text{Vakuum Energie}} \overbrace{-}^{\text{Teilchen fehlt}} \underbrace{(-E)}_{\text{En. des Teilchens}} = E > 0, \tag{6.71}$$

its momentum $0 - (-\boldsymbol{p}) = \boldsymbol{p}$ and its spin $0 - (-s) = s$. This hole theory of the Dirac equation is the relativistic generalization of the hole theory in semiconductors, as shown in Figure 6.1.

Furthermore, it is interesting to discuss the Klein paradox (scattering at a potential step) for fermions (even though pair production only really makes sense in many-body physics). One finds the following results: 1) For $eV > 2mc^2$ both bosons and fermions are totally reflected. 2) For fermions, due to the Pauli principle, no pair

can accompany the reflected particle. 3) Reflected bosons can be accompanied by further boson pairs (stimulated emission in the channel of the reflected boson). 4) For fermions, particle-antiparticle pairs can be spontaneously produced. As an analogue to spontaneous electron-positron (e^--e^+) pair production in a strong electric field, we mention particle-hole pair production via the tunnel effect at a semiconductor p-n junction in the breakdown regime, see Figure 6.2.

6.6 Nonrelativistic limit

We wish to derive the nonrelativistic limit up to order p^4/m^3, i.e. $p^2/2m \cdot (v/c)^2$, including first relativistic corrections. It is $p^2/2m \simeq e\phi \ll mc^2$. The simplest route unfortunately does not work: first we consider a solution

$$\Psi = \begin{pmatrix} \chi \\ \eta \end{pmatrix} e^{-iEt/\hbar} \tag{6.72}$$

with $E > 0$, $\chi = \chi(\boldsymbol{r})$ and $\eta = \eta(\boldsymbol{r})$. With $E = T + mc^2$ we can write

$$\begin{aligned} T\chi &= e\phi\chi + c\boldsymbol{\sigma} \cdot (\boldsymbol{p} - e\boldsymbol{A}/e)\,\eta, \\ (2mc^2 + T)\eta &= e\phi\eta + c\boldsymbol{\sigma} \cdot (\boldsymbol{p} - e\boldsymbol{A}/c)\,\chi. \end{aligned} \tag{6.73}$$

The second equation gives

$$\eta = \frac{1}{2mc^2 + T - e\phi} c\boldsymbol{\sigma} \cdot (\boldsymbol{p} - e\boldsymbol{A}/c)\chi \tag{6.74}$$

and inserting into (6.73) leads to

$$T\chi = \Big[e\phi + c\boldsymbol{\sigma} \cdot \big(\boldsymbol{p} - e\boldsymbol{A}/e\big) \frac{1}{2mc^2 + T - e\phi} c\boldsymbol{\sigma} \cdot \big(\boldsymbol{p} - e\boldsymbol{A}/e\big)\Big]\chi. \tag{6.75}$$

If we want to go to order p^4/m^3, we must expand the denominator accordingly and obtain

$$\begin{aligned} T\chi = \Big[&(e\phi + \frac{1}{2m}\Big[\boldsymbol{\sigma} \cdot \big(\boldsymbol{p} - e\boldsymbol{A}/c\big)\Big]^2 - \frac{1}{4m^2c^2}\boldsymbol{\sigma} \cdot \big(\boldsymbol{p} - e\boldsymbol{A}/c\big) \\ &\times \big(T - e\phi\big)\boldsymbol{\sigma} \cdot \big(\boldsymbol{p} - e\boldsymbol{A}/c\big)\Big]\chi = H(\text{Energie})\chi. \end{aligned} \tag{6.76}$$

The dependence on T on the right-hand side is undesirable since it forces an energy-dependent Hamiltonian upon us. Therefore we replace T by the lowest approximation

$$T \approx e\phi + \frac{1}{2m}\left[\boldsymbol{\sigma} \cdot \big(\boldsymbol{p} - e\boldsymbol{A}/c\big)\right]^2. \tag{6.77}$$

This step, however, is not unique; for the ordering $(\boldsymbol{p} - e\boldsymbol{A}/c)(T - e\phi)$ produces a term $\boldsymbol{p}\, e\phi \propto e\boldsymbol{E}$, while the expression $T - e\phi \approx \left[\boldsymbol{\sigma} \cdot \big(\boldsymbol{p} - e\boldsymbol{A}/c\big)\right]^2/2m$ with $\boldsymbol{\sigma} \cdot \big(\boldsymbol{p} - e\boldsymbol{A}/c\big)$ does not commute. Alternatively one can first commute the factor $\boldsymbol{\sigma} \cdot \big(\boldsymbol{p} - e\boldsymbol{A}/c\big)$

with $(T - e\phi)$ and only then use (6.77), but this procedure likewise is not unique and moreover produces non-Hermitian terms.

A unique nonrelativistic limit can be constructed with the Foldy-Wouthuysen transformation, a unitary transformation that renders the nondiagonal terms small. In the transformed system we can then solve the coupled system of equations. We start from the Dirac equation in the form

$$\begin{aligned} T\begin{pmatrix}\chi\\ \eta\end{pmatrix} &= \left[-mc^2 + e\phi + c\boldsymbol{\alpha}\cdot(\boldsymbol{p} - e\boldsymbol{A}/c) + mc^2\beta\right]\begin{pmatrix}\chi\\ \eta\end{pmatrix} \\ &= H\begin{pmatrix}\chi\\ \eta\end{pmatrix} \qquad (6.78) \\ &\widehat{=} mc^2\left[-1 + \left(\frac{v}{c}\right)^2 + \left(\frac{v}{c} - \left(\frac{v}{c}\right)^2\right) + 1\right]\begin{pmatrix}\chi\\ \eta\end{pmatrix}; \qquad (6.79) \end{aligned}$$

the last equation lists the order of the terms in the relativistic parameter v/c, where the term $c\boldsymbol{\alpha}\cdot\boldsymbol{p} \propto v/c$ is large. By contrast, in the Klein-Gordon equation only terms $\propto 1,\ (v/c)^2$ appear, which is why its nonrelativistic limit is unproblematic. It is precisely the term $c\boldsymbol{\alpha}\boldsymbol{p} \propto (v/c)mc^2$ that causes us trouble, and we will transform it away. To this end we assume a unitary transformation of the form

$$U = U^\dagger = \underbrace{\sqrt{1 - \frac{\lambda^2 p^2}{m^2c^2}}}_{\Lambda}\beta + \frac{\lambda}{mc}\boldsymbol{\alpha}\cdot\boldsymbol{p}, \qquad (6.80)$$

with a parameter λ to be chosen suitably; we will choose λ so that in the transformed system the term $c\boldsymbol{\alpha}\cdot\boldsymbol{p}$ vanishes[7]. We define the transformed spinor

$$\begin{pmatrix}\chi'\\ \eta'\end{pmatrix} = U\begin{pmatrix}\chi\\ \eta\end{pmatrix}, \qquad (6.81)$$

transform (6.78) with U,

$$UT\begin{pmatrix}\chi\\ \eta\end{pmatrix} = T\begin{pmatrix}\chi'\\ \eta'\end{pmatrix} \;=\; UHU^{-1}\begin{pmatrix}\chi'\\ \eta'\end{pmatrix}, \qquad (6.82)$$

and obtain the transformed Hamiltonian

$$H' = UHU^{-1}. \qquad (6.83)$$

[7]The ansatz (6.80) indeed defines a unitary transformation,

$$\begin{aligned} UU^\dagger &= \left(\Lambda\beta + \frac{\lambda}{mc}\boldsymbol{\alpha}\cdot\boldsymbol{p}\right)^2 \\ &= \Lambda^2\beta^2 + \frac{\lambda}{mc}\Lambda\{\beta,\alpha_i\}p_i + \frac{\lambda^2}{m^2c^2}[\{\alpha_i,\alpha_j\}\ p_ip_j/2 + p^2] \\ &= \Lambda^2 + \frac{\lambda^2}{m^2c^2}p^2 = 1. \end{aligned}$$

We transform the terms mc^2, $e\phi$, $c\boldsymbol{\alpha}\cdot(\boldsymbol{p}-e\boldsymbol{A}/c)$, and $mc^2\beta$ and obtain

$$
\begin{aligned}
&U\,[mc^2]\,U^{-1} = mc^2, && \text{(a)}\\
&U\,[e\phi]\,U^{-1} = \Lambda\, e\phi\,\Lambda + \beta\frac{\lambda}{mc}\,(\Lambda\, e\phi\,\boldsymbol{\alpha}\cdot\boldsymbol{p} - \boldsymbol{\alpha}\cdot\boldsymbol{p}\, e\phi\,\Lambda) && \text{(b)}\\
&\qquad + \frac{\lambda^2}{m^2c^2}\boldsymbol{\alpha}\cdot\boldsymbol{p}\, e\phi\,\boldsymbol{\alpha}\cdot\boldsymbol{p},\\
&U\,\big[c\boldsymbol{\alpha}\cdot\big(\boldsymbol{p}-e\boldsymbol{A}/c\big)\big]\,U^{-1} = \underline{-\Lambda c\,\boldsymbol{\alpha}\cdot(\boldsymbol{p}-e\boldsymbol{A}/c)\,\Lambda} && \text{(c)}\\
&\qquad + \beta\frac{\lambda}{m}[\Lambda\,\boldsymbol{\alpha}\cdot(\boldsymbol{p}-e\boldsymbol{A}/c)\;\boldsymbol{\alpha}\cdot\boldsymbol{p} + \boldsymbol{\alpha}\cdot\boldsymbol{p}\;\boldsymbol{\alpha}\cdot(\boldsymbol{p}-e\boldsymbol{A}/c)\,\Lambda] && (6.84)\\
&\qquad + \frac{\lambda^2}{m^2c}\boldsymbol{\alpha}\cdot\boldsymbol{p}\;\boldsymbol{\alpha}\cdot(\boldsymbol{p}-e\boldsymbol{A}/c)\,\boldsymbol{\alpha}\cdot\boldsymbol{p},\\
&U\,[mc^2\beta]\,U^{-1} = mc^2\beta\Lambda^2 + \underline{2\lambda\Lambda c\boldsymbol{\alpha}\cdot\boldsymbol{p}} - \beta\lambda^2p^2/m. && \text{(d)}
\end{aligned}
$$

Critical are the (underlined) nondiagonal terms of odd order in $\boldsymbol{\alpha}$ and of relative order v/c. With

$$\Lambda \approx 1 - \frac{\lambda^2p^2}{2m^2c^2} = 1 - \mathcal{O}[(v/c)^2], \tag{6.85}$$

we obtain (the index nd stands for "nondiagonal")

$$
\begin{aligned}
H'_{\mathrm{nd},v/c} &= -c\boldsymbol{\alpha}\cdot\big(\boldsymbol{p}-e\boldsymbol{A}/c\big) + 2\lambda c\boldsymbol{\alpha}\cdot\boldsymbol{p} + \mathcal{O}[(v/c)^3]\\
&\overset{\lambda=1/2}{=} e\boldsymbol{\alpha}\cdot\boldsymbol{A} \simeq\ \mathcal{O}[(v/c)^2].
\end{aligned}
\tag{6.86}
$$

Equation (6.82) then takes the form

$$T\begin{pmatrix}\chi'\\ \eta'\end{pmatrix} = \begin{pmatrix}H'_{11} & e\boldsymbol{\alpha}\cdot\boldsymbol{A}\\ e\boldsymbol{\alpha}\cdot\boldsymbol{A} & -2mc^2\end{pmatrix}\begin{pmatrix}\chi'\\ \eta'\end{pmatrix}, \tag{6.87}$$

with (we refer to (6.84); the terms $\propto mc^2$ in (a) and (d) combine to $(-I_4+\beta)mc^2$ and give no contribution to H'_{11} but the contribution $-2mc^2$ to H'_{22})

$$
\begin{aligned}
H'_{11} \approx\ & \underbrace{e\phi - \frac{p^2}{8m^2c^2}\,e\phi - e\phi\,\frac{p^2}{8m^2c^2}}_{\text{(b): }\Lambda\, e\phi\,\Lambda} + \underbrace{\frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}\, e\phi\,\boldsymbol{\sigma}\cdot\boldsymbol{p}}{4m^2c^2}}_{\text{(b)}} \qquad (6.88)\\
&+ \underbrace{\frac{1}{2m}\big[\boldsymbol{\sigma}\cdot\big(\boldsymbol{p}-e\boldsymbol{A}/c\big)\;\boldsymbol{\sigma}\cdot\boldsymbol{p} + \boldsymbol{\sigma}\cdot\boldsymbol{p}\;\boldsymbol{\sigma}\cdot\big(\boldsymbol{p}-e\boldsymbol{A}/c\big)\big] - \frac{p^4}{8m^3c^2}}_{\text{(c)}} - \underbrace{\frac{p^2}{2m}}_{\text{(d)}}.
\end{aligned}
$$

The remainder are rearrangements; with $\sigma_i\sigma_j = \delta_{ij} + i\varepsilon_{ijk}\sigma_k$ one obtains for the (c)

term

$$\begin{aligned}\boldsymbol{\sigma}\cdot\big(\boldsymbol{p}-e\boldsymbol{A}/c\big)\boldsymbol{\sigma}\cdot\boldsymbol{p}+\boldsymbol{\sigma}\cdot\boldsymbol{p}\;\boldsymbol{\sigma}\cdot\big(\boldsymbol{p}-e\boldsymbol{A}/c\big)&\\ =2p^2-\boldsymbol{\sigma}\cdot\boldsymbol{p}\;\boldsymbol{\sigma}\cdot(e\boldsymbol{A}/c)-\boldsymbol{\sigma}\cdot(e\boldsymbol{A}/c)\;\boldsymbol{\sigma}\cdot\boldsymbol{p}&\\ =2p^2-(e/c)\big(\boldsymbol{p}\cdot\boldsymbol{A}+\boldsymbol{A}\cdot\boldsymbol{p}\big)-i(e/c)\boldsymbol{\sigma}\cdot\big[\big(\boldsymbol{p}\wedge\boldsymbol{A}\big)-\big(\boldsymbol{A}\wedge\boldsymbol{p}\big)\big]&\\ =\big(\boldsymbol{p}-e\boldsymbol{A}/c\big)^2+p^2-\big(e\boldsymbol{A}/c\big)^2-(e\hbar/c)\boldsymbol{\sigma}\cdot\big(\boldsymbol{\nabla}\wedge\boldsymbol{A}\big)&\\ =\big(\boldsymbol{p}-e\boldsymbol{A}/c\big)^2+p^2-(e/c)^2A^2-(e\hbar/c)\boldsymbol{\sigma}\cdot\boldsymbol{B}.&\end{aligned}\tag{6.89}$$

The (b) terms can likewise be simplified according to

$$\begin{aligned}p^2\phi+\phi p^2&=(p^2\phi)+2(\boldsymbol{p}\phi)\cdot\boldsymbol{p}+2\phi p^2,\\ \boldsymbol{\sigma}\cdot\boldsymbol{p}\;\phi\;\boldsymbol{\sigma}\cdot\boldsymbol{p}&=\boldsymbol{\sigma}\cdot(\boldsymbol{p}\phi)\;\boldsymbol{\sigma}\cdot\boldsymbol{p}+\phi p^2\\ &=(\boldsymbol{p}\phi)\cdot\boldsymbol{p}+i\boldsymbol{\sigma}\cdot[(\boldsymbol{p}\phi)\wedge\boldsymbol{p}]+\phi p^2\end{aligned}$$

and we obtain

$$\begin{aligned}-(e/8m^2c^2)\{p^2\phi+\phi p^2-2\boldsymbol{\sigma}\cdot\boldsymbol{p}\;\phi\;\boldsymbol{\sigma}\cdot\boldsymbol{p}\}&\\ =-(e/8m^2c^2)\{(p^2\phi)-2i\boldsymbol{\sigma}\cdot[(\boldsymbol{p}\phi)\wedge\boldsymbol{p}]\}.&\end{aligned}$$

Inserting into (6.88) yields the Hamiltonian (the diamagnetic term $\propto A^2$ is obtained from $\eta'=(e/2mc^2)\boldsymbol{\alpha}\cdot\boldsymbol{A}\;\chi'+\mathcal{O}(v^4/c^4)$)

$$\begin{aligned}H'_{11}+\frac{e^2}{2mc^2}A^2=\frac{1}{2m}\left(\boldsymbol{p}-\frac{e}{c}\boldsymbol{A}\right)^2+e\phi-\frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot\boldsymbol{B}-\frac{p^4}{8m^3c^2}&\\ +\frac{e\hbar^2(\nabla^2\phi)}{8m^2c^2}+\frac{e\hbar}{4m^2c^2}\boldsymbol{\sigma}\cdot\big[(\boldsymbol{\nabla}\phi)\wedge\boldsymbol{p}\big].&\end{aligned}\tag{6.90}$$

The dynamics of a spin-1/2 particle in an electromagnetic field $\phi,\boldsymbol{A}$ with relativistic corrections up to order $(v/c)^2$ is given by (with $\phi(\boldsymbol{r})=\phi(r)$ we can replace $\boldsymbol{\nabla}\phi\wedge\boldsymbol{p}$ by $[r^{-1}(\partial_r\phi)\boldsymbol{L}$)

$$\begin{aligned}H=\underbrace{\frac{1}{2m}\left(\boldsymbol{p}-\frac{e}{c}\boldsymbol{A}\right)^2+e\phi}_{H_0}-\underbrace{\frac{1}{2mc^2}\left(\frac{p^2}{2m}\right)^2+\frac{e\hbar}{4m^2c^2}\frac{(\partial_r\phi)}{r}\boldsymbol{\sigma}\cdot\boldsymbol{L}+\frac{e\hbar^2\,(\nabla^2\phi)}{8m^2c^2}}_{\text{Feinstruktur}}&\\ \underbrace{-\frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot\boldsymbol{B}}_{\text{Zeeman}}.&\end{aligned}\tag{6.91}$$

Combining the orbital and spin Zeeman terms one obtains

$$H_{\text{Zeeman}}=-\frac{e}{2mc}\boldsymbol{B}\cdot(\boldsymbol{L}+2\boldsymbol{S}),\tag{6.92}$$

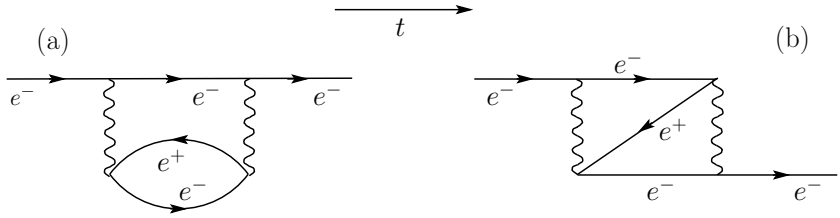


Figure 6.3: Feynman diagrams for virtual electron-hole pair production: (a) vacuum polarization (direct term) and (b) Zitterbewegung (exchange term of vacuum polarization). The electron disappears (via pair annihilation) to reappear at another location (via pair production).

and therefrom the gyromagnetic ratio of the electron spin, $g = 2$. One may argue that the Darwin term $\propto \nabla^2\phi$ stems from the smearing of the electron over the Compton scale[8] $\hbar/mc$: averaging the potential

$$\begin{aligned}\langle e\phi(\boldsymbol{r}+\delta\boldsymbol{r})\rangle &\approx \langle e\phi(\boldsymbol{r}) + e\delta\boldsymbol{r}\cdot\boldsymbol{\nabla}\phi(\boldsymbol{r}) + \frac{e}{2}(\delta\boldsymbol{r}\cdot\boldsymbol{\nabla})^2\phi(\boldsymbol{r})\rangle \\ &\approx e\phi(\boldsymbol{r}) + \frac{e}{6}\frac{\hbar^2}{m^2c^2}(\nabla^2\phi),\end{aligned} \tag{6.93}$$

yields a term of the corresponding structure and order of magnitude. The corresponding quantum fluctuations in the position of the electron go under the name "Zitterbewegung." Unpleasant is that this argument does not involve the spin, for the Klein-Gordon equation shows no corresponding term.

A deeper analysis shows that Zitterbewegung originates from interference between positive- and negative-energy states: a state with purely positive (negative) energy exhibits no Zitterbewegung. Scattering at a potential always mixes a state with $E > 0$ but always entails states with $E < 0$ as well. Within hole theory, Zitterbewegung is interpreted as follows: the potential mixes a $E > 0$ state with components $E < 0$. The interference effects that lead to Zitterbewegung correspond to virtual e^+-e^- pairs. The scattering e^- annihilates the e^+ and the e^- of the e^+-e^- pair takes over its role. This means that the electron jumps. The process sketched in Figure 6.3 is an exchange process.

The extent of the Zitterbewegung is found from the following estimate: the lifetime Δt of the virtual e^+-e^- pair is given by the uncertainty relation[9]: $\Delta t \simeq \hbar/2mc^2$, whence the distance $\Delta r \simeq c\Delta t \simeq \hbar/mc$ results.

The application to the hydrogen atom $\phi(r) = e/r$ yields a fine structure that involves only the quantum numbers n and j, although the Darwin term contributes only to the orbital angular momentum $l = 0$ and the spin-orbit term to $l \neq 0$. The compensation

[8] The Compton scale $\hbar/mc$ is the length scale of the relativistic differential equation.

[9] According to the Heisenberg uncertainty principle, $\Delta E\,\Delta t \geq \hbar$, where ΔE is set by the rest energy mc^2 of the electron.

of these contributions then gives the energy

$$T = E_{nj} = -mc^2 \left[\frac{(Z\alpha)^2}{2n^2} + \frac{(Z\alpha)^4}{2n^4} \left(\frac{n}{j+1/2} - \frac{3}{4} \right) \right]. \tag{6.94}$$

Appendix A

Spherical Bessel Functions

A.1 Definition

The Bessel functions J_n are the solutions of the differential equation

$$\left[\partial_x^2 + \frac{1}{x}\partial_x + \left(1 - \frac{n^2}{x^2}\right)\right] J_n(x) = 0. \tag{A.1}$$

For half-integer $n = l + \frac{1}{2}$ one usually considers the equivalent differential equation

$$\left[\partial_x^2 + \frac{2}{x}\partial_x + \left(1 - \frac{l(l+1)}{x^2}\right)\right] z_l(x) = 0. \tag{A.2}$$

with the spherical Bessel and Neumann functions,

$$\text{Bessel} \to j_l(x) = \sqrt{\frac{\pi}{2x}}\, J_{l+\frac{1}{2}}(x), \tag{A.3}$$

$$\text{Neumann} \to n_l(x) = (-1)^{l+1}\sqrt{\frac{\pi}{2x}}\, J_{-l-\frac{1}{2}}(x). \tag{A.4}$$

as solutions. Note that (A.2) is symmetric under $l + \frac{1}{2} \mapsto -l - \frac{1}{2}$, and thus the Neumann solutions go over into the Bessel solutions. Examples of the spherical Bessel and Neumann functions are (see also Figure A.1 and A.2)

$$\begin{aligned}
j_0 &= \frac{\sin x}{x}, & n_0 &= -\frac{\cos x}{x}, \\
j_1 &= \frac{\sin x}{x^2} - \frac{\cos x}{x}, & n_1 &= -\frac{\cos x}{x^2} - \frac{\sin x}{x}, \\
j_2 &= \left(\frac{3}{x^3} - \frac{1}{x}\right)\sin x - \frac{3}{x^2}\cos x, & n_2 &= -\left(\frac{3}{x^3} - \frac{1}{x}\right)\cos x - \frac{3}{x^2}\sin x.
\end{aligned} \tag{A.5}$$

The spherical Bessel functions exhibit the following behavior for small (series

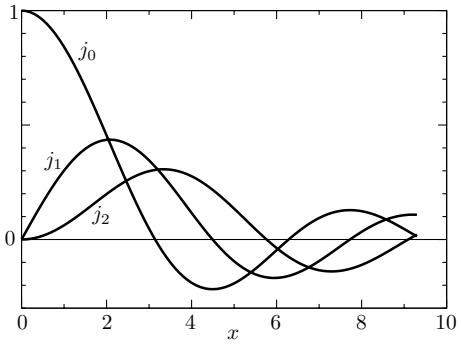


Figure A.1: Bessel functions $j_0(x)$, $j_1(x)$, $j_2(x)$.

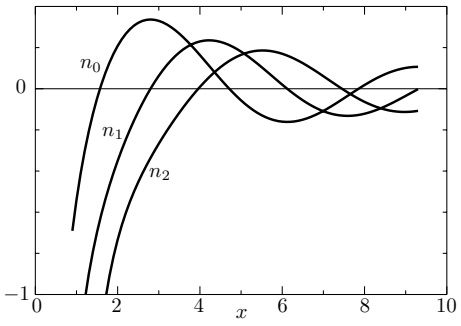


Figure A.2: Neumann functions $n_0(x)$, $n_1(x)$, $n_2(x)$.

expansion for $x \to 0$) and large (asymptotics for $x \to \infty$) arguments:

$$\begin{aligned} &x \to 0: && x \to \infty: \\ &j_l \sim \frac{x^l}{(2l+1)!!}, && j_l \sim \frac{1}{x}\cos[x-(l+1)\pi/2], \\ &n_l \sim -\frac{(2l-1)!!}{x^{l+1}}, && n_l \sim \frac{1}{x}\sin[x-(l+1)\pi/2]. \end{aligned} \tag{A.6}$$

where $(2l-1)!! = 1\cdot 3\cdot \ldots \cdot (2l-3)\cdot(2l-1) = \frac{(2l)!}{2^l l!}$ denotes the double factorial. For the spherical Bessel functions $z_l = j_l, n_l$ the following recursion formulas ($l > 0$)

$$\begin{aligned} z_{l-1} + z_{l+1} &= \frac{2l+1}{x}\, z_l, \\ \partial_x z_l &= z_{l-1} - \frac{l+1}{x}\, z_l. \end{aligned} \tag{A.7}$$

also hold. Note that n_l is always singular for $x \to 0$. We further define the spherical Hankel functions as linear combinations

$$\begin{aligned} h_l^{(1)}(x) &= j_l(x) + i\, n_l(x), \\ h_l^{(2)}(x) &= j_l(x) - i\, n_l(x), \end{aligned} \tag{A.8}$$

with the asymptotic behavior

$$h_l^{(1,2)} \sim \frac{1}{x}\, e^{\pm i[x-(l+1)\pi/2]}. \tag{A.9}$$

A.2 Expansion of the Plane Wave

The plane wave $e^{ikr\cos\theta}$ is a solution of (1.2) and hence of (1.33) with $V \equiv 0$. Since, as discussed in the previous chapter, $j_l(x = kr)$ solves the differential equation (1.33),

we may choose for the plane wave the ansatz

$$e^{ikr\cos\theta} = \sum_l A_l\, j_l(kr)\, \underbrace{P_l(\cos\theta)}_{\propto Y_{l0}} \tag{A.10}$$

which separates the radial and angular dependences. Finally, we must determine the coefficients A_l. Using (1.29), i.e. $P_l \perp P_{l'}$, we obtain by integration, exploiting the normalization $\int dz\, P_l(z)^2 = 2/(2l+1)$

$$\frac{2}{2l+1} A_l\, j_l(kr) = \int_{-1}^{1} dz\, P_l(z) e^{ikrz}. \tag{A.11}$$

We consider $r \to 0$, expand $\exp(ikrz)$ on the right-hand side and $j_l(kr)$ on the left-hand side in $x = kr$ and then compare coefficients. First the right-hand side: We expand $\exp(ikrz)$ and obtain

$$\int_{-1}^{1} dz\, P_l(z) \sum_{s=0}^{\infty} \frac{(ikrz)^s}{s!} \approx \int_{-1}^{1} dz\, P_l(z) \frac{(ikrz)^l}{l!}, \tag{A.12}$$

where we have used that in the limit of small x only the term $s = l$ gives a non-negligible contribution [terms with $s < l$ do not contribute because $P_l \perp z^s = \sum_{i \le s} c_i P_i$; terms with $s > l$ give a small correction of order $O(x^{l+2})$].

Next we replace z^l by Legendre polynomials $P_{l'}(z)$ with $l' \le l$. Under the integral, terms with $l' < l$ vanish (orthogonality of the Legendre polynomials). We have

$$P_l(z) = \frac{(2l-1)!!}{l!} z^l + O(z^{l-2}). \tag{A.13}$$

for the term with $l' = l$, so that we may replace z^l by $l! P_l(z)/(2l-1)!!$ under the integral. Finally, we use the normalization of $P_l(z)$ to evaluate the integral over z and obtain

$$\int_{-1}^{1} dz\, P_l(z) \sum_{s=0}^{\infty} \frac{(ikrz)^s}{s!} \approx \frac{2(ikr)^l}{(2l+1)!!} \tag{A.14}$$

For the left-hand side: The expansion of $j_l(x)$ for small x yields [see (A.6)]

$$\frac{2}{2l+1} A_l\, j_l(kr) \approx A_l \frac{2(kr)^l}{(2l+1)(2l+1)!!} \tag{A.15}$$

and the comparison of coefficients gives us the desired result

$$A_l = i^l (2l+1). \tag{A.16}$$

For the general case with $\boldsymbol{k}$ and $\boldsymbol{r}$ arbitrary, we may use the addition theorem for Y_{lm} and express $P_l(\cos\vartheta)$ in terms of the spherical harmonics

$$\sum_{m=-l}^{l} Y_{lm}(\theta,\varphi) Y_{lm}^*(\theta',\varphi') = \frac{2l+1}{4\pi} P_l(\cos\vartheta); \tag{A.17}$$

thereby obtaining for the plane wave the expression

$$e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{l} i^l \, j_l(kr) Y_{lm}^*(\Omega_{\boldsymbol{k}}) Y_{lm}(\Omega_{\boldsymbol{r}}). \tag{A.18}$$

Appendix B

Spin and Orbital Species

The analysis of the N spin-1/2 fermion problem is important for the discussion of atoms and very instructive with regard to understanding the role of permutation symmetry. The topic is moreover a nice application of group theory. Here is an attempt to present the problem as simply as possible. First we will briefly discuss the representation theory of S_N. Then we will explain the connection between spin and species in the reduction of $\mathcal{H}_{1/2}^{\otimes N}$.

The chapter introduces all necessary tools and shows by examples how to use them. The results are made plausible. Proofs of the statements can be found, for instance, in the books by Morton Hamermesh "Group theory and its application to physics problems" or Hermann Weyl "The classical groups: their invariants and representations".

B.1 Symmetric Group S_N

The elements π of the group S_N are the permutations of N objects,

$$\begin{aligned}\pi &= \begin{pmatrix} 1 & 2 & \cdots & N \\ \pi(1) & \pi(2) & \cdots & \pi(N) \end{pmatrix} \\ &= \begin{pmatrix} 2 & 4 & 3 & \cdots \\ \pi(2) & \pi(4) & \pi(3) & \cdots \end{pmatrix}; \end{aligned} \tag{B.1}$$

here π is a mapping from $\mathbb{Z}_N$ to $\mathbb{Z}_N$ that maps the top row to the bottom row, i.e. k to $\pi(k)$. The inverse of the permutation π is

$$\pi^{-1} = \begin{pmatrix} \pi(1) & \pi(2) & \cdots & \pi(N) \\ 1 & 2 & \cdots & N \end{pmatrix}. \tag{B.2}$$

The group S_N has $N!$ elements; the order g of S_N is $N!$.

An important class of permutations are the k-cycles

$$\pi_{(\alpha_1\cdots\alpha_k)} = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_k & \beta_1 & \cdots \beta_{n-k} \\ \alpha_2 & \alpha_3 & \alpha_4 & \cdots & \alpha_1 & \beta_1 & \cdots \beta_{n-k} \end{pmatrix}; \tag{B.3}$$

$\pi_{(\alpha_1\cdots\alpha_k)}$ is also denoted briefly by $(\alpha_1\alpha_2\ldots\alpha_k)$; the cycle maps α_1 to α_2, α_2 to α_3, $\ldots$, α_k to α_1 and leaves the elements β_k, which are not involved in the cycle, invariant. The simplest cycle is the two-cycle or transposition $\pi_{(\alpha_1\alpha_2)} = (\alpha_1\alpha_2)$, for example

$$\pi_{(35)} = (35) = \begin{pmatrix} 3 & 5 & \beta_1 & \cdots & \beta_{n-2} \\ 5 & 3 & \beta_1 & \cdots & \beta_{n-2} \end{pmatrix}. \tag{B.4}$$

Every permutation can be built up from transpositions. A permutation is called even (odd) if it can be expressed as an even (odd) number of transpositions. We denote

$$(-1)^\pi \text{ the signature of } \pi, \text{ with } \begin{cases} (-1)^\pi = 1, & \pi \text{ even}, \\ (-1)^\pi = -1, & \pi \text{ odd}. \end{cases} \tag{B.5}$$

The decomposition of π into disjoint cycles[1] determines the equivalence class of the permutation[2] For example S_3 has altogether three equivalence classes: the identity e (as always) forms one equivalence class, the three two-cycles $(\alpha_1\alpha_2)$ with $1 \le \alpha_1 < \alpha_2 \le 3$ form the second equivalence class, and the two three-cycles $(\alpha_1\alpha_2\alpha_3)$ $1 \le \alpha_1 < \alpha_2 < \alpha_2 \le 3$ likewise form their own equivalence class.

In general one can show that π and π' are equivalent if and only if they have the same cycle structure; consider a permutation with cycle structure $3+2$,

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 2 & 1 & 4 \end{pmatrix} = \left(\begin{array}{ccc|cc} 1 & 5 & 4 & 2 & 3 \\ 5 & 4 & 1 & 3 & 2 \end{array}\right) = (1\,5\,4)(2\,3). \tag{B.6}$$

Indeed the transformation $\nu\pi\nu^{-1}$ with $\nu \in S_N$ leaves the cycle structure invariant; we illustrate this with the above example using a specific ν,

$$\begin{aligned} \nu &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 3 & 5 \end{pmatrix} = \begin{pmatrix} 5 & 4 & 1 & 3 & 2 \\ 5 & 3 & 2 & 1 & 4 \end{pmatrix} \\ \nu^{-1} &= \begin{pmatrix} 2 & 4 & 1 & 3 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 2 & 5 & 3 & 4 & 1 \\ 1 & 5 & 4 & 2 & 3 \end{pmatrix}. \end{aligned} \tag{B.7}$$

Then $\nu\pi\nu^{-1}$ can be constructed simply,

$$\begin{aligned} \nu\pi\nu^{-1} &= \begin{pmatrix} 2 & 5 & 3 & 4 & 1 \\ 1 & 5 & 4 & 2 & 3 \\ 5 & 4 & 1 & 3 & 2 \\ 5 & 3 & 2 & 1 & 4 \end{pmatrix} \begin{matrix} \leftarrow & \nu^{-1} \\ \leftarrow & \pi \\ \leftarrow & \nu \end{matrix} \\ &= \begin{pmatrix} 2 & 5 & 3 & 4 & 1 \\ 5 & 3 & 2 & 1 & 4 \end{pmatrix} = (2\,5\,3)(4\,1), \end{aligned} \tag{B.8}$$

[1] that is, each number appears only once.

[2] Two permutations π and π' are equivalent if there exists a $\nu \in S_N$ such that $\nu\pi\nu^{-1} = \pi'$.

and the cycle structure is preserved. This fact can be proved in general and thereby the cycle structure defines an equivalence criterion.

To find the number of equivalence classes of S_N we therefore need all decompositions of N. We call $[\lambda_1, \lambda_2, \ldots, \lambda_k] \equiv \lambda$ a decomposition of N if $\lambda_i \geq 1$ are integers sorted according to $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_k$ and

$$\lambda_1 + \lambda_2 + \cdots + \lambda_k = N. \tag{B.9}$$

The decompositions of N yield the equivalence classes of S_N. For the example with $N = 5$ one finds by simple counting 7 equivalence classes,

$$\begin{aligned} S_5: \qquad 5 \; &= 5 \\ &= 4 + 1 \\ &= 3 + 2 \\ &= 3 + 1 + 1 \\ &= 2 + 2 + 1 \\ &= 2 + 1 + 1 + 1 \\ &= 1 + 1 + 1 + 1 + 1 \end{aligned} \Bigg\} \rightarrow 7 \text{ classes.}$$

The number of irreducible representations of S_N equals the number K of classes. Let d_i be the dimension of the i-th irreducible representation. Then the group order $g = N!$ is given by[3]

$$g = \sum_{i=1}^{K} d_i^2. \tag{B.10}$$

In concrete examples one finds by trial

$$\begin{aligned} S_2: &\qquad K = 2, &&\qquad 2! = 1 + 1, \\ S_3: &\qquad K = 3, &&\qquad 3! = 1 + 1 + 2^2, \\ S_4: &\qquad K = 5, &&\qquad 4! = 24 = 1 + 1 + 2^2 + 3^2 + 3^2, \\ S_5: &\qquad K = 7, &&\qquad 5! = 120 = 1 + 1 + 4^2 + 4^2 + 5^2 + 5^2 + 6^2. \end{aligned}$$

The two one-dimensional representations always occur. These are the antisymmetric (A) and the symmetric (S) representations. They are generated by the (anti-)symmetrization operators

$$\mathcal{A} = \frac{1}{N!} \sum_{\pi \in S_N} (-1)^{\pi} \pi, \tag{B.11}$$

$$\mathcal{S} = \frac{1}{N!} \sum_{\pi \in S_N} \pi, \tag{B.12}$$

[3]This relation between the group order and the sum of the squares of the dimensions of all irreducible representations holds in general for any finite group.

respectively their representations in the Hilbert space. We have already seen in Chapter 2.3 that the representation of the (anti-)symmetrization operator projects onto a one-dimensional subspace of the total Hilbert space. Applying an arbitrary permutation π then acts on this space only by multiplication with $(\pm 1)^\pi$. These properties can be derived simply from the multiplication rules

$$\pi\mathcal{S} = \mathcal{S}\pi = \mathcal{S} \quad \text{and} \quad \pi\mathcal{A} = \mathcal{A}\pi = (-1)^\pi \mathcal{A} \tag{B.13}$$

B.2 Example: 2 Spin-1/2 Fermions

In the study of N spin-1/2 fermions, the central question is how spin, spin symmetry, and orbital symmetry are related. To familiarize ourselves with the topic, we analyze the problem of two spin-1/2 fermions. We separate the total wave function into spin and orbital parts, $\Psi(x_1, x_2) = \varphi(\boldsymbol{r}_1, \boldsymbol{r}_2)\,\chi(s_1, s_2)$. The (total) spin determines the spin symmetry,

$$\begin{aligned} S = 1: \qquad & \chi_t = \begin{cases} |\uparrow\uparrow\rangle \\ \big(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\big)/\sqrt{2}, \\ |\downarrow\downarrow\rangle \end{cases} \quad \text{symmetric under } S_2, \\ S = 0: \qquad & \chi_s = \big(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\big)/\sqrt{2}, \quad \text{antisym. under } S_2. \end{aligned} \tag{B.14}$$

The fermion antisymmetry determines the orbital symmetry

$$S = 1 \qquad \to \chi_t \text{ symmetric} \qquad \to \varphi_{\mathcal{A}} \text{ antisymmetric} \tag{B.15}$$

$$S = 0 \qquad \to \chi_s \text{ antisymmetric} \qquad \to \varphi_{\mathcal{S}} \text{ symmetric.} \tag{B.16}$$

The connection between spin and orbital species, which follows from Weyl's theorem and the Pauli exchange principle, generalizes this situation to a system of N fermions. In general, other irreducible representations of S_N will also occur, which differ from the one-dimensional representations S and A. The total spin fixes the spin symmetry (under exchange with S_N) and, via the Pauli principle, the orbital symmetry (under S_N). Here the notions of spin and orbital species refer to the "quantum numbers" associated with the irreducible representation of the permutation symmetry (just as L denotes the quantum number associated with the irreducible representation of rotational symmetry). Thus the spin and orbital species determine, according to which irreducible representation of S_N the spin and orbital functions transform. We anticipate the result of the decomposition of the Hilbert space of N spin-1/2 particles.

B.3 N Spin-1/2 Particles

For the decomposition of the spin Hilbert space of N particles into irreducible representations one has

$$\begin{aligned} \big(\mathcal{H}_{1/2}\big)^{\otimes N} &= \oplus_S\, \mu_S(N)\, \mathcal{H}_S, \qquad S = \text{ spin quantum number}, \\ &= \oplus_S\, (2S+1)\, \mathcal{H}_{\mu_S}. \end{aligned} \tag{B.17}$$

Here $\mathcal{H}_{1/2}$ is the two-dimensional Hilbert space of a spin-1/2 particle, $\mathcal{H}_{\mu_S}$ the $\mu_S(N)$-dimensional representation space for the $\mu_S(N)$-dimensional irreducible representation of the permutations S_N and $\mathcal{H}_S$ the $(2S+1)$-dimensional representation space for the $(2S+1)$-dimensional irreducible representation of the rotation group $SU(2)$. The spin S fixes the spin species μ_S, i.e., it determines according to which irreducible representation of S_N the spin wave function transforms. If the spin S occurs $\mu_S(N)$ times in the decomposition of $\left(\mathcal{H}_{1/2}\right)^{\otimes N}$, we obtain $2S+1$ times a $\mu_S(N)$-dimensional irreducible representation of S_N, i.e., the irreducible representation of S_N with species μ_S occurs exactly $(2S+1)$ times. Moreover, the spin species μ_S, via the requirement of total antisymmetry for the total wave function Ψ, uniquely fixes the orbital species μ_B. We explain and illustrate this result in what follows. For this we need the irreducible representations of S_N.

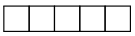
B.4 Irreducible Representations of S_N

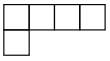
We consider the decomposition of permutations into cycles, for example

$$\begin{aligned}
\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} &= (1\,2\,3)(4) \\
\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} &= (1\,4)(2\,3) \\
\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix} &= (1\,4\,2\,3)
\end{aligned} \tag{B.18}$$

Each decomposition of N defines an equivalence class of S_N. For example, for $N=5$ we find 7 classes with the decompositions $5=5,\ 4+1,\ 3+2,\ 3+1+1,\ 2+2+1,\ 2+1+1+1,\ 1+1+1+1+1$. To each of these decompositions we assign a

Young diagram with h rows and k columns,

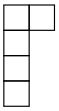
$S_5 : 5 = 5$  $[5]$

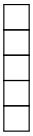
$= 4 + 1$  $[4, 1]$

$= 3 + 2$ $[3, 2]$

$= 3 + 1 + 1$ $[3, 1^2]$

$= 2 + 2 + 1$ $[2^2, 1]$

$= 2 + 1 + 1 + 1$  $[2, 1^3]$

$= 1 + 1 + 1 + 1 + 1$  $[1^5]$

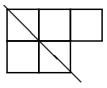
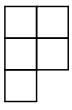
Each class or partition generates an irreducible representation of S_N. To each partition belongs an associated partition,

$$\text{Partition } \lambda \rightarrow \text{associated partition } \tilde{\lambda} \tag{B.19}$$

and an associated Young diagram,

$$\text{Young diagram } Y \rightarrow \text{ associated Young diagram } \tilde{Y} \tag{B.20}$$

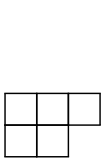
which is obtained by exchanging rows and columns, compare the following example with its corresponding axis of reflection,

$Y_\lambda \qquad \rightarrow \qquad \tilde{Y}_\lambda = Y_{\tilde{\lambda}}$

$\lambda = [3, 2] \qquad \tilde{\lambda} = [2^2, 1]$ (B.21)

In addition to the Young diagrams we define the Young tableaux by filling the Young diagrams Y with the numbers $1, 2, \ldots N$:



$\Longrightarrow$

θ_λ

$\rightarrow$

$\rightarrow$

$\tilde{\theta}_\lambda$ associated tableau

To each Young tableau we define the row symmetrizer

$$\mathcal{S}_\lambda = \frac{1}{\lambda_1! \cdots \lambda_h!} \sum_{\pi_\lambda} \pi_\lambda \tag{B.22}$$

and the antisymmetrizer

$$\mathcal{A}_\lambda = \frac{1}{\lambda_1! \cdots \lambda_h!} \sum_{\pi_\lambda} (-1)^{\pi_\lambda} \pi_\lambda. \tag{B.23}$$

Here π_λ permutes only elements within but not between the rows, i.e. each cycle in π_λ contains elements from only one row. As an example we consider the $\lambda = [3, 2]$ tableau[4]

. (B.24)

$\mathcal{S}_\lambda$ is given by

$$\begin{aligned} \mathcal{S}_\lambda = \frac{1}{3!\,2!} \big[& (1)(2)(3)(4)(5) + (12)(3)(4)(5) + (1)(23)(4)(5) \\ & + (13)(2)(4)(5) + (23)(1)(4)(5) \\ & + (123)(4)(5) + \text{ all with } (45) \big] \end{aligned} \tag{B.25}$$

It is $\mathcal{S}_{\lambda=[N]} = \mathcal{S}$, $\mathcal{A}_{\lambda=[N]} = \mathcal{A}$. The associated symmetrizers permute elements in the columns,

$$\begin{aligned} \tilde{\mathcal{S}}_\lambda &= \frac{1}{\tilde{\lambda}_1! \cdots \tilde{\lambda}_k!} \sum_{\tau_\lambda} \tau_\lambda, \\ \tilde{\mathcal{A}}_\lambda &= \frac{1}{\tilde{\lambda}_1! \cdots \tilde{\lambda}_k!} \sum_{\tau_\lambda} (-1)^{\tau_\lambda} \tau_\lambda; \end{aligned} \tag{B.26}$$

[4] For the $\lambda = [3, 2]$ tableau $\lambda_1 = 3$ and $\lambda_2 = 2$.

the permutations τ_λ leave the columns of the Young tableau θ_λ invariant, i.e. τ_λ permutes only elements within the columns. In the above example $\lambda = [3,2]$ becomes $\tilde{\lambda} = [2^2, 1]$ and

$$\tilde{\mathcal{S}}_\lambda = \frac{1}{2!\,2!\,1!}\big[(1)(2)(3)(4)(5) + (14)(2)(3)(5) + (14)(25)(3) + (1)(4)(25)(3)\big].$$

It is $\tilde{\mathcal{S}}_{\lambda=[1^N]} = \mathcal{S}$, $\tilde{\mathcal{A}}_{\lambda=[1^N]} = \mathcal{A}$.

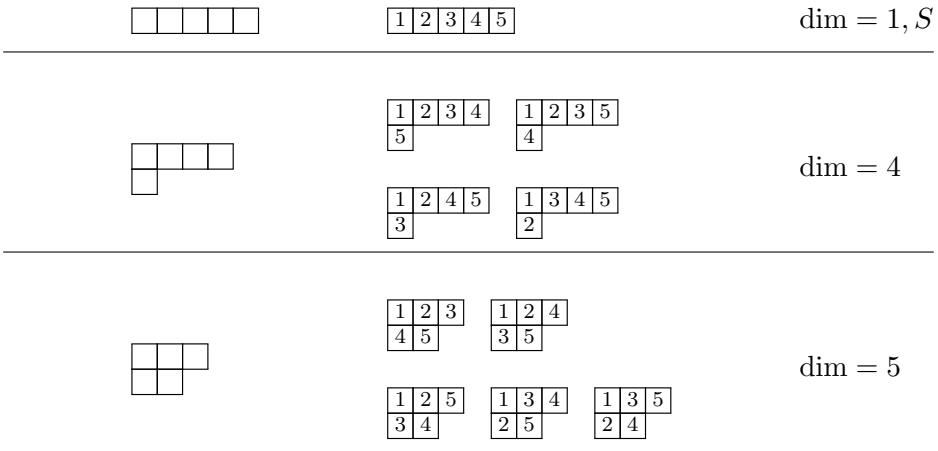
The irreducible symmetrizers

$$\begin{aligned} i_\lambda &= \mathcal{S}_\lambda \tilde{\mathcal{A}}_\lambda \quad \text{(irreducible symmetrizer)} \\ j_\lambda &= \tilde{\mathcal{A}}_\lambda \mathcal{S}_\lambda \quad \text{(irreducible antisymmetrizer)} \end{aligned} \tag{B.27}$$

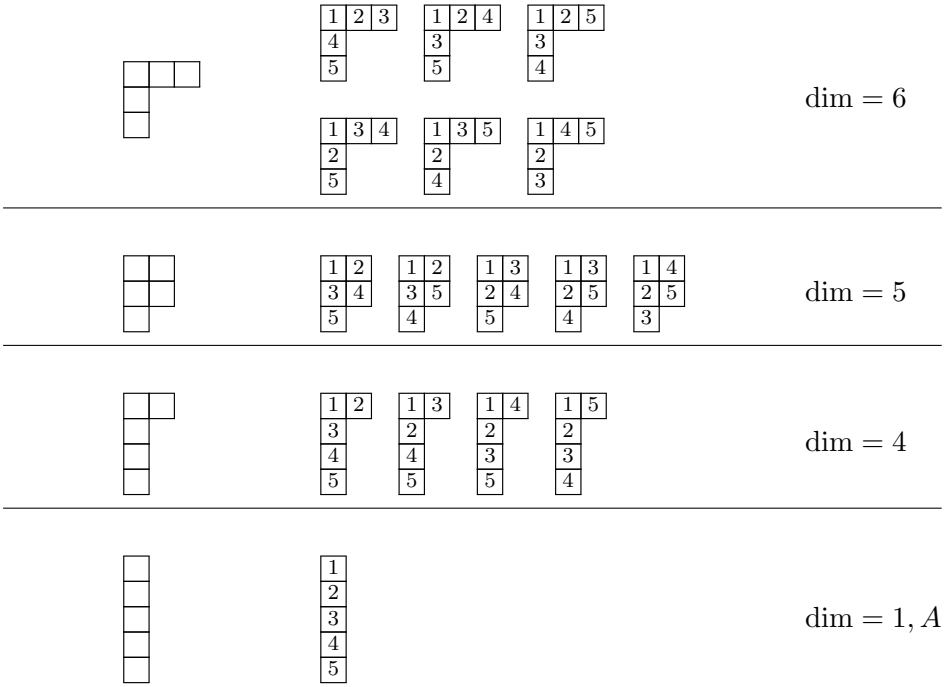
define the irreducible representations of S_N (i_λ and j_λ define equivalent representations). The operators i_λ and j_λ are the generalizations of $\mathcal{S}$ and $\mathcal{A}$, which define the one-dimensional representations S and A. It is[5]

$$\begin{aligned} i_{\lambda=[N]} &= \mathcal{S}_{[N]}\tilde{\mathcal{A}}_{[N]} = \mathcal{S}e = \mathcal{S}, \\ j_{\lambda=[N]} &= \tilde{\mathcal{A}}_{[N]}\mathcal{S}_{[N]} = e\mathcal{S} = \mathcal{S}, \\ i_{\lambda=[1^N]} &= \mathcal{S}_{[1^N]}\tilde{\mathcal{A}}_{[1^N]} = e\mathcal{A} = \mathcal{A}, \\ j_{\lambda=[1^N]} &= \tilde{\mathcal{A}}_{[1^N]}\mathcal{S}_{[1^N]} = \mathcal{A}e = \mathcal{A}. \end{aligned} \tag{B.28}$$

To find the dimension of the irreducible representation associated with Y we list all standard tableaux for Y. A standard tableau Y is a tableau in which the numbering of the cells is carried out so that in each column the numbers increase from top to bottom and in each row from left to right; we again consider S_5 as an example,



[5] e is the identity.



Here the first-mentioned representation is the so-called "normal-ordered tableau." As expected, we obtain the usual relation between the order of the group $N!$, the number of classes K, and the dimensions d_i of the irreducible representations, see (B.10),

$$N! = \sum_{i=1}^{K} d_i^2: \qquad 5! = 120 = 1 + 4^2 + 5^2 + 6^2 + 4^2 + 5^2 + 1. \tag{B.29}$$

Next we use the above structures to define spatial and spin wavefunctions with well-defined symmetry under permutations.

B.5 Position-space functions with defined symmetry

Let $u(x)$, $v(x)$, $w(x), \ldots$ be single-particle wavefunctions. We seek many-particle functions of the type[6]

$$\Psi(1, \cdots, N) = \sum \cdots u(\,)v(\,)w(\,)\cdots \tag{B.30}$$

with well-defined species, i.e. defined symmetry under permutation. We assume that the functions $u(x)$, $v(x)$, $w(x), \ldots$ are orthonormal. In what follows we consider the cases $N = 2$ and $N = 3$.

$N = 2$: Y is either $Y = \square\square$, i.e. S, or $Y = \begin{smallmatrix}\square\\\square\end{smallmatrix}$ and hence A. The basis vectors of the irreducible representation are obtained by applying the (anti-)symmetrizer to

[6] For the general case $u \neq v$, $v \neq w$, $u \neq w$, ...

the product state $\varphi(1,2) = u(1)\,v(2) = uv$[7] We suppress henceforth the arguments of the functions and understand that the first function takes argument 1, the second argument 2, and so on. Thus

$$\begin{aligned}
S\colon \boxed{1}\boxed{2} \qquad & i_{[2]} = \mathcal{S}_{[2]} = [e + (12)]/2, \qquad e = (1)(2) \\
& i_{[2]}\varphi = (uv + vu)/2 \equiv \varphi_{\mathcal{S}}/\sqrt{2} \\
A\colon \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline \end{array} \qquad & i_{[1^2]} = \tilde{\mathcal{A}}_{[1^2]} = [e - (12)]/2 \\
& i_{[1^2]}\varphi = (uv - vu)/2 \equiv \varphi_{\mathcal{A}}/\sqrt{2}.
\end{aligned} \tag{B.31}$$

$N = 3$: There are three possible arrangements of Young tableau Y: $Y =$ $\square\square\square$ (S), $Y =$ $\begin{array}{|c|}\hline \\ \hline \\ \hline \end{array}$ (A), and the third variant $Y =$ $\begin{array}{|c|c|}\hline \\ \hline \end{array}$ which we denote by Z.

$$\begin{aligned}
S\colon \boxed{1}\boxed{2}\boxed{3} \qquad & i_{[3]} = [e + (12) + (23) + (13) + (123) + (132)]/6 \\
& i_{[3]}\varphi = (uvw + vuw + uwv + wvu + wuv + vwu)/6 \\
& \equiv \varphi_{\mathcal{S}}/\sqrt{6}, \\
A\colon \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline 3 \\ \hline \end{array} \qquad & i_{[1^3]} = [e - (12) - (23) - (13) + (123) + (132)]/6 \\
& i_{[1^3]}\varphi = (uvw - vuw - uwv - wvu + wuv + vwu)/6 \\
& \equiv \varphi_{\mathcal{A}}/\sqrt{6}, \\
Z\colon \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} \end{array} \qquad & i_{[21]} = \mathcal{S}_{[21]}\tilde{\mathcal{A}}_{[21]} \\
& = \frac{1}{2!\,1!}[e + (12)]\frac{1}{2!\,1!}[e - (13)] \\
& = [e + (12) - (13) - (132)]/4 = e_{\mathcal{Z}}/4 \\
& i_{[21]}\varphi = (uvw + vuw - wvu - vwu)/4 \\
& \equiv \varphi_{\mathcal{Z}}/2, \\
\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1} \end{array} \qquad & i'_{[21]} = [e + (13) - (12) - (123)]/4 = e'_{\mathcal{Z}}/4 \\
& i'_{[21]}\varphi = (uvw + wvu - vuw - wuv)/4 \\
& \equiv \varphi'_{\mathcal{Z}}/2.
\end{aligned}$$

[7] In this chapter we let the operator act on the product function. In doing so it permutes the factors $u, v, \ldots$ with π.

Applying all group elements to $\varphi_{\mathcal{Z}}$ generates a two-dimensional representation of S_3; we write $4\, i_{[21]} = e_{\mathcal{Z}}$, then

$$\begin{aligned} e\, e_{\mathcal{Z}} &= e_{\mathcal{Z}}, \\ (12)e_{\mathcal{Z}} &= e_{\mathcal{Z}}, \\ (23)e_{\mathcal{Z}} &= (23) + (132) - (123) - (12) \equiv \overline{e}_{\mathcal{Z}}, \\ (31)e_{\mathcal{Z}} &= -e_{\mathcal{Z}} - \overline{e}_{\mathcal{Z}}, \\ (123)e_{\mathcal{Z}} &= -e_{\mathcal{Z}} - \overline{e}_{\mathcal{Z}}, \\ (132)e_{\mathcal{Z}} &= \overline{e}_{\mathcal{Z}}. \end{aligned} \tag{B.32}$$

Hence $e_{\mathcal{Z}} = e + (12) - (13) - (132)$ and $\overline{e}_{\mathcal{Z}} = (23) + (132) - (123) - (12)$ respectively $\varphi_{\mathcal{Z}} = e_{\mathcal{Z}}\varphi/2$ and $\overline{\varphi}_{\mathcal{Z}} = \overline{e}_{\mathcal{Z}}\varphi/2$ span a two-dimensional irreducible representation (we call it Z) of S_3. Likewise

$$\begin{aligned} e'_{\mathcal{Z}} &= e + (13) - (12) - (123), \\ \overline{e}'_{\mathcal{Z}} &= (23) + (123) - (132) - (23), \end{aligned} \tag{B.33}$$

spans a second two-dimensional irreducible representation Z' of S_3; with $(e'_{\mathcal{Z}} e_{\mathcal{Z}}) = 0$ this representation differs from Z. The reduction of the 6-dimensional representation space spanned by the product wave functions $u(\pi_1)v(\pi_2)w(\pi_3)$, $\pi \in S_3$ yields thus

$$S \oplus A \oplus Z \oplus Z'. \tag{B.34}$$

Each standard tableau defines a generating vector in each of these irreducible representations.

B.6 Spin functions with defined symmetry

We consider the spin problem for N spin-1/2 fermions. The states $\chi_\uparrow$ and $\chi_\downarrow$ span the Hilbert space $\mathcal{H}_{1/2}$ and the product functions

$$\chi_{\sigma_1} \otimes \chi_{\sigma_2} \otimes \cdots \chi_{\sigma_N} \tag{B.35}$$

define a basis in the 2^N-dimensional product space $(\mathcal{H}_{1/2})^{\otimes N}$. The operators S^2, S_z, $\pi \in S_N$, with $\boldsymbol{S} = \sum_{i=1}^{N} \boldsymbol{s}_i$ the total spin, commute pairwise and can therefore be simultaneously "diagonalized," i.e. we can simultaneously reduce the representations of $SU(2)$ (with quantum numbers S for $\boldsymbol{S}^2$ and M_S for S_z) and of S_N (with quantum numbers given by the standard tableaux).

***SU*(2):** $\mathcal{H}_{1/2}^{\otimes N} = \oplus_S \mu_S(N)\mathcal{H}_S$ according to the rules of spin addition. As examples we again consider the cases $N = 2, \ldots, 5$:

$$
\begin{aligned}
N=2\colon \quad & \mathcal{H}_{1/2}^{\otimes 2} = \mathcal{H}_0 \oplus \mathcal{H}_1, \\
& \mu_0(2) = \mu_1(2) = 1. \\
N=3\colon \quad & \mathcal{H}_{1/2}^{\otimes 3} = \Big(\mathcal{H}_0 \oplus \mathcal{H}_1\Big) \otimes \mathcal{H}_{1/2}, \\
& = 2\mathcal{H}_{1/2} \oplus \mathcal{H}_{3/2}, \\
& \mu_{1/2}(3) = 2, \quad \mu_{3/2}(3) = 1. \\
N=4\colon \quad & \mathcal{H}_{1/2}^{\otimes 4} = \Big(\mathcal{H}_{1/2} \oplus \mathcal{H}_{1/2} \oplus \mathcal{H}_{3/2}\Big) \otimes \mathcal{H}_{1/2}, \\
& = 2\mathcal{H}_0 \oplus 3\mathcal{H}_1 \oplus \mathcal{H}_2, \\
& \mu_0(4) = 2, \quad \mu_1(4) = 3, \quad \mu_2(4) = 1. \\
N=5\colon \quad & \mathcal{H}_{1/2}^{\otimes 5} : \mu_{1/2}(5) = 5 \quad \mu_{3/2}(5) = 4, \quad \mu_{5/2}(5) = 1.
\end{aligned}
\tag{B.36}
$$

In general

$$
\mu_S(N) = \frac{(2S+1)N!}{(N/2-S)!(N/2+S+1)!} \tag{B.37}
$$

and hence

$$
\sum_S \mu_S(N) \cdot (2S+1) = 2^N, \qquad \mu_{\frac{N}{2}}(N) = 1. \tag{B.38}
$$

***S*$_N$:** In the spin problem only two basis vectors $\chi_\uparrow$ and $\chi_\downarrow$ exist. Hence the Young tableaux can have at most two rows; otherwise the corresponding projector on the product basis vanishes. If, for example, there were three rows,

$$
\begin{array}{|c|}\hline \uparrow \\ \hline \downarrow \\ \hline \downarrow \\ \hline \end{array} = i_{[1^3]}\, \chi_\uparrow\chi_\downarrow\chi_\downarrow = \mathcal{A}\, \chi_\uparrow\chi_\downarrow\chi_\downarrow
$$

antisymmetrizing the two spin-$\downarrow$ states would yield 0. Furthermore, in every column the state $\uparrow$ must be paired with a state $\downarrow$, because the wave function likewise vanishes in this case. For instance, the operation i_Z applied to $Z = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} \end{array}$ acting on $\chi_\uparrow\chi_\downarrow\chi_\uparrow$ with 1 and 3 both in state $\uparrow$ yields

$$
\begin{aligned}
i_{[2,1]}\, \chi_\uparrow\chi_\downarrow\chi_\uparrow &= \frac{1}{4}[e + (12)][e - (13)]\chi_\uparrow\chi_\downarrow\chi_\uparrow \\
&= \frac{1}{4}[e + (12)](\chi_\uparrow\chi_\downarrow\chi_\uparrow - \chi_\uparrow\chi_\downarrow\chi_\uparrow) = 0.
\end{aligned}
\tag{B.39}
$$

We parameterize a general Young diagram with two rows

$$
\mu_S = \overbrace{\begin{array}{|c|c|c|c|}\hline & \cdots & & \\ \hline & \cdots & \\ \cline{1-3} \end{array}}^{N/2+S}
\underbrace{\qquad\qquad}_{N/2-S} \overbrace{\qquad\qquad}^{2S}
\tag{B.40}
$$

by the number

$$S = \begin{cases} 0, 1, \dots \frac{N}{2}, & \text{N even}, \\ \frac{1}{2}, \frac{3}{2}, \dots \frac{N}{2}, & \text{N odd}, \end{cases} \tag{B.41}$$

which later turns out to be the spin quantum number S. By elementary combinatorics one obtains the result that there are exactly $\mu_S(N)$ normal-ordered Young tableaux of species μ_S. In the following we prove that the representation space $\mathcal{H}_{1/2}^{\otimes N}$ contains the representation for the Young diagram μ_S exactly $(2S+1)$ times.

The space of fixed magnetic quantum number (eigenvalue of S_z)

$$M_S = \nu \cdot \frac{1}{2} - (N-\nu) \cdot \frac{1}{2} = \nu - \frac{N}{2} \in \left\{ -\frac{N}{2}, -\frac{N}{2}+1, \dots, \frac{N}{2} \right\}. \tag{B.42}$$

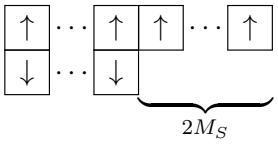
is spanned by product states with ν spin-$\uparrow$ and $(N-\nu)$ spin-$\downarrow$ states, $\nu \in \{0, \dots, N\}$. The vector space at fixed S_z is $N!/\nu!\,(N-\nu)!$-dimensional, since it is determined by specifying the particles with spin-$\uparrow$ (without regard to order). Within this space we may look for orthogonal subspaces of fixed $\boldsymbol{S}^2$. The allowed values of S are

$$S = \frac{N}{2}, \frac{N}{2} - 1, |M_S| + 1, |M_S|; \tag{B.43}$$

i.e. $N/2 - |M_S| + 1$ distinct possibilities.[8]

As we saw in the reduction $SU(2)$, the subspace of fixed S, M_S occurs exactly $\mu_S(N)$ times. It therefore defines a $\mu_S(N)$-dimensional representation of S_N, because a permutation $\pi \in S_N$ neither changes S nor M_S. Hence the subspace in $\mathcal{H}_{1/2}^{\otimes N}$ of fixed S, M_S generates a representation space of S_N. We have seen that $\mu_S(N)$ also equals the number of normal-ordered Young tableaux of species μ_S. Therefore the (correct) conjecture that the representation is irreducible is plausible, and we now proceed to prove it.

Let us first consider the case $S = M_S$. We require a wavefunction with maximal spin projection. For a fixed Young diagram of the form $\mu_S = [N/2 + S, N/2 - S]$ this is achieved by the following wavefunction

$\chi_{S,M_S} \propto$ 

(B.44)

. Applying permutations π from S_N to the wavefunction yields the $\mu_S(N)$-dimensional subspace for fixed S and M_S, which transforms according to the irreducible representation μ_S of S_N. By applying the operator S_-, which commutes with all $\pi \in S_N$, we obtain equivalent representations, namely exactly $2S+1$ copies of $M_S = -S, \dots, S$.

[8] It holds that $\sum_{S=|M_S|}^{N/2} \mu_S(N) = N!/\nu!(N-\nu)!$

Hence $\mathcal{H}_{1/2}^{\otimes N}$ decomposes as

$$\mathcal{H}_{1/2}^{\otimes N} = \bigoplus_S (2S+1)\mathcal{H}_{\mu_S}, \tag{B.45}$$

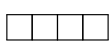
with $\mathcal{H}_{\mu_S}$ the irreducible representation corresponding to μ_S.
We verify the result on the usual examples $N = 2, \ldots, 5$,

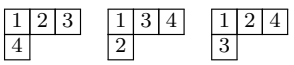
$N=2$ $\dim = 1$, $S = 1$, triplet $\begin{cases}\mathcal{S}|\uparrow\uparrow\rangle,\\ \mathcal{S}|\uparrow\downarrow\rangle,\\ \mathcal{S}|\downarrow\downarrow\rangle,\end{cases}$

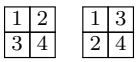
$\dim = 1$, $S = 0$, singlet $\{\mathcal{A}|\uparrow\downarrow\rangle$.

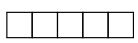
$N=3$ $\dim = 1$, $S = \frac{3}{2}$,

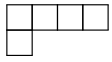
$\dim = 2 = \mu_{\frac{1}{2}}(3)$, $S = \frac{1}{2}$.

$N=4$  $\dim = 1 = \mu_2(4)$, $S = 2$,


$\dim = 3 = \mu_1(4)$, $S = 1$,


$\dim = 2 = \mu_0(4)$, $S = 0$.

$N=5$  $\mu_{\frac{5}{2}}(5) = 1$, $S = \frac{5}{2}$,

 $\mu_{\frac{3}{2}}(5) = 4$, $S = \frac{3}{2}$,

$\mu_{\frac{1}{2}}(5) = 5$, $S = \frac{1}{2}$.

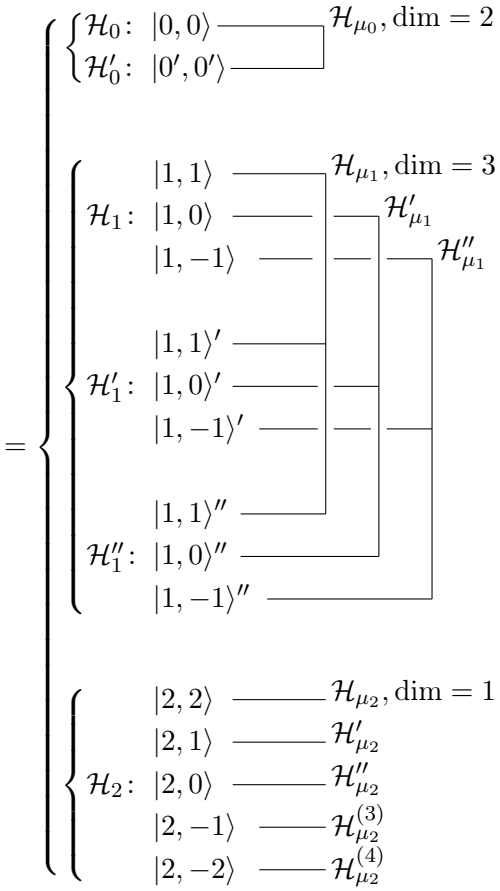
. Thus we have the decompositions (Weyl's theorem, Schur-Weyl duality)

$$\mathcal{H}_{1/2}^{\otimes N} = \bigoplus_S \mu_S(N)\mathcal{H}_S = \bigoplus_S (2S+1)\mathcal{H}_{\mu_S}; \tag{B.46}$$

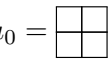
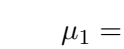
the spin S fixes the Young diagram μ_S and hence the spin symmetry. The representation space $\mathcal{H}_{\mu_S}$ is spanned by the $|S, M_S\rangle$ for fixed S and M_S.

As a final example, consider the case of 4 particles with spin 1/2,

$$\begin{aligned}\mathcal{H}_{1/2}^{\otimes 4} &\overset{SU(2)}{=} 2\mathcal{H}_0 \oplus 3\mathcal{H}_1 \oplus \mathcal{H}_2 \\ &\overset{S_4}{=} \mathcal{H}_{\mu_0} \oplus 3\mathcal{H}_{\mu_1} \oplus 5\mathcal{H}_{\mu_2}\end{aligned}$$

$=$ 

with

$\mu_0 =$, $\mu_1 =$  , $\mu_2 =$  ,

$\mu_0(4) = 2$, $\mu_1(4) = 3$, $\mu_2(4) = 1$.

B.7 Spin and Orbit

We consider a N-particle system with total spin S. The spin wavefunction has symmetry μ_S. Let the symmetry of the orbital wavefunction be given by the diagram μ_B. It holds: The tensor product $\mu_B \otimes \mu_S$ of the two representations contains the

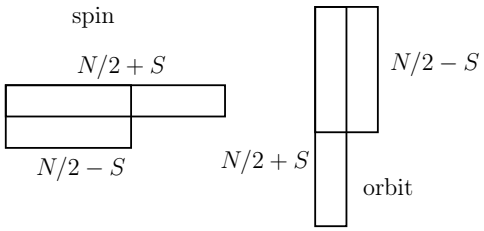


Figure B.1: Young tableaux associated with the spin and orbital wavefunctions for N spin-1/2 fermions with total spin S.

representation

$$\begin{aligned} &\text{(bosons)} \quad S \Leftrightarrow \mu_B = \mu_S \\ &\text{(fermions)} \quad A \Leftrightarrow \mu_B = \tilde{\mu}_S. \end{aligned} \tag{B.47}$$

It follows that the spin *species*, which is given by μ_S, determines the orbital *species* μ_B: For bosons, the orbital *species* corresponds to the spin *species*, $\mu_B = \mu_S$, whereas for fermions the orbital and spin *species* are associated with each other, $\mu_B = \tilde{\mu}_S$, see Figure B.1.

As an example, consider $N = 3$ electrons in the p shell for which we wish to find explicitly a wave function with $S = 1/2$. As we know, the spin fixes the spin

$$\mu_S = [21] = \begin{array}{|c|c|}\hline & \\ \hline \\ \cline{1-1} \end{array} \tag{B.48}$$

species. We also know that the state $M_S = 1/2$ corresponds to the Young tableau $\begin{array}{|c|c|}\hline \uparrow & \uparrow \\ \hline \downarrow \\ \cline{1-1} \end{array}$. Hence we obtain a wave function for $S = 1/2$, $M_S = 1/2$

$$\begin{aligned} \chi_{1/2,1/2}(s_1,s_2,s_3)/\sqrt{2} &= i_{[21]}\,\chi_\uparrow(s_1)\chi_\uparrow(s_2)\chi_\downarrow(s_3) = \mathcal{S}_{\mu_S}\tilde{\mathcal{A}}_{\mu_S}\chi_\uparrow(s_1)\chi_\uparrow(s_2)\chi_\downarrow(s_3) \\ &= \frac{1}{2}\mathcal{S}_{\mu_S}[\chi_\uparrow(s_1)\chi_\uparrow(s_2)\chi_\downarrow(s_3) - \chi_\uparrow(s_3)\chi_\uparrow(s_2)\chi_\downarrow(s_1)] \\ &= \frac{1}{4}[\chi_\uparrow(s_1)\chi_\uparrow(s_2)\chi_\downarrow(s_3) + \chi_\uparrow(s_2)\chi_\uparrow(s_1)\chi_\downarrow(s_3) \\ &\quad - \chi_\uparrow(s_3)\chi_\uparrow(s_2)\chi_\downarrow(s_1) - \chi_\uparrow(s_2)\chi_\uparrow(s_3)\chi_\downarrow(s_1)] \\ &= \frac{1}{2}[\chi_\uparrow(s_1)\chi_\uparrow(s_2)\chi_\downarrow(s_3) - \chi_\downarrow(s_1)\chi_\uparrow(s_2)\chi_\uparrow(s_3)] \end{aligned}$$

We know that the space at fixed S is exactly $[\mu_S(N) = 2]$-fold degenerate. A second basis state is obtained by applying the transposition $(s_2 \leftrightarrow s_3)$ to $\chi_{1/2,1/2}$, yielding

$$\chi'_{1/2,1/2}(s_1,s_2,s_3) = \frac{1}{\sqrt{2}}[\chi_\uparrow(s_1)\chi_\downarrow(s_2)\chi_\uparrow(s_3) - \chi_\downarrow(s_1)\chi_\uparrow(s_2)\chi_\uparrow(s_3)]; \tag{B.49}$$

The two states χ and χ' are linearly independent and therefore lie in the distinct representations Z and Z' of (B.34). Applying other elements of S_3 to χ, however, always gives linear combinations of χ and χ'. Of course, the spin wave functions for a given S and M_S can also be found with a Clebsch-Gordan table.

Next we wish to determine the possible orbital wave functions. We know that the orbital μ_B species must correspond to the associated diagram $\tilde{\mu}_S$. In our case we have

$$\mu_B = \tilde{\mu}_S = \mu_S = \begin{array}{|c|c|}\hline & \\ \hline \\ \cline{1-1} \end{array}\,. \tag{B.50}$$

In the p shell there are wave functions with $M_L = -1, 0, 1$. We denote the wave functions by φ_{M_L}. Since in any column of the Young tableau each quantum number may appear only once (see the discussion for the spin), the following states are allowed

$$\begin{array}{lcccccccc} & \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 0 \\ \cline{1-1}\end{array} & \begin{array}{|c|c|}\hline 1 & 1 \\ \hline -1 \\ \cline{1-1}\end{array} & \begin{array}{|c|c|}\hline 1 & 0 \\ \hline 0 \\ \cline{1-1}\end{array} & \begin{array}{|c|c|}\hline 1 & 0 \\ \hline -1 \\ \cline{1-1}\end{array} & \begin{array}{|c|c|}\hline 1 & -1 \\ \hline 0 \\ \cline{1-1}\end{array} & \begin{array}{|c|c|}\hline 1 & -1 \\ \hline -1 \\ \cline{1-1}\end{array} & \begin{array}{|c|c|}\hline 0 & 0 \\ \hline -1 \\ \cline{1-1}\end{array} & \begin{array}{|c|c|}\hline 0 & -1 \\ \hline -1 \\ \cline{1-1}\end{array} \\ M_L = & 2 & 1 & 1 & 0 & 0 & -1 & -1 & -2 \end{array}$$

8 possibilities (corresponding to $L = 1$ and $L = 2$). (B.51)

We consider only the case $M_L = 2$.[9] We must therefore symmetrize the orbital wave function $\varphi_1(\boldsymbol{r}_1)\varphi_0(\boldsymbol{r}_2)\varphi_1(\boldsymbol{r}_3)$ with the associated Young tableau $\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array}$ and obtain

$$\begin{aligned} \varphi_{2,2}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3) &= \frac{1}{\sqrt{2}} i'_{[2,1]} \varphi_1(\boldsymbol{r}_1)\varphi_0(\boldsymbol{r}_2)\varphi_1(\boldsymbol{r}_3) \\ &= \frac{1}{\sqrt{2}}[\varphi_1(\boldsymbol{r}_1)\varphi_0(\boldsymbol{r}_2)\varphi_1(\boldsymbol{r}_3) - \varphi_0(\boldsymbol{r}_1)\varphi_1(\boldsymbol{r}_2)\varphi_1(\boldsymbol{r}_3)]. \end{aligned} \tag{B.52}$$

We know that the product $\Psi = \varphi_{2,2}\chi_{1/2,1/2}$ must contain the antisymmetric representation. To obtain an antisymmetric wave function, we therefore need only antisymmetrize the wave function further

$$\begin{aligned} \Psi^{\mathcal{A}} &\propto \mathcal{A}\varphi_{2,2}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3)\chi_{1/2,1/2}(s_1, s_2, s_3) \\ &= \mathcal{A}\varphi_{1\uparrow}(x_1)\varphi_{0\uparrow}(x_2)\varphi_{1\downarrow}(x_3) - \mathcal{A}\varphi_{1\downarrow}(x_1)\varphi_{0\uparrow}(x_2)\varphi_{1\uparrow}(x_3) \\ &\quad - \mathcal{A}\varphi_{0\uparrow}(x_1)\varphi_{1\uparrow}(x_2)\varphi_{1\downarrow}(x_3) + \underbrace{\mathcal{A}\varphi_{0\downarrow}(x_1)\varphi_{1\uparrow}(x_2)\varphi_{1\uparrow}(x_3)}_{=0} \\ &= 3\mathcal{A}\varphi_{1\uparrow}(x_1)\varphi_{0\uparrow}(x_2)\varphi_{1\downarrow}(x_3) \end{aligned} \tag{B.53}$$

where $\varphi_{M_L M_S}(\boldsymbol{r}, s) = \varphi_{M_L}(\boldsymbol{r})\chi_{M_S}(s)$. The wave function $\Psi^{\mathcal{A}}$ is now the fermionic three-particle wave function for $M_S = S = 1/2$ and $M_L = L = 2$.

[9] Since M_L is maximal, we also know that this is a state of $L = 2$.

Appendix C

$SU(n)$ with Young Diagrams

We have seen in Appendix B that the representations of $SU(2)$ can be characterized by means of Young diagrams. To do so, we used N copies of $SU(2)$, so that in addition to the action of $SU(2)$ (on all copies simultaneously) the action of the symmetric group S_N is also defined. The representations of $SU(2)$ then follow from those of S_N (described by the Young diagrams with the help of Weyl's theorem). This procedure can be generalized to $SU(n)$. In this chapter we will understand aspects of the representation theory of $SU(n)$ and, in particular, $SU(3)$ with the aid of Young diagrams. As a physical application we finally derive the multiplet structure of the quark model (the "Eightfold Way"). We begin this chapter by formulating the results of the last section for the case of $SU(n)$ in a somewhat more general way.

C.1 $SU(n)$ tensors and the Schur-Weyl duality

To describe the action of the group $SU(n)$, we need an n-dimensional Hilbert space $\mathcal{H}$ with basis $|1\rangle, \ldots, |n\rangle$. Thus we have

$$U|j\rangle = \sum_{i=1}^{n} U_{ij}|i\rangle, \qquad U \in SU(n), \quad j = 1, \ldots, n\,. \tag{C.1}$$

This representation is called the defining (or fundamental) representation of $SU(n)$. For the case of $SU(2)$ the above representation is the spin-$1/2$ representation with basis vectors $|1\rangle = |\uparrow\rangle$ and $|2\rangle = |\downarrow\rangle$.

For systems N of identical particles, one considers the product space $\mathcal{H}^{\otimes N} = \mathcal{H} \otimes \cdots \otimes \mathcal{H}$. This transforms under the so-called product representation $U^{\otimes N}$ of $SU(n)$, i.e. each basis vector transforms with the defining representation and hence

$$U^{\otimes N} = \underbrace{U \otimes \cdots \otimes U}_{N \text{ times}}\,. \tag{C.2}$$

This product representation is usually reducible, and an important task in physics is now to ask which irreducible representations occur in $\mathcal{H}^{\otimes N}$. In the context of $SU(2)$

this question means which spins can occur in a N-particle system. More generally, one can use this procedure to obtain (all) irreducible representations of the Lie group $SU(n)$ with the help of (known) representation theory of S_N.

To this end we define the action of the symmetric group S_N on the product space $\mathcal{H}^{\otimes N}$ by[1]

$$\pi|j_1,\ldots,j_N\rangle = |\pi(j_1),\ldots,\pi(j_N)\rangle. \tag{C.3}$$

For what follows it is crucial that the actions of S_N and $SU(n)$ commute. Indeed we obtain

$$\begin{aligned}
\pi U^{\otimes N}\,|j_1,\ldots,j_N\rangle &= \sum_{i_1,\ldots,i_N} \pi\,|i_1,\ldots,i_N\rangle U_{i_1 j_1}\cdots U_{i_N j_N} \\
&= \sum_{i_1,\ldots,i_N} |\pi(i_1),\ldots,\pi(i_N)\rangle U_{i_1 j_1}\cdots U_{i_N j_N} \\
&= \sum_{i_1,\ldots,i_N} |\pi(i_1),\ldots,\pi(i_N)\rangle U_{\pi(i_1)\pi(j_1)}\cdots U_{\pi(i_N)\pi(j_N)} \\
&= U^{\otimes N}|\pi(j_1),\ldots,\pi(j_N)\rangle = U^{\otimes N}\pi\,|j_1,\ldots,j_N\rangle\,.
\end{aligned}$$

The decisive step occurs from the second to the third line, where the U matrices are reordered by exploiting the fact that π is a bijection. Thus on the product space $\mathcal{H}^{\otimes N}$ the representations of S_N can be reduced simultaneously with those of $SU(n)$.

From the last chapter we know that the irreducible representations of S_N are determined by Young diagrams λ. The corresponding basis vectors in $\mathcal{H}^{\otimes N}$ are generated by the irreducible symmetrizers i_λ. Here the number of standard tableaux fixes the dimension $\lambda(N)$ of the irreducible subspace. If one projects the vector space $\mathcal{H}^{\otimes N}$ onto a particular standard tableau, one obtains in general an d_λ-dimensional subspace with $d_\lambda = 0, 1, \ldots$. Each basis vector in this subspace now generates under S_N its own irreducible representation of species λ, see Chapter B.5. Hence d_λ does not depend on the chosen tableau, and we obtain the decomposition

$$\mathcal{H}^{\otimes N} = \bigoplus_\lambda d_\lambda \mathcal{H}_\lambda \tag{C.4}$$

into irreducible representations $\mathcal{H}_\lambda$ of S_N. Of course $\sum_\lambda d_\lambda \lambda(N) = \dim(\mathcal{H}^{\otimes N}) = n^N$. It is even possible to determine d_λ in closed form. After a longer derivation[2] one finds

$$d_\lambda = \frac{D(l_1,l_2,\ldots l_N)}{D(n-1,n-2,\ldots,0)} \quad . \quad D(l_1,l_2,\ldots,l_N) = \prod_{i<j}(l_i - l_j) \tag{C.5}$$

[1] By the right-hand side we denote the result after interchanging the factors, i.e. $(123)|u,v,w\rangle = |w,u,v\rangle$; in other words $\pi|j_1,\ldots,j_N\rangle = |j_{\pi^{-1}(1)},\ldots j_{\pi^{-1}(N)}\rangle$.

[2] See T. Inui, Y. Tanabe, Y. Onodera "Group Theory and Its Applications in Physics", Chapter 15.5.

with $l_j = \lambda_j + n - j$ and λ_j the length of the j-th row in λ.

Since π and $U^{\otimes N}$ commute, π cannot change the quantum numbers that characterize the irreducible representations of $SU(n)$. Hence the irreducible representations of S_N also carry fixed quantum numbers under $SU(n)$. This is exactly what we saw in $SU(2)$: the irreducible representations of S_N, described by μ_S, have a fixed S. Furthermore, $U^{\otimes N}$ does not change the species of a state, so that the representations of $SU(n)$ carry a fixed species. For $SU(2)$ this even means that the spin S uniquely determines the species $\lambda = \mu_S$ of the state. Thus the dimension $(2S+1)$ of the representation $\mathcal{H}_S$ under $SU(2)$ is precisely the multiplicity d_λ with which the species μ_S occurs. This unique correspondence between the irreducible representations extends to the groups $SU(n)$.

From (C.4) we know that the $\lambda(N)$-dimensional irreducible representation $\mathcal{H}_\lambda$ of S_N with species λ occurs d_λ times. Under the action of $SU(n)$ the species λ is assigned the irreducible representation $\mathcal{H}^\lambda$.[3] Since the action of S_N commutes with that of $SU(n)$, the irreducible representation $\mathcal{H}^\lambda$ occurs $\lambda(N)$ times. Thus we obtain the decomposition (Schur-Weyl duality)

$$\mathcal{H}^{\otimes N} = \underbrace{\bigoplus_\lambda \lambda(N)\, \mathcal{H}^\lambda}_{SU(n)} = \underbrace{\bigoplus_\lambda d_\lambda\, \mathcal{H}_\lambda}_{S_N}\,. \tag{C.6}$$

From this the dimension $\lambda(N)\dim(\mathcal{H}^\lambda) = d_\lambda \dim(\mathcal{H}_\lambda) = d_\lambda \lambda(N)$ follows. Hence the dimension $\dim(\mathcal{H}^\lambda) = d_\lambda$ of the irreducible representations of $SU(n)$ is fixed by (C.5). The representation theory of $SU(n)$ thereby follows from that of S_N.

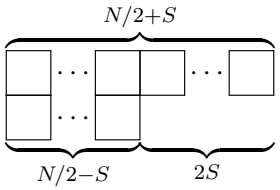
The connection between Young diagrams and the irreducible representations of $SU(n)$ is called Schur-Weyl duality. What is remarkable here is that one can determine the irreducible representations of $SU(n)$ without studying continuous Lie groups. Everything follows from the representation theory of the finite-dimensional symmetric group S_N. In the following sections we will discuss this in more detail using $SU(2)$ and $SU(3)$.

C.2 $SU(2)$ with Young diagrams

We first examine $SU(2)$ using the methods developed in the previous section, in order to prepare the analysis of $SU(3)$. The defining representation of $SU(2)$ is spin-1/2, so there are only two basis vectors $|\uparrow\rangle, |\downarrow\rangle$. Hence the Young diagrams can have at most two rows, since otherwise the antisymmetrizer always produces a zero. Thus for a fixed N each Young diagram is completely characterized by a single constant S

[3] It can be shown that this assignment is unique.

(the spin). As already shown in (B.40), we adopt the convention [4]

$\mu_S =$ 

. (C.7)

For fixed N, $2S$ thus describes the "overhang" of the Young diagram. The spin $S = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$ is directly related to the Casimir operator $\boldsymbol{S}^2$ of $SU(2)$, which determines the irreducible representations of $SU(2)$. A Casimir operator is an operator that commutes with all generators of the Lie group, here therefore with S_x, S_y, S_z. To label the states in the representations of $SU(2)$ one chooses, besides $\boldsymbol{S}^2$, a further operator, conventionally S_z, that uniquely fixes the various basis states in the representation. Here the basis states $|S, M_S\rangle$ in $\mathcal{H}^{\mu_S} \equiv \mathcal{H}_S$ carry the additional quantum number $M_S = -S, \ldots, S$. The dimension of the representation associated with S is obtained at once from equation (C.5)

$$\begin{aligned} d_{\mu_S} &= \frac{D(l_1, l_2)}{D(1,0)} = l_1 - l_2 = (\mu_1 + 2 - 1) - (\mu_2 + 2 - 2) \\ &= (\mu_1 - \mu_2) + 1 = 2S + 1. \end{aligned}$$

By elementary combinatorics (exactly as in B) one can now also determine $\mu_S(N) = \dim(\mathcal{H}_\lambda)$. In summary we recover the already familiar decomposition:

$$\mathcal{H}^{\otimes N} = \bigoplus_S \mu_S(N)\, \mathcal{H}_S = \bigoplus_S (2S+1)\, \mathcal{H}_{\mu_S} \tag{C.8}$$

As seen in the previous chapter, Schur-Weyl duality also allows one to determine basis vectors for the spin-S representation.

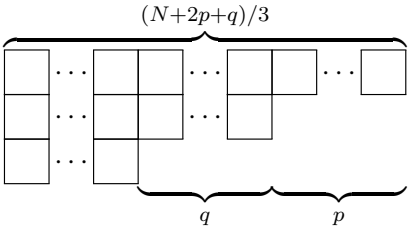
C.3 $SU(3)$ with Young diagrams

For the group $SU(3)$ the procedure is completely analogous. Physically this group, for example, describes the flavour symmetry of quarks. The fundamental representation of $SU(3)$ is three-dimensional and the basis vectors are denoted by $|u\rangle$ (up-quark), $|d\rangle$ (down-quark) and $|s\rangle$ (strange-quark).

Since there are only three distinct basis vectors, the Young diagrams can have at most three rows. In this case, for fixed N two numbers p, q are needed to specify a Young diagram uniquely. In the literature of Lie-group theory the tuple (p, q) is

[4] The constant must be chosen so that the result for d_λ is independent of N.

called the Dynkin indices. We define the species as

$$\lambda_{(p,q)} = (q,p) = $$ 

, (C.9)

where $p, q = 0, 1, 2, \ldots$ describe the lengths of the overhangs.

Exactly as for $SU(2)$, the numbers (p, q) are directly related to Casimir operators of the group. Hence the irreducible representations of $SU(3)$ are fixed by two Casimir operators, one for p and one for q. We denote this irreducible representation by $\mathcal{H}^{\lambda} = \mathcal{H}^{(p,q)}$. To label the states within a representation $\mathcal{H}^{(p,q)}$ one then chooses an arbitrary further element of the Lie algebra, thereby obtaining a complete set of quantum numbers.[5]

The dimension of the representation $\mathcal{H}^{(p,q)}$ is again determined using equation (C.5) (again $d_{(p,q)}$ does not depend on N)

$$\begin{aligned} d_{(p,q)} &= \frac{D(l_1,l_2,l_3)}{D(2,1,0)} = \frac{(l_1-l_2)\cdot(l_1-l_3)\cdot(l_2-l_3)}{(2-1)\cdot(2-0)\cdot(1-0)} \\ &= \frac{1}{2}(p+1)(q+1)(p+q+2). \end{aligned} \tag{C.10}$$

This central result in Lie group theory we obtain again directly from representation theory of S_N.

As a simple example we consider the fundamental representation of $SU(3)$, described by the diagram $\square$. This diagram belongs to $N = 1, (p,q) = (1,0)$ so that the dimension, using equation (C.10), is as expected $d_{(1,0)} = 3$.

For $N = 2$ there are now two possible diagrams: on the one hand $\square\square$ with $(p,q) = (2,0)$ and $d_{(2,0)} = 6$ and on the other $\begin{smallmatrix}\square\\\square\end{smallmatrix}$ with $(p,q) = (0,2)$ and $d_{(0,2)} = 6$. Just as for $SU(2)$ one can now determine basis vectors with Young symmetrizers that transform under an irreducible representation of $SU(3)$.

The extension of this procedure to $SU(n)$ is relatively straightforward. The Young diagrams of $SU(n)$ can have at most n rows, since the fundamental representation has exactly n basis vectors. For fixed N, one needs exactly $n-1$ constants to describe a Young diagram of $SU(n)$ completely. Hence $SU(n)$ has in total $(n-1)$ Casimir operators. A Young diagram is labeled by an index $(p_1, p_2, \ldots, p_{n-1})$, by which it is uniquely described. The constants $(p_1, p_2, \ldots, p_{n-1})$ must again be defined so that the dimension d_λ is independent of N, which concretely means that each constant describes the "overhang" in the diagram. Then the dimensions of the representations, as well as basis vectors, can be determined simply with the help of the Young diagrams.

[5] In the literature this is also called the Cartan subalgebra.

C.4 Conjugate representations

In addition to the ordinary irreducible representations, the conjugate irreducible representations $\overline{U} = U^*$ are also important in physics.[6] Here U^* denotes the complex conjugate matrix. Starting from an irreducible representation $\mathcal{H}_\lambda$, $\overline{\mathcal{H}_\lambda}$ is of course also irreducible. One is therefore immediately led to the question whether $\mathcal{H}_\lambda$ and $\mathcal{H}_{\overline{\lambda}} \equiv \overline{\mathcal{H}_\lambda}$ are the same irreducible representation.[7]

The conjugate representation $\overline{\lambda}$ can also be described simply with Young diagrams. In $SU(n)$, for fixed N a Young diagram λ is described by a tuple $\boldsymbol{p} = (p_1, \ldots, p_{n-1})$ of "overhangs." The conjugate representation then corresponds simply to the conjugate Young diagram $\overline{\lambda}$, which is given by $\overline{\boldsymbol{p}} = (p_{n-1}, p_{n-2}, \ldots, p_1)$, where possibly another N is needed to obtain an admissible Young diagram. Note that conjugation is an involution, i.e., $\overline{\overline{\lambda}} = \lambda$.

In the example of $SU(2)$ there is only one overhang. Hence the conjugate representation is equivalent to the ordinary representation. In fact, for the fundamental representation [8] of $SU(2)$,

$$U^* = \sigma_2 U \sigma_2, \qquad U \in SU(2)\,. \tag{C.11}$$

This nontrivial relation is also written as $\mathcal{H}_{1/2} \simeq \overline{\mathcal{H}_{1/2}}$. To show it, one can use the parametrization $U = \exp(i\boldsymbol{n}\cdot\boldsymbol{\sigma})$, $\boldsymbol{n} \in \mathbb{R}^3$, together with the property $\sigma_j^* = -\sigma_2\sigma_j\sigma_2$ of the Pauli matrices.

For $SU(3)$, however, the conjugate representation differs from the ordinary one. As an example consider the (1,0) and the (1,1) representations of $SU(3)$. We obtain

$\overline{(1,0)} =$ $=$ $= (0,1)$ (C.12)

$\overline{(1,1)} =$ $=$ $= (1,1)$ (C.13)

In general $\overline{(p,q)} = (q,p)$. The indices p and q are thus conjugate to each other.

While in $SU(2)$ the various representations can be identified solely by their dimension, this is not possible for $SU(3)$. In fact, different combinations of (p,q) yield the same dimension. With the concept of conjugate representations it is, however, possible to describe the physically relevant "small" representations (p,q) with $pq, \leq 3$ solely by their dimension $d_{(p,q)} = \dim \mathcal{H}^{(p,q)}$. In particle physics one often uses the convention

$$\mathcal{H}^{(p,q)} \equiv \begin{cases} \boldsymbol{d}_{(p,q)}, & p \geq q \\ \overline{\boldsymbol{d}_{(p,q)}}, & p < q\,. \end{cases} \tag{C.14}$$

[6] If $|j\rangle$ transforms with U, then $\langle j|$ transforms with U^*.

[7] Formaly this means that there exists a unitary transformation T such that $TUT^\dagger = U^*$ for all U holds.

[8] This suffices, since the "higher" representations are generated by taking tensor products of the fundamental representation.

That is, one writes $\mathcal{H}^{(1,0)} = \mathbf{3}$ and $\mathcal{H}^{(0,1)} = \overline{\mathbf{3}}$ instead of specifying the tuple (p, q). This notation is particularly useful when discussing the transformation of particles and antiparticles in product representations. In fact, the up, down, and strange quarks (by definition) transform under the fundamental representation **3**. Hence the corresponding antiparticles transform under $\overline{\mathbf{3}}$.

C.5 Product representations and the quark model

Young diagrams allow one to graphically reduce product representations. We wish to explain this in this chapter using the example of $SU(3)$ in order to better understand the quark model. Up to now we have always reduced tensor products of the fundamental representation of $SU(3)$. Graphically this means that we have decomposed tensor products of $\square$ into direct sums. For example, for $SU(3)$

$$3 \otimes 3 = \square \otimes \square = \square\square \oplus \begin{array}{|c|}\hline \\ \hline \\ \hline \end{array}. \tag{C.15}$$

This result is obtained directly from the following rule for the graphical tensor product of an irreducible representation with the fundamental representation (a single box):

> The additional box $\square$ is added in all positions, subject to the condition that an admissible Young diagram results. The result is the direct sum of all such diagrams.

A nice application of the methods introduced in the previous sections is found in the (static) quark model of particle physics. In the static quark model one distinguishes two types of particles: baryons (consisting of 3 quarks) and mesons (consisting of a quark and an antiquark). As already mentioned above, a quark transforms under the fundamental representation **3**, while an antiquark transforms under the conjugate representation $\overline{\mathbf{3}}$. This transformation behavior is understood in modern particle physics as a consequence of the flavor symmetry of the quarks. The strong interaction depends only on the color charge of the quarks and not on the flavor. Hence, if one replaces a quark by another while keeping the color the same, the mass of the hadron does not change significantly, provided the respective quark masses are much smaller than the interaction scale of the strong interaction. This is true for the light quarks (u, s, d), so that these can be exchanged among themselves without significantly changing the mass of the hadron. This exchange is described by the representation **3** and $\overline{\mathbf{3}}$. Historically, this symmetry was inferred from experimental data.

It now arises the question why stable states consist only of 3 quarks ($\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$) or a quark-antiquark pair ($\mathbf{3} \otimes \overline{\mathbf{3}}$) and not, for example, only of two quarks ($\mathbf{3} \otimes \mathbf{3}$). The reason for this lies in the strong interaction. Due to "color confinement," all matter must be colorless. That is, the wave function must be invariant (a singlet) under transformations in color space (r, g, b). Since the color charge is also described by the symmetry group $SU(3)$, the representations can likewise be described by Young diagrams. The only one-dimensional representation is obtained with $p = q = 0$, which

corresponds to the Young tableau $\begin{array}{|c|}\hline r \\ \hline g \\ \hline b \\ \hline \end{array}$. Hence one needs 3 quarks for a color singlet, which corresponds to the baryons. The only other way to produce a colorless wave function is the combination of color with anticolor, since $\square \otimes \overline{\square}$ contains a singlet. In fact, the state $|r\bar{r}\rangle + |g\bar{g}\rangle + |b\bar{b}\rangle$ forms a color singlet, which corresponds to the mesons.

The stable particles are now represented by basis vectors of irreducible representations in the tensor product spaces $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$ (baryons) and $\mathbf{3} \otimes \overline{\mathbf{3}}$ (mesons). Thus we must decompose the product representations into irreducible ones in order to find the corresponding particles. We use the graphical method for this.

We begin with the baryons:

$$\begin{aligned} \mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} &= \square \otimes \square \otimes \square = \left(\begin{array}{|c|}\hline \\ \hline \end{array} \oplus \square\square \right) \otimes \square \\ &= \left(\begin{array}{|c|}\hline \\ \hline \end{array} \otimes \square \right) \oplus \left(\square\square \otimes \square \right) \\ &= \begin{array}{|c|}\hline \\ \hline \\ \hline \end{array} \oplus \begin{array}{|c|c|}\hline \\ \hline \end{array} \oplus \begin{array}{|c|c|}\hline \\ \hline \end{array} \oplus \square\square\square \end{aligned} \tag{C.16}$$

In conventional notation one writes this result as

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}. \tag{C.17}$$

Here only **8** and **10** correspond to physical particles. The singlet state **1** is forbidden by the Pauli principle.

The Pauli principle requires that the total wave function (spin, flavor and color) be antisymmetric. As we have already seen, the color degree of freedom, because of color confinement, is a singlet and hence antisymmetric. Hence the spin-flavor wave function must be symmetric. Each quark has spin 1/2, so that the 3 quarks combine to total spin $S = \frac{1}{2}$ (with species $\begin{array}{|c|c|}\hline \\ \hline \end{array}$) or to total spin $S = \frac{3}{2}$ (with species $\square\square\square$). With (B.47) we know that the associated flavor wave functions must have the same species in order that the spin-flavor wave function as a whole has a symmetric representation. Hence we obtain only the particles **8** for spin-1/2 and **10** for spin-3/2.

The decomposition is shown graphically in Figure C.1. By symmetry we expect the particles in **8** (the light baryons) and in **10** (the heavy baryons), respectively, to have the same mass (up to the small mass differences of the quarks). This statement holds except for effects of the electroweak interaction. The flavour wave functions belonging to the individual particles can now be determined with Young tableaux. As an example we consider the Δ^+ particle of spin-3/2 composed of the quarks uud (strangeness 0, isospin 1/2). The wave function is given by $\boxed{u}\boxed{u}\boxed{d}$. In comparison, the proton (p) of spin-1/2 (with the same quark content) has the flavour wave function $\begin{array}{|c|c|}\hline u & u \\ \hline d \\ \cline{1-1} \end{array}$.

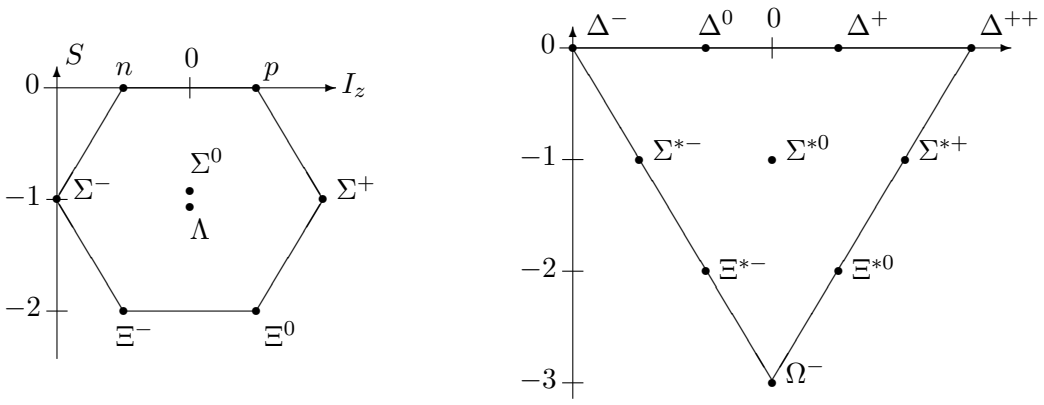


Figure C.1: Baryon octet **8** for spin-1/2 (left) and baryon decuplet **10** for spin-3/2 (right). The x axis in this graph denotes the isospin quantum number I_z (proportional to the number of up quarks minus the number of down quarks), while the y axis denotes strangeness S (respectively hypercharge Y). We will not pursue this characterization further here (for more information see S. Coleman "Aspects of Symmetry").

The procedure for the mesons is analogous. Here the decomposition

$$3 \otimes \overline{3} = \square \otimes \overline{\square} = \square \otimes \begin{array}{|c|}\hline\\\hline\end{array} = \begin{array}{|c|}\hline\\\hline\\\hline\end{array} \oplus \begin{array}{|c|c|}\hline\\\hline\end{array}. \tag{C.18}$$

holds. In conventional notation one writes this result as

$$\mathbf{3} \otimes \overline{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8}. \tag{C.19}$$

In contrast to the baryons, the singlet **1** is now allowed. It describes a η' meson. The reason is that mesons are composed of distinguishable constituents (quark and antiquark), so that the Pauli principle does not apply. We have collected the meson octet in Figure C.2. Again, the mesons in the octet have similar masses. With this we conclude our excursion into the world of representation theory. Representation theory is a field of great relevance to physics, of which we have shown only a small excerpt here. A good starting point for further study of this topic is the book "Group theory and its application to physical problems" by Morton Hamermesh, where you can also find proofs of the methods motivated here.

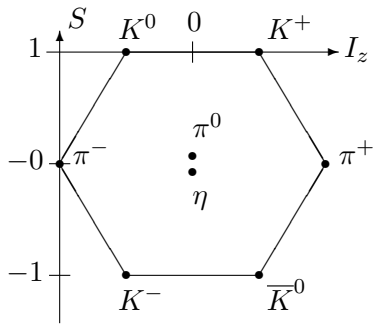


Figure C.2: meson octet **8**