

Quantum Mechanics II

Gianni Blatter

2020

Professor Gianni Blatter
Institute for Theoretical Physics
ETH Zurich

Warning: This document is originally based on a handwritten script in German. Errors may appear in translation and we appreciate much your help in their identification. Please send corrections and suggestions to

`blatterj@phys.ethz.ch`, Subject: Comment to QM-II-Script

future generation of students will be grateful for your help.

Thanks: I wish to thank Xueqing Chen, Miranda Davis, Chaoxin Ding, Anqi Gong, Beltran de Otalora Marcos Liberto, Beat Nairz, Yongxin Song, Jun Yamamoto, Runze Zhang, Yu Zhu, for their help in writing this English translation of the script and numerous students for their valuable comments.

Gianni Blatter

Contents

12 Entanglement and the EPR experiment	1
12.1 EPR Experiment	2
12.2 Bell Test and Bell Inequality	6
12.3 Short summary of facts, notions, and history	9
13 Identical Particles	11
13.1 Description of n Particles	12
13.2 Description of n identical particles	14
13.3 Symmetric Group S_n	16
13.4 Fermions and bosons	19
13.5 Spin & Statistics	20
13.6 Applications	20
13.7 Spin and race (german ‘Rasse’)	32
13.8 Formalism of spin and race* (german ‘Rasse’)	34
14 Atoms	47
14.1 Hartree & Hartree-Fock Approximation	49
14.2 Thomas-Fermi Screening	54
14.3 Thomas-Fermi Atom	57
14.4 Ground state configuration of atoms	63
14.5 Wigner-Eckart theorem	73
14.6 Spin-Orbit coupling in HF Theory*	79

14.7 Zeeman splitting of the fine structure in a magnetic field . . .	82
15 Molecules	85
15.1 Born-Oppenheimer approximation	86
15.2 Electronic problem	90
15.3 H ₂ molecule	94
15.4 The nuclear problem	98
15.5 Chemical bonds	99
16 Second Quantization	109
16.1 Problem in first-quantized form	110
16.2 Problem in second-quantized form	111
16.3 Second quantization: derivation*	115
16.4 Change of basis	124
16.5 Field operators	126
16.6 Equation of motion	133
17 Fermi sea: electrons	137
17.1 Fermi sea	138
17.2 Pair correlations	141
17.3 Ground state energy for interacting electrons	144
18 Phonons	149
18.1 One-dimensional chain/string	150
18.2 Quantization	154
18.3 Path integrals	161
19 Photons	171
19.1 Maxwell equations	171
19.2 Lagrange function	175
19.3 Quantization	180

19.4 Massive photons	189
20 Radiation and Matter	193
20.1 Consistent formulation	194
20.2 H-Atom in the radiation field	199
20.3 Light scattering	208
20.4 Lamb-shift	212
21 Quantum Statistical Mechanics	221
21.1 Expectation values	221
21.2 Ensembles	225
21.3 Concerning the third law of thermodynamics	226
22 Gases	227
23 Fermions	237
23.1 Dilute Fermi gas, $\delta = n\lambda^3 \ll 1$	239
23.2 Degenerate Fermi gas, $\delta = n\lambda^3 \gg 1$	240
24 Bosons	247
24.1 Photons	247
24.2 Phonons	252
24.3 Ideal Bose-gas: Bose-Einstein-condensation	255
25 Rel. particles, spin 0	263
25.1 Klein-Gordon equation	264
25.2 Pair generation at potential barriers	269
25.3 Two-component representation	275
25.4 Non-relativistic limit	278
25.5 Charge conjugation	279
25.6 Coulomb scattering and anti-particles	281

26 Rel. Spin $\frac{1}{2}$-Particles	289
26.1 Dirac equation	290
26.2 Charge conjugation	298
26.3 Non-relativistic limits	301
26.4 Covariance of the Dirac equation	308
27 Rel. massless Spin $\frac{1}{2}$-particles	317
27.1 Alternative derivation of the Dirac equation	320
27.2 Massless fermions in Weyl representation	322

Chapter 12

Entanglement and the EPR experiment

We have already seen that quantum mechanics brings some surprises to our classical world view; examples are the uncertainty principle, the quantization of orbits and energies, or the tunnel effect. The most spectacular new phenomenon, however, is the entanglement of degrees of freedom of two objects (and their generalization): two quantum mechanical degrees of freedom can be more strongly correlated than allowed in classical physics. The (experimentally provable) existence of such ‘strong’ correlations or entanglement shows that our world cannot be explained in a classical manner but is based (at least) on quantum mechanics.

The notion of entanglement, introduced in a work by Einstein, Podolsky, and Rosen (1935), further developed into a quantitative and experimentally testable statement by John Bell (1964), and experimentally tested by Freedman-Clauser (1972) and Aspect, Dalibard, and Roger (1982), was long-time an issue at the border between physics and philosophy (1935+). With the theoretical and experimental work of the 1960s and 1980s, the topic became more specific and even develops towards technological applications today: quantum cryptography and quantum computers are based on protocols and algorithms that use fundamental resources of quantum mechanics, the coherent superposition of states and the entanglement of separate degrees of freedom. The new research area of quantum information theory is developing rapidly and provides results that span the entire spectrum from fundamental to technological aspects. It is expected that these topics will be conquered by the engineering domain in a few years.

The simplest entanglement of objects can be formulated on the example of two spins (alternatively, the polarization of photons can be used; nowadays, spins/photons are often replaced by artificial two-level systems, so-called qubits): the spin singlet state

$$\begin{aligned} |0, 0\rangle &= \frac{1}{\sqrt{2}}[|\uparrow\rangle_1 \otimes |\downarrow\rangle_2 - |\downarrow\rangle_1 \otimes |\uparrow\rangle_2] \\ &= \frac{1}{\sqrt{2}}[|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle] \end{aligned} \quad (12.1)$$

is maximally entangled (other maximally entangled states are the Bell states $\frac{1}{\sqrt{2}}[|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle]$ and $\frac{1}{\sqrt{2}}[|\uparrow, \uparrow\rangle \pm |\downarrow, \downarrow\rangle]$). In contrast, the product state $|\pi\rangle = (a|\uparrow\rangle_1 + b|\downarrow\rangle_1) \otimes (c|\uparrow\rangle_2 + d|\downarrow\rangle_2)$ is not entangled. The original article by Einstein, Podolsky, and Rosen (EPR) deals with position and momentum variables instead of spin variables, the formulation in terms of spins originates from David Bohm. This formulation is particularly attractive today with a view to the new research field of ‘quantum engineering’: a quantum bit (qubit), the elementary unit of a quantum computer, can be understood as a system with two states (or two levels) and thus as a spin-1/2 system. While a classical bit has the values 1 or 0, a quantum bit can be in a superposition state $[a|0\rangle + \sqrt{1-a^2}e^{i\phi}|1\rangle] \leftrightarrow [a|\uparrow\rangle + \sqrt{1-a^2}e^{i\phi}|\downarrow\rangle]$; what we can learn about spin-1/2 systems also applies to qubits.

In the following, we analyze the properties of the state (12.1) (in an EPR experiment) and formalize the findings in a classical world view (local realism). We extend the simple EPR experiment to a Bell test with four detector settings and derive the Bell inequality for a local theory. We start with a modern language/understanding and at the end place our knowledge within the context of the historical discussion.

12.1 EPR Experiment

We make use of the established language of quantum information theory, where Alice and Bob share an EPR pair; we imagine that a machine generates spin-singlet pairs of the form (12.1), of which the first spin (1) goes to Alice, the second (2) to Bob, cf. Fig. 12.1. Natural sources of EPR pairs are nonlinear crystals in which a photon (frequency ω) is converted into two entangled photons ($\omega/2$, through a so-called down-conversion) or a pion that decays into an electron-positron pair. Today, entangled two-qubit states can be generated in a controlled manner through a carefully designed interaction between the qubits. Alice and Bob each have a measuring device with

which they detect their spins; furthermore, they can set the (spin/photon-)polarization directions \vec{a} (Alice) and \vec{b} (Bob) of their measuring devices; the measuring devices deliver results ± 1 if the spin is detected parallel/anti-parallel to the selected axis.

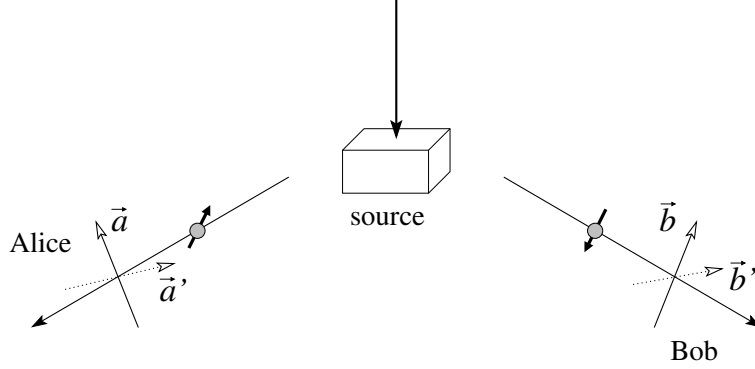


Fig. 12.1: EPR experiment with spins/photons/electron-positron pairs: a source generates a spin- or electron-positron pair in an entanglement state. Alice and Bob measure the polarization of these particles after they have passed through a polarization filter directed along \vec{a} or \vec{a}' (in the case of Alice) and \vec{b} or \vec{b}' (for Bob). In local realism, these multi-particle states will be described through a multivariate distribution function $P(\sigma_a, \sigma_b, \sigma_{a'}, \sigma_{b'})$.

Let us first consider the situation where Alice and Bob both measure their spin along the \vec{z} axis. When comparing their measurement results, we find Alice's result is randomly distributed with respect to the value ± 1 , which in turn is perfectly anti-correlated with Bob's result: every measurement $+1$ (-1) by Alice corresponds to a measurement -1 ($+1$) by Bob. This result, we can understand as follows: when Alice's measurement delivers the result $1 \leftrightarrow \uparrow$, then the state $|0, 0\rangle = \frac{1}{\sqrt{2}}[|\uparrow, \downarrow\rangle_z - |\downarrow, \uparrow\rangle_z]$ is projected to the state $|\uparrow, \downarrow\rangle_z$ and Bob's measurement (immediately after Alice's) produces the result $-1 \leftrightarrow \downarrow$. On the other hand, Alice's measurement of -1 will collapse the wave function to $|\downarrow, \uparrow\rangle_z$ and Bob thus measures $+1 \leftrightarrow \uparrow$. Finally, since the state $|0, 0\rangle$ is equally weighted between \downarrow and \uparrow , Alice's measurements will be randomly distributed between ± 1 .

Next, Alice measures along \vec{z} and Bob measures along $-\vec{z}$: now their results are perfectly correlated. In the third type of experiment, Alice measure along \vec{x} and Bob along $-\vec{x}$. Since we can also write the singlet state in the

basis of the quantization axis \vec{x} ,

$$|0, 0\rangle = \frac{1}{\sqrt{2}}[|\uparrow, \downarrow\rangle_x - |\downarrow, \uparrow\rangle_x], \quad (12.2)$$

we obtain the same (maximally correlated) result.

The situation becomes more interesting when Alice measures along \vec{z} and Bob along \vec{x} . We can write the singlet state in the form (using $|\uparrow\rangle_z = (|\uparrow\rangle_x + |\downarrow\rangle_x)/\sqrt{2}$, etc.)

$$|0, 0\rangle = \frac{1}{\sqrt{2}} \left[|\uparrow\rangle_z \otimes \frac{|\uparrow\rangle_x + |\downarrow\rangle_x}{\sqrt{2}} - |\downarrow\rangle_z \otimes \frac{|\uparrow\rangle_x - |\downarrow\rangle_x}{\sqrt{2}} \right]. \quad (12.3)$$

From this, we read that if Alice measures the value 1 or -1 then in both cases Bob will measure a result ± 1 with equal probability, i.e., the measurements of Alice and Bob are now not correlated at all. The results predicted (and experimentally verifiable, e.g., using quantum technology) are therefore ± 1 randomly distributed both for Alice and for Bob, with no correlations between them.

Hence, depending on the detector settings, the result of experiments can range from correlated (parallel axes), via anti-correlated (anti-parallel axes), to uncorrelated (perpendicular axes). Furthermore, these outcomes are independent of the amount of spatial separation between the two measurements/spins—Alice and Bob can travel to other galaxies with their spins before performing their measurements, and yet quantum mechanics still fixes the state of the second spin (with suitably chosen axes) immediately after the first measurement has been done. This thought tells us, that the *information* gained in the first measurement cannot affect the second spin, since information does not travel faster than the speed of light.

If we try to understand these results in a classic framework, we are forced to assume that the values of the spins are determined in advance regardless of the detector settings, while still being randomly distributed. The associated random process (e.g., in the source) must then generate the observed correlations. Such a random process is described by a multivariate distribution (joint probability distribution). In order to describe the above data (of three experiments with axes (\vec{z}, \vec{z}) , $(\vec{z}, -\vec{z})$, and (\vec{z}, \vec{x})) as a whole, this probability function must describe the results of all the various detector settings in one go. Below, we will analyze an experiment with four polarization settings, the so-called Bell test. In this experiment, we consider two axes for each experimenter, (\vec{a}, \vec{a}') for Alice and (\vec{b}, \vec{b}') for Bob, resulting in four

possible pairings (\vec{a}, \vec{b}) , (\vec{a}, \vec{b}') , (\vec{a}', \vec{b}) , and (\vec{a}', \vec{b}') ; such a situation should then be described by a multivariate distribution function (we drop vector symbols and other indices when not needed)

$$P(\sigma_{\vec{a}}^A, \sigma_{\vec{b}}^B, \sigma_{\vec{a}'}^A, \sigma_{\vec{b}'}^B) = P(\sigma_{\vec{a}}, \sigma_{\vec{b}}, \sigma_{\vec{a}'}, \sigma_{\vec{b}'}). \quad (12.4)$$

The object (12.4) should be thought of as a $2 \times 2 \times 2 \times 2$ table or as a list of 16 numbers. The existence of such a distribution function is an expression of local realism, the basis of our classical understanding of the world (following Einstein, ‘the moon there, also when no one is looking’), here, formulated as the assumption that the values σ_a , σ_b , $\sigma_{a'}$, and $\sigma_{b'}$ exist before they are measured and regardless of whether they are measured in the future. The adjective ‘local’ assumes that Alice’s results do not depend on the settings of Bob’s detector and vice versa.¹ In the second expression in (12.4), we have greatly simplified the notation by having one index define both the observer and his detector setting. The correlations K measured (and computable) by Alice and Bob then are given by

$$K(\vec{a}, \vec{b}) = \sum_{\sigma_a, \sigma_b} P_{ab}(\sigma_a, \sigma_b) \sigma_a \sigma_b \quad (12.6)$$

that involves the reduced probabilities

$$P_{ab}(\sigma_a, \sigma_b) = \sum_{\sigma_{a'}, \sigma_{b'} \in \{\pm 1\}} P(\sigma_a, \sigma_b, \sigma_{a'}, \sigma_{b'}) \quad (12.7)$$

associated with the specific setting (\vec{a}, \vec{b}) out of the four combinations (\vec{a}, \vec{b}) , (\vec{a}, \vec{b}') , (\vec{a}', \vec{b}) , and (\vec{a}', \vec{b}') . Note that, while in an experiment (\vec{a}, \vec{b}) only the reduced probabilities $P_{ab}(\sigma_a, \sigma_b)$ are relevant (the object $P_{ab}(\sigma_a, \sigma_b)$ corresponds to a 2×2 table), the definition of physical realism requires the existence of the full distribution function involving all four detector settings.

The question arises whether there is a classical multivariate distribution function of the form (12.4), which represents the result of an experiment that deeply roots in quantum mechanics, or formulated differently, is there

¹If we ignore the condition of locality, the description involves a more complex formalism with a distribution function

$$P(\sigma_{\vec{a}, \vec{b}}^A, \sigma_{\vec{a}, \vec{b}}^B; \sigma_{\vec{a}, \vec{b}'}^A, \sigma_{\vec{a}, \vec{b}'}^B; \sigma_{\vec{a}', \vec{b}}^A, \sigma_{\vec{a}', \vec{b}}^B; \sigma_{\vec{a}', \vec{b}'}^A, \sigma_{\vec{a}', \vec{b}'}^B) \quad (12.5)$$

that involves 8 variables. The random space then grows from 16 to 64 elements. Of the $4 \times 4 = 16$ reduced probabilities only 4 are physical (instead of the $2 \times 2 = 4$ reduced probabilities in the local theory, which all are physically relevant).

a real local physical world according to our classical view in the sense that all variables have a definitive value even if we do not measure them? The answer to this problem is provided by the Bell test.

12.2 Bell Test and Bell Inequality

We consider an EPR experiment with four settings of the polarizers (\vec{a}, \vec{b}) , (\vec{a}, \vec{b}') , (\vec{a}', \vec{b}) , and (\vec{a}', \vec{b}') . The measured values ($\pm \frac{\hbar}{2}$ for spins, polarization '+' or '-' for photons, here normalized to 1 or -1) are (σ_a, σ_b) , $(\sigma_a, \sigma_{b'})$, $(\sigma_{a'}, \sigma_b)$, $(\sigma_{a'}, \sigma_{b'})$. In a physical real and local world-view, these measurement data are described by the multivariate distribution function (12.4). We combine the outcome of the measurements into the expression

$$\sigma_a \sigma_b + \sigma_a \sigma_{b'} + \sigma_{a'} \sigma_b - \sigma_{a'} \sigma_{b'} = \pm 2. \quad (12.8)$$

The right hand side in (12.8) derives from rewriting the expression in the form²

$$\sigma_a \sigma_b + \sigma_a \sigma_{b'} + \sigma_{a'} \sigma_b - \sigma_{a'} \sigma_{b'} = \sigma_a (\sigma_b + \sigma_{b'}) + \sigma_{a'} (\sigma_b - \sigma_{b'}) \quad (12.10)$$

and noting, that either the sum $\sigma_b + \sigma_{b'} = 0$ or the difference $\sigma_b - \sigma_{b'} = 0$ disappears (since both measured values can only be 1 or -1), hence, one of the terms in (12.10) vanishes while the other term assumes one of the values ± 2 .

Next, we average the combination (12.8) over the multipartite distribution function (12.4) to arrive at the Bell inequality (note that while each individual term involves only one of four reduced probability distributions of the type (12.7), the entire expression is determined by the full distribution function (12.4); you *cannot* evaluate the average of (12.8) out of the four reduced probabilities alone)

$$\mathcal{B} = |K(\vec{a}, \vec{b}) + K(\vec{a}, \vec{b}') + K(\vec{a}', \vec{b}) - K(\vec{a}', \vec{b}')| \leq 2 = \mathcal{B}_{l-R} \quad (12.11)$$

involving four correlators of the type (12.6). This inequality (a simplification of the original Bell inequality, which is easier to access experimentally) originates from the work by Clauser, Horne, Shimony, and Holt (1969) and

²In a non-local theory, the expression has the form

$$\sigma_{\vec{a}, \vec{b}}^A \sigma_{\vec{a}, \vec{b}}^B + \sigma_{\vec{a}, \vec{b}'}^A \sigma_{\vec{a}, \vec{b}'}^B + \sigma_{\vec{a}', \vec{b}}^A \sigma_{\vec{a}', \vec{b}}^B - \sigma_{\vec{a}', \vec{b}'}^A \sigma_{\vec{a}', \vec{b}'}^B \quad (12.9)$$

and such a rewriting is no longer possible.

can be experimentally checked. Basically, the statement (12.11) is a statistical one: every multivariate local distribution function of the type (12.4) generates correlators that observe the restriction (12.11). The assumption of local physical realism means that a multi-variable distribution function (12.4) exists, which describes the result of an EPR experiment (with the settings (\vec{a}, \vec{b}) , (\vec{a}, \vec{b}') , (\vec{a}', \vec{b}) and (\vec{a}', \vec{b}') of the detector variables).

The question arises as to whether quantum mechanics satisfies the inequality (12.11) and is therefore compatible with the classical view of local realism—this is *not* the case. Calculating the correlator K for the four settings using the rules of quantum mechanics (instead of using a local multivariate distribution function) and maximizing the expression over all possible settings of the polarizers \vec{a} and \vec{b} , we find settings such that

$$\max_{\vec{a}, \vec{b}, \vec{a}', \vec{b}'} [|K(\vec{a}, \vec{b}) + K(\vec{a}, \vec{b}') + K(\vec{a}', \vec{b}) - K(\vec{a}', \vec{b}')|] = 2\sqrt{2} = \mathcal{B}_{QM} \quad (12.12)$$

The \mathcal{B}_{QM} value is the Tsirelson limit (Tsirelson / Cirel'son, 1980). Thus quantum mechanics does not follow the principle of physical realism: the values of an observable can be undetermined and only manifest when they are measured.³

To prove the statement (12.12), we calculate the correlators $K(\vec{a}, \vec{b})$ according to the rules of quantum mechanics (instead of classically—statistically via (12.6)),

$$K(\vec{a}, \vec{b}) = \langle (\vec{a} \cdot \vec{\sigma}_1)(\vec{b} \cdot \vec{\sigma}_2) \rangle \quad (12.13)$$

with $\vec{\sigma}_i$, $i = 1, 2$, the (vector of) Pauli matrices of the two spins (since quantum mechanics is linear, we can calculate the four terms individually in the Bell test). We use the fact that for the singlet state $|0, 0\rangle$ and each direction \vec{n} , we have (choosing the direction \vec{n} as the quantization axis of the spins)

$$\vec{n} \cdot (\vec{\sigma}_1 + \vec{\sigma}_2)|0, 0\rangle = 0, \quad (12.14)$$

$$\langle 0, 0 | \vec{n} \cdot \vec{\sigma}_1 | 0, 0 \rangle = 0. \quad (12.15)$$

The correlator then simplifies to

$$K(\vec{a}, \vec{b}) = \langle 0, 0 | (\vec{a} \cdot \vec{\sigma}_1)(\vec{b} \cdot \vec{\sigma}_2) | 0, 0 \rangle = -\langle 0, 0 | (\vec{a} \cdot \vec{\sigma}_1)(\vec{b} \cdot \vec{\sigma}_1) | 0, 0 \rangle$$

³If we omit the condition of locality, it can be shown that there are partial distributions that produce a Bell parameter $\mathcal{B}_{nl-R} = \mathcal{B}_{max} = 4$, the maximum possible value of the expression (12.9), even if the relativistic causality requires that Alice's measurement must not depend on Bob's attitude, $P_{ab}(\pm 1, ?) = P_{ab}(\pm 1, 1) + P_{ab}(\pm 1, -1) = P_{ab'}(\pm 1, ?)$ (Popescu, Rohrlich, 1994). Hence, it makes no sense to test quantum mechanics with a Bell test against non-local realism, since the latter does not provide any restriction.

$$= -\langle 0, 0 | \vec{a} \cdot \vec{b} + i(\vec{a} \wedge \vec{b} \cdot \vec{\sigma}_1 | 0, 0 \rangle = -\vec{a} \cdot \vec{b}, \quad (12.16)$$

where we used the results (12.14) and (12.15) and the relation, $(\vec{a} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma}) = (\vec{a} \cdot \vec{b} + i(\vec{a} \wedge \vec{b}) \cdot \vec{\sigma})$. The Bell parameter (12.11) can then be written as

$$\mathbf{B} = |\vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{b}' + \vec{a}' \cdot \vec{b} - \vec{a}' \cdot \vec{b}'| \quad (12.17)$$

If we choose the polarization as shown in Figure 12.2, we find with $\vec{a} \cdot \vec{b} = \vec{a} \cdot \vec{b}' = \vec{a}' \cdot \vec{b} = \cos \frac{\pi}{4} = \frac{1}{\sqrt{2}}$ and $\vec{a}' \cdot \vec{b}' = \cos \frac{3\pi}{4} = -\frac{1}{\sqrt{2}}$ the value $\mathbf{B}_{QM} = \frac{4}{\sqrt{2}} = 2\sqrt{2}$. Therefore, quantum mechanical degrees of freedom can be more strongly correlated than classically allowed. Accordingly, quantum mechanics cannot be interpreted as a local realistic theory and there is no distribution function $P(\sigma_a, \sigma_b, \sigma_{a'}, \sigma_{b'})$ that can explain the experimental outcome. The deeper reason for the non-existence of this distribution function lies in the Heisenberg uncertainty of non-commuting variables: The axes \vec{a} and \vec{a}' occurring in the Bell test do not allow simultaneous determination of the results σ_a and $\sigma_{a'}$ (the corresponding spin operators do not commute), which is why we cannot write down a distribution function where these variables are simultaneously and sharply defined.

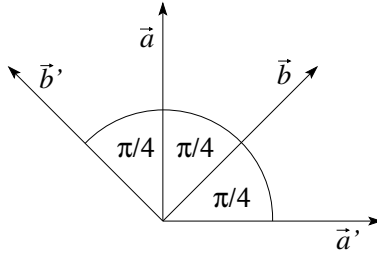


Fig. 12.2: Angles between the settings \vec{a} , \vec{b} , \vec{a}' , \vec{b}' which produce the maximum Bell parameter $\mathcal{B}_{QM} = 2\sqrt{2}$ (Tsirelson limit).

The further study of the phenomenon of 'entanglement' is the subject of quantum information theory; here are some comments. We have seen that classical physics generates the limit $\mathcal{B}_{l-R} = 2$ for the correlations of a Bell test within the framework of local realism, while quantum mechanics raises this value to $\mathcal{B}_{QM} = 2\sqrt{2}$. However, the maximum possible value for the Bell parameter (achieved in a non-local but relativistic causal theory) is $\mathcal{B}_{nl-R} = 4$ and the question arises as to the meaning of the gap $[2\sqrt{2}, 4]$ between the quantum mechanical and the maximum value of \mathcal{B} . It turns out that theories characterized by Bell parameters $\mathcal{B} > 2\sqrt{2}$ have quite astonishing properties. For example, the communication complexity can become trivial: in a distributed calculation (where the data is distributed between two parties), the necessary communication can be reduced to one

bit if so-called non-local boxes are available (Brassard, Buhrman, Linden, Method, Tapp, Unger (2006)). Similarly, due to the expanded informational causality, Bob would have had extended access to Alice's data (Pawlowski, Paterek, Kaszlikowski, Scarani, Winter, Zukowski (2009)).

12.3 Short summary of facts, notions, and history

Facts:

Quantum mechanics is not deterministic (hence, only makes statements about probabilities), quantum mechanics is non-local (in the sense of classical physical realism) but relativistic causal, quantum correlations can be stronger than classical ones, although there is no information transmission faster than light (no-signaling). Theories with correlations going beyond the Tsirelson limit (more strongly non-local than quantum mechanics) have expanded capacities in information processing. Quantum mechanics does not allow quantum information to be copied (No-Cloning Theorem, Wootters and Zurek, 1982). With the help of entangled states, quantum mechanics allows the teleportation of quantum states (Peres, Wootters, Brassard, Bennett, Josza, Crepeau, 1993).

Notions:

Realism: Physical variables are defined, even if they have not been measured. There is a distribution function that predicts the distribution of the values of variables.

Local theory: Reality in one place is not affected by reality in another place.

Local realism: There is a distribution function with local variables, e.g., $\sigma_{\vec{a}}$, which predicts the distribution of the values of the variables.

Non-local realism: the variables are not local, e.g., $\sigma_{\vec{a}, \vec{b}}$. This increases the probability space, e.g., from 16 to 64 possible events in the case of the Bell test.

Relativistic causality: introduces a condition to the distribution function which guarantees that the distribution function for Alice does not depend on Bob's distribution function. A local theory is relativistic causal.

History:

At the beginning of the last century, the philosophical-ideological implications of quantum mechanics were the subject of intense debate. The central

element in this discussion was the 1935 work by Einstein, Podolsky, and Rosen, in which they presented a thought experiment on the correlations of two particles that exemplarily shows that quantum mechanics can only give an incomplete description of reality. In John Bell's 1964 work, he developed these ideas further: assuming that causal relativity holds and also demanding that the physical world follows realism (taken together, assuming the principle of local realism), there is a contradiction to quantum mechanics; quantum mechanics is non-local in the sense of classical realism. It is important to understand that realism encompasses all settings of the detectors, since the particle pair does not know which variables are to be measured. The entire multivariant distribution function $P(\sigma_a, \sigma_b, \sigma_{a'}, \sigma_{b'})$, see (12.4), and not only their reductions $P(\sigma_a, \sigma_b)$, see (12.7), should exist.

Einstein was dissatisfied with quantum mechanics in many ways, on the one hand because he wanted a deterministic and non-probabilistic theory ('God doesn't throw dice'), and on the other hand because the physical world should be (locally) real ('The moon is there also with nobody watching'). In contrast to Einstein, Bohr always believed that the variables of a system are actually undefined as long as they have not been measured; Einstein tried to complete quantum mechanics until the end of his life and replace it with a deterministic theory. A well-known approach is the idea of hidden variables (hidden-variable theory), which cannot describe equantum mechanics, but which would describe the world as a real-local-deterministic one in a corresponding theory. John Bell formalized this idea in 1964 by assuming that the statistics of the process could be represented by a distribution function $\rho(\lambda)$ of a hidden variable λ . The measured values of Alice $[\sigma^A(\vec{a}, \lambda), \sigma^A(\vec{a}', \lambda)]$ and of Bob $[\sigma^B(\vec{b}, \lambda), \sigma^B(\vec{b}', \lambda)]$ then depend on the hidden variable λ ; evaluating their correlators

$$K(\vec{a}, \vec{b}) = \int d\lambda \rho(\lambda) \sigma^A(\vec{a}, \lambda) \sigma^B(\vec{b}, \lambda) \quad (12.18)$$

etc., and combining them into the Bell inequality, one arrives at the classical result, thus refuting the concept of a hidden-variable theory. Again, it is important that the distribution $\rho(\lambda)$ applies to all detector settings. The relationship between $\rho(\lambda)$ and $P(\sigma_a, \sigma_b, \sigma_{a'}, \sigma_{b'})$ can be found by writing down all tables (these are 16 tables that are numbered by λ) for the selected setting $\vec{a}, \vec{b}, \vec{a}', \vec{b}'$ and values $\sigma^A \in \{\pm 1\}, \sigma^B \in \{\pm 1\}$ and assigning each of these tables the probability $\rho(\lambda)$ (\rightarrow sums instead of integrals); these probabilities are precisely the 16 values of $P(\sigma_a, \sigma_b, \sigma_{a'}, \sigma_{b'})$.

Chapter 13

Identical Particles

We encounter another, classically inapproachable but mathematically consistent element, besides those of superposition and entanglement, when we want to describe several identical particles within quantum mechanics. We can easily see that this is indeed a problem that requires new structures: consider n identical particles, then, from a classical view, we can identify each particle, e.g., by marking each particle. In fact, even the markings are not really needed, since in classical mechanics, we can follow the trajectories of the particles and therefore identify them—classically the physics of distinguishable and identical particles is the same.

Within quantum mechanics, we can neither apply markings nor follow the trajectories, the latter being a consequence of the Heisenberg uncertainty principle. We thus cannot distinguish between identical particles and the quantum mechanical formalism must take this fact into account. The consistent inclusion of the indistinguishability of particles into the theory produces astonishing physical effects, which we will describe below. We first give a simple overview and then go a little deeper into the group theory of symmetric groups and the consequences for the physical world; in particular, we will deal with the problem of n spin-1/2 fermions (the problem of spin and race (in german ‘Rasse’), see the book of Res Jost).

The mathematical basis / the mathematical tool for this discussion is the symmetric group S_n of n objects, the permutation group. Basically, we need to know the behavior of the operators and the wave function under S_n to solve the problem.

13.1 Description of n Particles

We start with the problem of how to describe n particles. The following assumptions look reasonable for both wave functions and operators,

- the wave function shall depend on the n particle-coordinates, for example on their position and spin coordinates $\vec{r}_1, s_1; \dots; \vec{r}_n, s_n$, or in a more compact notation x_1, \dots, x_n , with $x_i = \vec{r}_i, s_i$,

$$\Psi(x_1, \dots, x_n) = \langle x_1, \dots, x_n | \Psi \rangle. \quad (13.1)$$

Accordingly, we define a Hilbert space \mathcal{H} over the complex functions $\Psi(x_1, \dots, x_n)$ of n arguments x_n with the scalar product

$$\langle \Psi, \Phi \rangle = \int dx_1 \cdots dx_n \Psi^*(x_1, \dots, x_n) \Phi(x_1, \dots, x_n). \quad (13.2)$$

A trivial basis in \mathcal{H} is the product basis deriving from a one-particle basis $\{\varphi_\alpha(x)\}$,

$$\Psi_{\alpha_1, \dots, \alpha_n}(x_1, \dots, x_n) = \varphi_{\alpha_1}(x_1) \cdots \varphi_{\alpha_n}(x_n), \quad (13.3)$$

but other bases are also possible.

- The operators act on the n -particle wave functions. The simplest case is the generalization of the one-particle operators,

$$A = \sum_i A_i, \quad \text{where } A_i \text{ acts on particle } i. \quad (13.4)$$

In a product basis, this operator acts in the following way,

$$\begin{aligned} \langle x_1, \dots, x_n | A \Psi \rangle &= A \Psi(x_1, \dots, x_n) \\ &= [A_1 \varphi_{\alpha_1}(x_1)] \varphi_{\alpha_2}(x_2) \cdots \varphi_{\alpha_n}(x_n) \\ &\quad + \varphi_{\alpha_1}(x_1) [A_2 \varphi_{\alpha_2}(x_2)] \cdots \varphi_{\alpha_n}(x_n) + \cdots, \end{aligned} \quad (13.5)$$

i.e., the operator $A = A_1 \otimes \mathbb{1}_2 \otimes \cdots \otimes \mathbb{1}_n + \mathbb{1}_1 \otimes A_2 \otimes \cdots \otimes \mathbb{1}_n + \cdots$ acts on the vector $|\alpha_1, \alpha_2, \dots, \alpha_n\rangle = |\varphi_{\alpha_1}\rangle \otimes |\varphi_{\alpha_2}\rangle \otimes \cdots$ in the space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots$.

Operators can, of course, also act on two (or more) particles at the same time and create interactions between them, for example pair-interactions,

$$V = \frac{1}{2} \sum_{i \neq j} V_{ij}, \quad \text{where } V_{ij} \text{ acts on the particle pair } i, j. \quad (13.6)$$

As an example, we consider n identical particles in \mathbb{R}^3 without spin.

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

$$|\vec{\alpha}_1, \dots, \vec{\alpha}_n\rangle = |\varphi_{\vec{\alpha}_1}\rangle \otimes |\varphi_{\vec{\alpha}_2}\rangle \otimes \dots \otimes |\varphi_{\vec{\alpha}_n}\rangle; \quad (13.7)$$

in the position representation,

$$\begin{aligned} \langle \vec{r}_i | \varphi_{\vec{\alpha}_i} \rangle &= \varphi_{\vec{\alpha}_i}(\vec{r}_i), \\ \langle \vec{r}_1, \dots, \vec{r}_n | \vec{\alpha}_1, \dots, \vec{\alpha}_n \rangle &= \varphi_{\vec{\alpha}_1}(\vec{r}_1) \dots \varphi_{\vec{\alpha}_n}(\vec{r}_n). \end{aligned} \quad (13.8)$$

- With the plane waves basis, $\vec{\alpha}_i \rightarrow \vec{k}_i$ and

$$\begin{aligned} \varphi_{\vec{k}}(\vec{r}) &= e^{i\vec{k} \cdot \vec{r}} / \sqrt{V}, \\ \langle \vec{r}_1, \dots, \vec{r}_n | \vec{k}_1, \dots, \vec{k}_n \rangle &= \prod_i e^{i\vec{k}_i \cdot \vec{r}_i} / \sqrt{V}. \end{aligned} \quad (13.9)$$

spans the product Hilbert space

$$\mathcal{H}^n = \mathbb{L}_2(\mathbb{R}^3) \otimes \dots \otimes \mathbb{L}_2(\mathbb{R}^3) = \mathbb{L}_2(\mathbb{R}^3)^{\otimes n} \quad (13.10)$$

- Examples of operators are

- the total momentum

$$P = \sum_i \vec{p}_i, \quad \vec{p}_i = \frac{\hbar}{i} \vec{\nabla}_i, \quad (13.11)$$

- the energy

$$\begin{aligned} H_0 &= \sum_i \frac{p_i^2}{2m}, \quad \text{for free particles,} \\ H &= H_0 + \frac{1}{2} \sum_{i \neq j} V(\vec{r}_i - \vec{r}_j) \quad \text{for int particles,} \end{aligned} \quad (13.12)$$

- the density

$$\rho(\vec{r}) = \sum_i \delta(\vec{r}_i - \vec{r}), \quad (13.13)$$

- the current density

$$\vec{j}(\vec{r}) = \sum_i \left[\frac{\vec{p}_i}{2m} \delta(\vec{r}_i - \vec{r}) + \delta(\vec{r}_i - \vec{r}) \frac{\vec{p}_i}{2m} \right]. \quad (13.14)$$

- Given these wave functions and operators, we form the expectation values (for the plane waves basis with states $\Psi = \Psi_{\vec{k}_1, \dots, \vec{k}_n}$)

$$\begin{aligned}\langle \vec{P} \rangle &= \sum_i \int d^3 r_1 \cdots d^3 r_n \Psi^*(\vec{r}_1, \dots, \vec{r}_n) \frac{\hbar}{i} \vec{\nabla}_i \Psi(\vec{r}_1, \dots, \vec{r}_n), \\ &= \sum_i \int d^3 r_i \varphi_{\vec{k}_i}^*(\vec{r}_i) \frac{\hbar}{i} \vec{\nabla}_i \varphi_{\vec{k}_i}(\vec{r}_i) = \sum_i \hbar \vec{k}_i,\end{aligned}\quad (13.15)$$

$$\langle H_0 \rangle = \sum_i \frac{\hbar^2 k_i^2}{2m}, \quad (13.16)$$

$$\begin{aligned}\langle H \rangle &= \langle H_0 \rangle + \frac{1}{2} \sum_{i \neq j} \int d^3 r_1 \cdots d^3 r_n \\ &\quad \times \Psi^*(\vec{r}_1, \dots, \vec{r}_n) V(\vec{r}_i - \vec{r}_j) \Psi(\vec{r}_1, \dots, \vec{r}_n) \\ &= \langle H_0 \rangle + \frac{1}{2} \sum_{i \neq j} \int d^3 r_i d^3 r_j \\ &\quad \times \varphi_{\vec{k}_i}^*(\vec{r}_i) \varphi_{\vec{k}_j}^*(\vec{r}_j) V(\vec{r}_i - \vec{r}_j) \varphi_{\vec{k}_i}(\vec{r}_i) \varphi_{\vec{k}_j}(\vec{r}_j) \\ &= \langle H_0 \rangle + \frac{1}{2} \sum_{i \neq j} \int d^3 r_i d^3 r_j \\ &\quad \times |\varphi_{\vec{k}_i}(\vec{r}_i)|^2 V(\vec{r}_i - \vec{r}_j) |\varphi_{\vec{k}_j}(\vec{r}_j)|^2.\end{aligned}\quad (13.17)$$

Above, we have trivially integrated over all coordinates that do not involve the momentum \vec{p}_i or in interaction $V(\vec{r}_i - \vec{r}_j)$ (with a unit weight for properly normalized wave functions).

13.2 Description of n identical particles

Since the particles are identical, there is no operator that allows for the identification of the individual particles. This implies that all allowed operators are symmetric, that is, they treat all particles equally. In a coordinate representation, the allowed operators are symmetric in the coordinates,

$$A(x_1, \dots, x_n) = A(x_{\pi_1}, \dots, x_{\pi_n}), \quad (13.18)$$

where $\pi(i)$ denotes a permutation

$$\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_n \end{pmatrix}. \quad (13.19)$$

Next, we analyze the expectation values of A ,

$$\begin{aligned}
 \langle A \rangle &= \int dx_1 \cdots dx_n \Psi^*(x_1, \dots, x_n) A(x_1, \dots, x_n) \Psi(x_1, \dots, x_n) \\
 &= \int dx_1 \cdots dx_n \Psi^*(x_1, \dots, x_n) A(x_{\pi_1}, \dots, x_{\pi_n}) \Psi(x_1, \dots, x_n) \\
 &= \int dx_1 \cdots dx_n \Psi^*(x_{\pi_1^{-1}}, \dots, x_{\pi_n^{-1}}) A(x_1, \dots, x_n) \\
 &\quad \times \Psi(x_{\pi_1^{-1}}, \dots, x_{\pi_n^{-1}}), \tag{13.20}
 \end{aligned}$$

where π^{-1} is the inverse of the permutation π . Obviously, we can swap the arguments in Ψ arbitrarily; since A is symmetric, $\langle A \rangle$ remains invariant. This is consistent with the interchangeability of the particles, i.e., the permutation of the coordinates in Ψ ; however, we do not know how Ψ behaves when exchanged, for example, in the case of a degeneracy where Ψ could rotate in the Hilbert space. The invariance of $\langle A \rangle$ therefore does not give us a condition or a further insight regarding the behavior of the wave function Ψ when exchanging the arguments.

We can formalize the above discussion in the following way. In dealing with a new symmetry, S_n , we seek some inspiration from previous discussions of symmetries. In analogy to the case of translations and rotations and their representations in Hilbert space,

$$\begin{array}{ll}
 \text{translation} & U_{\vec{a}} \Psi(\vec{r}) = e^{-i\vec{a} \cdot \vec{p}/\hbar} \Psi(\vec{r}) = \Psi(\vec{r} - \vec{a}), \\
 \text{rotation} & U_{\vec{\omega}} \Psi(\vec{r}) = e^{-i\vec{\omega} \cdot \vec{L}/\hbar} \Psi(\vec{r}) = \Psi(\mathbf{R}_{-\vec{\omega}} \vec{r}), \tag{13.21}
 \end{array}$$

we introduce the (unitary) representation of the permutation group S_n ¹,

$$U_{\pi} \Psi(x_1, \dots, x_n) = \Psi(x_{\pi_1^{-1}}, \dots, x_{\pi_n^{-1}}). \tag{13.22}$$

Making use of the definition of U_{π} , we can define a symmetric operator A by requiring that for all $\pi \in S_n$ the relation (with $U_{\pi}^{\dagger} = U_{\pi}^{-1}$)

$$U_{\pi}^{\dagger} A U_{\pi} = A \quad \Leftrightarrow \quad [A, U_{\pi}] = 0 \tag{13.23}$$

holds true. The central question we tackle next is that of finding the permuted wave function $U_{\pi} \Psi$. This takes us into the realm of representation theory, here of S_n ; obviously, U_{π} , $\pi \in S_n$ defines a representation of S_n in

¹Note that S_n is not a continuous group and thus there are no infinitesimal generators.

\mathcal{H} . For all (permitted, i.e., symmetric) operators A , and for the Hamilton operator H in particular, the relation $[A, U_\pi] = 0$ holds. As a result, we can diagonalize A and U_π at the same time and exploit the structure defined through the representations of S_n .

The common lore then goes like this: Let $A\Psi_a = a\Psi_a$, then $AU_\pi\Psi_a = aU_\pi\Psi_a$ and $\{U_\pi\Psi_a | \pi \in S_n\}$ spans an eigenspace Eig_a of A .² Given the eigenspace Eig_a with $\dim \text{Eig}_a = D_a$, we then can define a D_a -dimensional representation of S_n and reduce it further (e.g., in the presence of further symmetries). We then can apply the whole machinery of group theory on this problem, e.g., in order to efficiently reduce these representations. To do this, we require the knowledge about the group S_n and its irreducible representations, how many of them there are, what dimensions they have, etc. The effect of U_π on Ψ can then be calculated in each of these spaces. Fortunately, very few representations of S_n are realized in nature. In order to find them and understand the amazing restriction imposed by nature, we must briefly go through the representation theory of S_n .

13.3 Symmetric Group S_n

Elements of the group S_n are the permutation of n objects,

$$\begin{aligned} \pi &= \begin{pmatrix} 1 & 2 & \cdots & n \\ \pi_1 & \pi_2 & \cdots & \pi_n \end{pmatrix} \\ &= \begin{pmatrix} 2 & 4 & 3 & \cdots \\ \pi_2 & \pi_4 & \pi_3 & \cdots \end{pmatrix}. \end{aligned} \quad (13.24)$$

The inverse of the permutation π is

$$\pi^{-1} = \begin{pmatrix} \pi_1 & \pi_2 & \cdots & \pi_n \\ 1 & 2 & \cdots & n \end{pmatrix}. \quad (13.25)$$

The group S_n has $n!$ elements, and $g = n!$ is the order of S_n .

The permutation $\pi_{\alpha_1, \dots, \alpha_k}$ is called a k -cycle,

$$\pi_{\alpha_1, \dots, \alpha_k} = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_k & \beta_1 & \cdots & \beta_{n-k} \\ \alpha_2 & \alpha_3 & \alpha_4 & \cdots & \alpha_1 & \beta_1 & \cdots & \beta_{n-k} \end{pmatrix}; \quad (13.26)$$

²In the case of $A = H$, we consider Ψ_E with $H\Psi_E = E\Psi_E$ and find that $U_\pi\Psi_E$ also lies in Eig_E .

$\pi_{\alpha_1, \dots, \alpha_k}$ is often abbreviated with $(\alpha_1 \alpha_2 \dots \alpha_k)$. The simplest cycle is the two-cycle or the transposition $\pi_{\alpha_1, \alpha_2} = (\alpha_1 \alpha_2)$, e.g.,

$$\pi_{35} = (35) = \begin{pmatrix} 3 & 5 & \beta_1 & \cdots & \beta_{n-2} \\ 5 & 3 & \beta_1 & \cdots & \beta_{n-2} \end{pmatrix}. \quad (13.27)$$

Every permutation can be written as a product of transpositions. A permutation is called even (odd) if it can be written as a product of an even (odd) number of transpositions. We call

$$\begin{aligned} (-1)^\pi & \text{ the sign of } \pi, \\ \pi \text{ even} & \Rightarrow (-1)^\pi = 1, \\ \pi \text{ odd} & \Rightarrow (-1)^\pi = -1. \end{aligned} \quad (13.28)$$

The decomposition of π into independent cycles³ defines an equivalence class⁴; π and π' are equivalent if they have the same cycle structure. As an example, we consider a permutation with the cycle $3 + 2$,

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 2 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 5 & 4 & 2 & 3 \\ 5 & 4 & 1 & 3 & 2 \end{pmatrix} = (1\ 5\ 4)(2\ 3). \quad (13.29)$$

Indeed, the transformation $\nu\pi\nu^{-1}$ with $\nu \in S_n$ leaves the cycle structure invariant; we illustrate this fact with the above permutation $(1\ 5\ 4)(2\ 3)$ and a specific ν ,

$$\begin{aligned} \nu & \equiv \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 3 & 5 \end{pmatrix} = \begin{pmatrix} 5 & 4 & 1 & 3 & 2 \\ 5 & 3 & 2 & 1 & 4 \end{pmatrix} \\ \nu^{-1} & = \begin{pmatrix} 2 & 4 & 1 & 3 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 2 & 5 & 3 & 4 & 1 \\ 1 & 5 & 4 & 2 & 3 \end{pmatrix}. \end{aligned} \quad (13.30)$$

The transformed permutation $\nu\pi\nu^{-1}$ is constructed straightforwardly,

$$\begin{aligned} \nu\pi\nu^{-1} & = \begin{pmatrix} 2 & 5 & 3 & 4 & 1 \\ 1 & 5 & 4 & 2 & 3 \\ 5 & 4 & 1 & 3 & 2 \\ 5 & 3 & 2 & 1 & 4 \end{pmatrix} \begin{array}{l} \leftarrow \nu^{-1} \\ \leftarrow \pi \\ \leftarrow \nu \end{array} \\ & = \begin{pmatrix} 2 & 5 & 3 & 4 & 1 \\ 5 & 3 & 2 & 1 & 4 \end{pmatrix} = (2\ 5\ 3)(4\ 1), \end{aligned} \quad (13.31)$$

and the cycle structure is preserved. This fact can be proved for a general case and thus the cycle structure defines an equivalence criterion.

³that is, each number occurs only once.

⁴two permutations π and π' are equivalent if there is a $\nu \in S_n$, such that $\nu\pi\nu^{-1} = \pi'$.

To find the number K of equivalence classes of S_n , we need to know all possible cycle structures, that follow from the decompositions of n . We call $[\lambda_1, \lambda_2, \dots, \lambda_k] \equiv \lambda$ a decomposition of n if $\lambda_i \geq 1$ are integers, sorted according to $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_k$, and satisfy

$$\lambda_1 + \lambda_2 + \dots + \lambda_k = n. \quad (13.32)$$

The decompositions of n provide us with the equivalence classes of S_n . For the example with $n = 5$ one finds $K = 7$ equivalence classes by simple counting,

$$\begin{aligned} S_5 : \quad 5 &= 5 \\ &= 4 + 1 \\ &= 3 + 2 \\ &= 3 + 1 + 1 \\ &= 2 + 2 + 1 \\ &= 2 + 1 + 1 + 1 \\ &= 1 + 1 + 1 + 1 + 1 \quad \rightarrow \quad 7 \text{ classes.} \end{aligned} \quad (13.33)$$

The number of irreducible representations of S_n is equal to the number K of classes. Let d_i be the dimension of the i -th irreducible representation. Then (this relationship applies in general; for S_n , we have $g = n!$)

$$g = \sum_{i=1}^K d_i^2. \quad (13.34)$$

You can try out this relation on a few examples,

$$\begin{aligned} S_2 &: \quad K = 2, \quad 2! = 1 + 1, \\ S_3 &: \quad K = 3, \quad 3! = 1 + 1 + 2^2, \\ S_4 &: \quad K = 5, \quad 4! = 24 = 1 + 1 + 2^2 + 3^2 + 3^2, \\ S_5 &: \quad K = 7, \quad 5! = 120 = 1 + 1 + 4^2 + 4^2 + 5^2 + 5^2 + 6^2. \end{aligned} \quad (13.35)$$

The two 1-dimensional representations show up in all decompositions of S_n . These are the anti-symmetric (A) and the symmetric (S) representation, which are generated by the following (anti-)symmetrization operators in Hilbert space,

$$\begin{aligned} U_a &= \frac{1}{n!} \sum_{\pi \in S_n} (-1)^\pi U_\pi, \\ U_s &= \frac{1}{n!} \sum_{\pi \in S_n} U_\pi. \end{aligned} \quad (13.36)$$

One easily shows that

$$U_\pi U_s = U_s U_\pi = U_s \quad \text{and} \quad U_\pi U_a = U_a U_\pi = (-1)^\pi U_a. \quad (13.37)$$

13.4 Fermions and bosons

It is an experimental fact that in nature only two representations of S_n in \mathcal{H} are realized, the two 1-dimensional representations S and A . Let \mathcal{H} be the Hilbert space of the n -particle functions; then only the sectors

$$\begin{aligned} \mathcal{H}_A &= \{U_a \Psi | \Psi \in \mathcal{H}\}, \\ \mathcal{H}_S &= \{U_s \Psi | \Psi \in \mathcal{H}\}, \end{aligned} \quad (13.38)$$

of \mathcal{H} are relevant.⁵ The functions in \mathcal{H}_A satisfy

$$\langle x_1, \dots, x_n | U_\pi \Psi \rangle = \langle x_{\pi_1^{-1}}, \dots, x_{\pi_n^{-1}} | \Psi \rangle = (-1)^\pi \langle x_1, \dots, x_n | \Psi \rangle, \quad (13.39)$$

while for those in \mathcal{H}_S , we have

$$\langle x_1, \dots, x_n | U_\pi \Psi \rangle = \langle x_{\pi_1^{-1}}, \dots, x_{\pi_n^{-1}} | \Psi \rangle = \langle x_1, \dots, x_n | \Psi \rangle. \quad (13.40)$$

Hence, nature realizes only a small part of \mathcal{H} , namely its symmetric and antisymmetric sectors. We call particles whose wave functions lie in $\mathcal{H}_A/\mathcal{H}_S$ (and thus are anti-symmetric/symmetric under the exchange of their arguments) fermions/bosons.

\mathcal{H}_A		\mathcal{H}_S
anti-symmetric	\leftrightarrow	symmetric
fermions		bosons

This takes us to the following problem: Given an ensemble of particles, should we describe them with a wave function Ψ in \mathcal{H}_A or in \mathcal{H}_S ? Are the particles fermions or bosons? The answer is given by the theorem of spin and statistics.

⁵The basis/product states Ψ in (13.38) have to be properly normalized, e.g., with an additional factor $\sqrt{n!}$ for the case of fermions.

13.5 Spin & Statistics

The following relationship between a particle's spin and its symmetry properties under S_n (statistics) derives from relativistic quantum field theory,

$$\begin{array}{ccc} \text{half-integer spin} & & \text{integer Spin} \\ \text{fermions} & \leftrightarrow & \text{bosons} \end{array} \quad (13.41)$$

Examples for particles and their corresponding statistics are,

fermions	bosons
– leptons e^- , μ , τ	– Higgs (spin 0)
– lepton-neutrinos ν_e , ν_μ , ν_τ	– mesons π, K with spin 0, or ρ, ω with spin 1, with two 2 quarks each.
– baryons $p, n, \Lambda, \Sigma, \Xi, \dots$ with 3 quarks each	– gauge bosons (Yang-Mills)
– ^3He	–electro-weak interaction: γ, W^\pm, Z^0 with spin 1.
	–strong interaction: gluons with spin 1.
	– ^4He (spin 0).

All fermions listed here have spin = 1/2.

13.6 Applications

13.6.1 Wave function of two spin-1/2 fermions

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

$$\Psi = \varphi_\nu(\vec{r}_1, s_1) \varphi_\mu(\vec{r}_2, s_2). \quad (13.42)$$

The physical state has to be properly symmetrized

$$\sqrt{2} U_a \Psi = \frac{1}{\sqrt{2!}} \sum_{\pi \in S_2} (-1)^\pi U_\pi \Psi(\vec{r}_1, s_1; \vec{r}_2, s_2)$$

$$\begin{aligned}
&= \frac{1}{\sqrt{2!}} [\Psi(\vec{r}_1, s_1; \vec{r}_2, s_2) - \Psi(\vec{r}_2, s_2; \vec{r}_1, s_1)] \\
&= \frac{1}{\sqrt{2!}} [\varphi_\nu(\vec{r}_1, s_1) \varphi_\mu(\vec{r}_2, s_2) - \varphi_\nu(\vec{r}_2, s_2) \varphi_\mu(\vec{r}_1, s_1)]. \quad (13.43)
\end{aligned}$$

If $\mu = \nu$, then $U_a \Psi = 0$ and hence, no two fermions can be in the same state: this is the famous **Pauli exclusion principle**. The Pauli exclusion principle for fermions applies generally to n fermions.

13.6.2 Spin & orbital wave functions for two spin-1/2 fermions

Next, we separate the spin and orbital parts of Ψ ,

$$\Psi(\vec{r}_1, s_1; \vec{r}_2, s_2) = \varphi(\vec{r}_1, \vec{r}_2) \chi(s_1, s_2). \quad (13.44)$$

The total wave function Ψ is anti-symmetric if

$$\left. \begin{array}{l} \varphi \text{ symmetric} \\ \chi \text{ anti-symmetric} \end{array} \right\} \text{ or } \left\{ \begin{array}{l} \varphi \text{ anti-symmetric} \\ \chi \text{ symmetric} \end{array} \right. \quad (13.45)$$

The spin component $\chi(s_1, s_2)$ can be reduced according to $\vec{S} = \vec{S}_1 + \vec{S}_2$, $\mathcal{H}_{1/2} \otimes \mathcal{H}_{1/2} = \mathcal{H}_0 \oplus \mathcal{H}_1$. In the Spin-Hilbert space \mathcal{H}_0 , the singlet state $\chi_{0,0}(s_1, s_2)$ is anti-symmetric in s_1, s_2 ,⁶

$$\begin{aligned}
\chi_{0,0}(s_1, s_2) &= \langle s_1, s_2 | 0, 0 \rangle \\
&= \frac{1}{\sqrt{2}} [\langle s_1 | \uparrow \rangle \langle s_2 | \downarrow \rangle - \langle s_1 | \downarrow \rangle \langle s_2 | \uparrow \rangle] \\
&= \frac{1}{\sqrt{2}} [\chi_\uparrow(s_1) \chi_\downarrow(s_2) - \chi_\downarrow(s_1) \chi_\uparrow(s_2)]. \quad (13.46)
\end{aligned}$$

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,

$$\begin{aligned}
\chi_{1,1}(s_1, s_2) &= \langle s_1, s_2 | 1, 1 \rangle = \chi_\uparrow(s_1) \chi_\uparrow(s_2), \\
\chi_{1,0}(s_1, s_2) &= \langle s_1, s_2 | 1, 0 \rangle = \frac{1}{\sqrt{2}} [\chi_\uparrow(s_1) \chi_\downarrow(s_2) + \chi_\downarrow(s_1) \chi_\uparrow(s_2)], \\
\chi_{1,-1}(s_1, s_2) &= \langle s_1, s_2 | 1, -1 \rangle = \chi_\downarrow(s_1) \chi_\downarrow(s_2). \quad (13.47)
\end{aligned}$$

We want to combine the above symmetric and anti-symmetric spin functions χ with properly symmetrized orbital functions φ and thus define the combinations

⁶in our notation, $\chi_\uparrow(s = \uparrow) = 1$ and $\chi_\uparrow(s = \downarrow) = 0$, etc.

$$\begin{aligned} \varphi \text{ symmetric:} \quad \varphi_s &= [\varphi_1(\vec{r}_1)\varphi_2(\vec{r}_2) + \varphi_2(\vec{r}_1)\varphi_1(\vec{r}_2)]/\sqrt{2}, \\ \varphi \text{ anti-symmetric:} \quad \varphi_a &= [\varphi_1(\vec{r}_1)\varphi_2(\vec{r}_2) - \varphi_2(\vec{r}_1)\varphi_1(\vec{r}_2)]/\sqrt{2}. \end{aligned}$$

In this way, we can diagonalize S^2 , S_z , and $U_\pi \in S_2$ at the same time and obtain wave functions Ψ with defined spin and symmetry. Furthermore, the spin determines the symmetry of the orbit φ ,

$$\begin{aligned} S = 1 &\Rightarrow \chi = \text{symmetric}, \quad \varphi = \text{anti-symmetric}, \\ S = 0 &\Rightarrow \chi = \text{anti-symmetric}, \quad \varphi = \text{symmetric}. \end{aligned} \quad (13.48)$$

A similar principle (theorem of spin and race (in german ‘Rasse’)) applies to n spin-1/2 fermions. Obviously, fermions interfere with each other via the Pauli principle and the question arises whether we have to antisymmetrize all the electrons in the world. Luckily, this is not the case. Consider two electrons, one in Zürich and one in Vienna, each described by a one-particle state $\varphi_z(\vec{r}_1, s_1)$ and $\varphi_v(\vec{r}_1, s_1)$. It is

$$\Psi = \frac{1}{\sqrt{2}}[\varphi_z(x_1)\varphi_v(x_2) - \varphi_v(x_1)\varphi_z(x_2)]. \quad (13.49)$$

If A is a 1-particle operator, its expectation value is given by

$$\begin{aligned} \langle A \rangle &= \frac{1}{2} \int dx_1 dx_2 \left[\varphi_z^*(x_1)\varphi_v^*(x_2) A \varphi_z(x_1)\varphi_v(x_2) \right. \\ &\quad + \varphi_v^*(x_1)\varphi_z^*(x_2) A \varphi_v(x_1)\varphi_z(x_2) \\ &\quad - \varphi_z^*(x_1)\varphi_v^*(x_2) A \varphi_v(x_1)\varphi_z(x_2) \\ &\quad \left. - \varphi_v^*(x_1)\varphi_z^*(x_2) A \varphi_z(x_1)\varphi_v(x_2) \right] \approx 0 \\ &\cong \langle A \rangle_z + \langle A \rangle_v. \end{aligned} \quad (13.50)$$

We see that the interference terms carry the symmetrization properties.⁷ For non-overlapping wave functions these terms disappear and we can work without (anti-)symmetrization of the wave function at large separation.

13.6.3 Wave function of two spin-0 bosons

Let $\varphi_\mu(x)$ be a basis of a one-particle wave functions. The physically allowed two-particle wave functions Ψ are

$$\Psi(\vec{r}_1, \vec{r}_2) = \frac{1}{\sqrt{2}}[\varphi_\mu(\vec{r}_1)\varphi_\nu(\vec{r}_2) + \varphi_\nu(\vec{r}_1)\varphi_\mu(\vec{r}_2)]. \quad (13.51)$$

⁷Note that i) for bosons the last two terms in (13.50) are added, ii) there are no physically realistic 2-particle operators that act over large distances, iii) consider the problem of entangled states.

For bosons, a state with $\mu = \nu$ is allowed,

$$\Psi(\vec{r}_1, \vec{r}_2) \Big|_{\mu=\nu} = \varphi_\mu(\vec{r}_1) \varphi_\mu(\vec{r}_2), \quad (13.52)$$

and hence, both bosons may occupy the same place because, i.e., $\Psi(\vec{r}, \vec{r}) = [\varphi_\mu(\vec{r})]^2 \neq 0$.

Note that Fermions avoid each other, while bosons ‘like’ each other

$$\begin{aligned} \Psi(x_1, x_2) &= \frac{1}{\sqrt{2}} \left(\varphi_\nu(x_1) \varphi_\mu(x_2) \mp \varphi_\mu(x_1) \varphi_\nu(x_2) \right) \\ &\downarrow x_1 = x_2 \equiv x \\ &= \begin{cases} 0, & \text{fermions,} \\ \sqrt{2} \varphi_\nu(x) \varphi_\mu(x), & \text{bosons.} \end{cases} \end{aligned} \quad (13.53)$$

□

13.6.4 Composite particles

Consider two H-atoms as illustrated in figure 13.1. The wavefunction $\Psi(x_1, X_1, x_2, X_2)$ must be anti-symmetric in the electron coordinates x_1, x_2

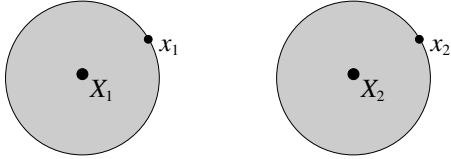


Fig. 13.1: About the symmetry of the wave function of two hydrogen atoms: X_1, X_2 denote the coordinates of the protons, x_1, x_2 the coordinates of the electrons.

and also anti-symmetric in the coordinates X_1, X_2 of the fermionic protons. Swapping coordinates $x_1, X_1; x_1, X_2; x_2, X_1; x_2, X_2$ remains mathematically unrestricted but is physically not meaningful. Exchanging the atoms gives (first exchange the electrons, then the protons)

$$\begin{aligned} \Psi(x_2, X_2, x_1, X_1) &= -\Psi(x_1, X_2, x_2, X_1) \\ &= (-1)^2 \Psi(x_1, X_1, x_2, X_2). \end{aligned} \quad (13.54)$$

As a result, we find that the H-atoms, each consisting of 2 fermions, behave like bosons. Note the compatibility with the spin & statistics theorem from section 13.5,

$$\text{spin of the H-atom} \hat{=} \mathcal{H}_{1/2} \otimes \mathcal{H}_{1/2} = \mathcal{H}_0 \oplus \mathcal{H}_1, \quad (13.55)$$

the H-atoms have an integer spin, and thus are bosons. Other composite particles are treated similarly, for example

$$\begin{aligned} {}^3\text{He} &= 2p + 1n + 2e = \text{fermion} \rightarrow \text{pairing transition}, \\ {}^4\text{He} &= 2p + 2n + 2e = \text{boson} \rightarrow \lambda \text{ transition}. \end{aligned} \tag{13.56}$$

Both He liquids become superfluid at low temperatures: While ${}^4\text{He}$ ‘Bose condenses’ (at low temperatures $T < 2.17$ K, a Bose-Einstein condensate (BEC) is formed, all ${}^4\text{He}$ atoms occupy the same one-particle state, this is a so-called condensation in k -space), the fermionic ${}^3\text{He}$ atoms must first pair up (\rightarrow bosons), and then change into the superfluid state via a so-called BCS (Bardeen-Cooper-Schrieffer) transition at very low temperatures $T \sim \text{few mK}$. The amazing fact is that the interaction between the atoms in the liquid is carried by the electrons, which have identical shells for ${}^3\text{He}$ and ${}^4\text{He}$. The different types of nuclei ‘hide’ inside these shells and do not see each other. Nevertheless, the liquids behave completely differently!

13.6.5 Symmetry under time evolution

It is easy to check that the time evolution preserves the symmetry, since if

$$U_\pi \Psi = (\pm)^\pi \Psi, \quad \begin{array}{l} \text{bosons (+)} \\ \text{fermions (-)} \end{array} \tag{13.57}$$

then

$$\begin{aligned} U_\pi \Psi(t + dt) &= U_\pi \Psi(t) + U_\pi [\partial_t \Psi(t)] dt \\ &= (\pm)^\pi \Psi - i U_\pi (H/\hbar) \Psi dt \\ &= (\pm)^\pi \Psi - i (H/\hbar) U_\pi \Psi dt \\ &= (\pm)^\pi [\Psi + (\partial_t \Psi) dt] = (\pm)^\pi \Psi(t + dt). \end{aligned} \tag{13.58}$$

13.6.6 Many-body states of non-interacting identical particles

For non-interacting particles, we write the Hamiltonian in the form

$$H = \sum_{i=1}^N H_0(x_i), \quad x_i = \vec{r}_i, s_i, \tag{13.59}$$

with the 1-particle operator

$$H_0(x) = \frac{p^2}{2m} + V(\vec{r}), \quad (13.60)$$

(we neglect spin dependencies in H_0). With the 1-particle solutions

$$H_0\varphi_\varepsilon = \varepsilon\varphi_\varepsilon, \quad \text{and } \{\varphi_\varepsilon\} \text{ a complete orthonormal set,} \quad (13.61)$$

we can define N -particle product eigenstates of H ,

$$\Psi(\vec{r}_1, \dots, \vec{r}_N) = \varphi_{\varepsilon_1}(\vec{r}_1)\varphi_{\varepsilon_2}(\vec{r}_2) \cdots \varphi_{\varepsilon_N}(\vec{r}_N) \quad (13.62)$$

with the total energy $E = \sum_i \varepsilon_i$. (13.62) describes an N -particle ensemble with particles in the states $\varphi_{\varepsilon_1}, \varphi_{\varepsilon_2}, \dots, \varphi_{\varepsilon_N}$, for example for $N = 4$ and V a potential well

$$\Psi_{1356} = \varphi_{\varepsilon_1}(\vec{r}_1)\varphi_{\varepsilon_3}(\vec{r}_2)\varphi_{\varepsilon_5}(\vec{r}_3)\varphi_{\varepsilon_6}(\vec{r}_4), \quad (13.63)$$

as shown in Figure 13.2. Note that Ψ_{1356} is not yet a physical 4-particle state because it does not yet have a well-defined symmetry under particle exchange. We would now like to describe fermions with spin = 1/2.

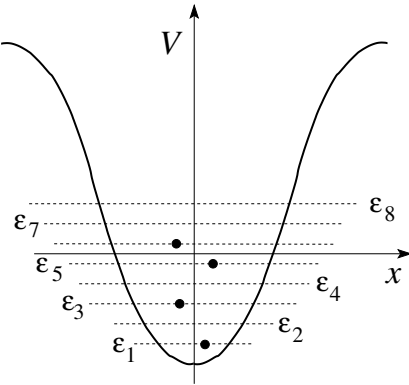


Fig. 13.2: Occupation of a fermionic system. The Pauli principle (the anti-symmetry of the wave function) allows at most one particle per state.

1. We have to take into account the spin wavefunction

$$\begin{aligned} \varphi_\varepsilon(\vec{r}) &\rightarrow \varphi_{\varepsilon,\sigma}(x) = \varphi_{\varepsilon,\sigma}(\vec{r}, s) = \varphi_\varepsilon(\vec{r})\chi_\sigma(s), \\ \chi_\sigma &\in \{\chi_\uparrow, \chi_\downarrow\}, \\ \chi_\uparrow(s = \hbar/2) &= 1, \quad \chi_\uparrow(s = -\hbar/2) = 0. \end{aligned} \quad (13.64)$$

2. We have to anti-symmetrize

$$\begin{aligned}\Psi &\rightarrow \Psi_a = \frac{1}{\sqrt{N!}} \sum_{\pi \in S_N} (-1)^\pi \varphi_{\varepsilon_1 \sigma_1}(x_{\pi_1}) \cdots \varphi_{\varepsilon_N \sigma_N}(x_{\pi_N}) \\ &= \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_{\varepsilon_1 \sigma_1}(x_1) & \cdots & \varphi_{\varepsilon_1 \sigma_1}(x_N) \\ \varphi_{\varepsilon_2 \sigma_2}(x_1) & \cdots & \varphi_{\varepsilon_2 \sigma_2}(x_N) \\ \vdots & & \vdots \\ \varphi_{\varepsilon_N \sigma_N}(x_1) & \cdots & \varphi_{\varepsilon_N \sigma_N}(x_N) \end{vmatrix}. \quad (13.65)\end{aligned}$$

(13.65) is called a Slater determinant. If there is a pair $i \neq j$ with $\varepsilon_i, \sigma_i = \varepsilon_j, \sigma_j$, then $\Psi_a \equiv 0$. Thus the Pauli exclusion principle⁸ requires that no two fermions occupy the same state in spin and orbit.

The ground state for $N(+1)$ fermions (N even) results from the occupation of every (orbital) state $\varphi_\varepsilon(\vec{r})$ with two electrons. In accordance with the Pauli principle, one electron occupies the state with spin $= \uparrow$ and the other one with spin $= \downarrow$. In the energy spectrum the following states then are occupied,

$$\varepsilon_1 \uparrow, \varepsilon_1 \downarrow; \varepsilon_2 \uparrow, \varepsilon_2 \downarrow, \dots, \varepsilon_{N/2} \downarrow, \varepsilon_{N/2} \uparrow \quad (\varepsilon_{N/2+1} \uparrow \text{ or } \downarrow). \quad (13.66)$$

This is the fermion ground state configuration with the energy

$$E_0 = \sum_i 2\varepsilon_i \quad (+ \varepsilon_{N/2+1}). \quad (13.67)$$

13.6.7 Fermi sea

In the simplest, i.e., non-interacting, case, the electrons in a solid are described by

- inner or valence electrons; these are electrons firmly bound to the atoms (or in an atomic group) and define full electron bands.
- conduction electrons; these are quasi-free moving electrons in partially filled bands, cf. Fig. 13.3.

The bands are filled to the Fermi level. If the Fermi level is in a band, the solid is a metal, if the Fermi level is in an energy gap, we deal with a

⁸1925 postulated for the first time by Wolfgang Pauli to explain the periodic table of the elements.

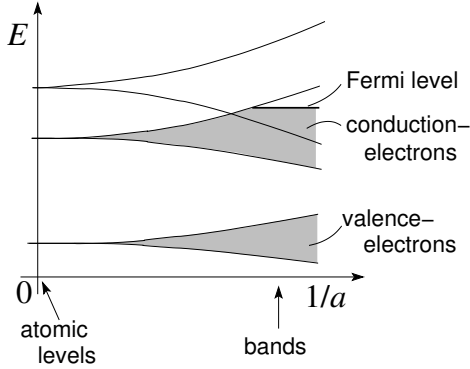


Fig. 13.3: Electron states in the solid form bands. Well separated atoms with $a \rightarrow \infty$, a the lattice distance, show a discrete spectrum. At finite (small) distances, the electrons are able to hop between atoms and energy bands are formed which are filled to the Fermi level.

semiconductor or an insulator. The conduction electrons are conveniently described as gas of spin = 1/2 fermions. The ground state $|\Psi_0\rangle$ has the form of a Slater determinant with states filled up to the Fermi level k_F ,

$$\Psi_0 = \text{Slater determinant with } \varphi_{\vec{k},\sigma} = e^{i\vec{k}\cdot\vec{r}}\chi_\sigma(s)/\sqrt{V}, \quad (13.68)$$

and \vec{k} in the first Brillouin zone. When exchanging two particles, Ψ_0 behaves anti-symmetrically,

$$\Psi_0(\dots, x_{i+k}, \dots, x_i, \dots) = -\Psi_0(\dots, x_i, \dots, x_{i+k}, \dots), \quad (13.69)$$

see also (13.37).

The energy density associated with the ground state is given by (with the density of states $\rho(\varepsilon)$ following from the relation $\rho(\varepsilon)d\varepsilon = \rho(k)dk$ and the free dispersion $\varepsilon_k = \hbar^2 k^2/2m$; here we consider the density of states in energy per volume with units of density/energy)

$$e_0 = \frac{E_0}{V} = 2 \int_0^{\varepsilon_F} d\varepsilon \rho(\varepsilon) \varepsilon, \quad \text{with } \rho(\varepsilon) = \frac{m}{2\pi^2 \hbar^3} \sqrt{2m\varepsilon}, \quad (13.70)$$

where the factor 2 before the integral takes the spin degeneracy into account.⁹ The relationship between the electron density n

$$n = \frac{N}{V} = 2 \int_0^{\varepsilon_F} d\varepsilon \rho(\varepsilon), \quad (13.71)$$

and the parameters $\varepsilon_F = p_F^2/2m$, $p_F = \hbar k_F$, $k_F = \sqrt{2m\varepsilon_F}/\hbar$ of the Fermi gas is given by

$$k_F^3 = 3\pi^2 n. \quad (13.72)$$

⁹Alternatively, spin degeneracy is often taken into account in the density of states ρ .

This results in the ground state energy density

$$e_0 = \frac{3}{5} n \varepsilon_F \quad (13.73)$$

and the density of states

$$\rho(\varepsilon) = \rho(\varepsilon_F) \left(\frac{\varepsilon}{\varepsilon_F} \right)^{1/2}, \quad \rho(\varepsilon_F) = \frac{3}{4} \frac{n}{\varepsilon_F}. \quad (13.74)$$

The ground state energy $\propto N\varepsilon_F$ of N fermions is always high because particles avoid each other and therefore occupy levels at high energies. We find the following energy scale in typical metals: With

$$\begin{aligned} \varepsilon_F &= \frac{\hbar^2 k_F^2}{2m_e} = 3.81 \text{ eV} \cdot k_F^2 [\text{\AA}^{-1}] \quad \text{and} \\ k_F &\sim \frac{\pi}{a}, \quad a = \text{lattice constant} \sim 3\text{\AA}, \end{aligned} \quad (13.75)$$

we obtain $k_F \sim 1 \text{ \AA}^{-1}$ and $\varepsilon_F \sim$ a few eV typically, that is a few 10^4 K (Kelvin). Electrons in a metal at room temperature are not 300 K, but $\sim 30'000$ K ‘hot’. Why do the metals still feel ‘cold’?

13.6.8 Bose condensate

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

$$\Psi_0 = \prod_{i=1}^N \varphi_{\varepsilon_1}(\vec{r}_i) \quad (13.76)$$

is already symmetric and the associated energy is $E_0 = N\varepsilon_1$. The bosons thus condense in the state with the lowest single-particle energy ε_1 . This condensation phenomenon finally is responsible for the superfluidity of a bosonic liquid at low temperatures.¹⁰ Fermions can only condense if they pair first, for example in the form of Cooper pairs in superconductivity in metals.¹¹

¹⁰A bosonic gas without interaction condenses but does not become superfluid—the latter requires a finite interaction, repulsive at small-distances to avoid a collapse. Alternatively, a bosonic system also may crystallize.

¹¹This pairing phenomenon is more complex in nature and we refer to the corresponding standard literature on BCS (Bardeen-Cooper-Schrieffer) theory.

13.6.9 Excited states

Fermions In an excited state, we do not select the $N/2$ lowest levels in the Slater determinant, but instead include a set of higher levels (excited electrons), leaving some of the lower levels empty (excited holes when energies are measured with respect to the Fermi level), as illustrated in Figure 13.4.

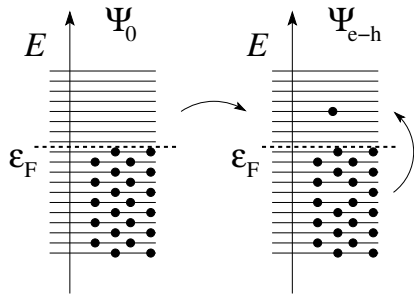


Fig. 13.4: Excited states in the spin-1/2 Fermi system. In the ground state Ψ_0 , states are occupied up to the Fermi energy ε_F . In an excited state Ψ_{eh} (at fixed particle number), a one-particle state below ε_F remains unoccupied (\rightarrow hole ‘h’) and the missing particle (‘e’) is filled into an unoccupied state above ε_F .

Bosons We occupy the i^{th} level N_i times and obtain the excited state¹²

$$\Psi_s = \sqrt{\frac{N_1! \cdots N_k!}{N!}} \sum'_{\pi \in S_N} \varphi_{\varepsilon_1}(x_{\pi_1}) \cdots \varphi_{\varepsilon_k}(x_{\pi_N}). \quad (13.77)$$

Here, the sum \sum' only extends over those $\pi \in S_N$ which lead to different terms.¹³ The ground state is obtained for $N_1 = N$ and all other states $N_{i>1} = 0$ unoccupied. This is consistent with (13.76).

13.6.10 Scattering of identical particles

Consider the collision or scattering of two particles. We separate the center of mass and relative coordinates $\vec{R} = \vec{r}_1 + \vec{r}_2$ and $\vec{r} = \vec{r}_1 - \vec{r}_2$. This makes \vec{R} symmetric and \vec{r} anti-symmetric in \vec{r}_1 and \vec{r}_2 . We split the total wave function into spin and orbital parts,

$$\Psi(x_1, x_2) = e^{i\vec{P} \cdot \vec{R}} \psi(\vec{r}) \chi(s_1, s_2). \quad (13.78)$$

¹²Note that $0! = 1! = 1$

¹³The factor $[\Pi_i N_i!]/N!$ takes into account the number of identical terms; alternatively, Ψ_s can be written in the form $\Psi_s = \sqrt{N!/N_1! \cdots N_k!} s[\varphi_{\varepsilon_1}(x_{\pi_1}) \cdots \varphi_{\varepsilon_k}(x_{\pi_N})]$.

For distinguishable particles, the scattering solution has the asymptotic form

$$\psi(\vec{r}) \sim e^{i\vec{k}\cdot\vec{r}} + f(\theta) \frac{e^{ikr}}{r}. \quad (13.79)$$

Spin-0 Bosons: For bosons with spin 0, we have $\chi = 1$. Given the symmetry of Ψ under exchange of particles, we require that $\psi(\vec{r}) = \psi(-\vec{r})$. Accordingly, we have to symmetrize the scattering solution; using the fact that the exchange of the particle via $\vec{r} \rightarrow -\vec{r}$ in polar coordinates is equivalent to the transformation $\theta \rightarrow \pi - \theta$, $r \rightarrow r$, we request that

$$\psi \sim \left(e^{i\vec{k}\cdot\vec{r}} + e^{-i\vec{k}\cdot\vec{r}} \right) + [f(\theta) + f(\pi - \theta)] \frac{e^{ikr}}{r}. \quad (13.80)$$

For the differential cross section, we find the expression

$$\begin{aligned} \frac{d\sigma}{d\Omega} &= |f(\theta) + f(\pi - \theta)|^2 \\ &= \underbrace{|f(\theta)|^2 + |f(\pi - \theta)|^2}_{\text{classical term}} + \underbrace{2\text{Re}[f^*(\theta)f(\pi - \theta)]}_{\text{interference term}}. \end{aligned} \quad (13.81)$$

The interference terms appear as a result of particle statistics. The angles showing up in (13.81) are illustrated in figure 13.5. The interference terms for bosons produce a doubling of the cross section when $\theta = \pi/2$ as compared with the classical result,

$$\theta = \frac{\pi}{2} : \quad \frac{d\sigma}{d\Omega} = 4|f(\pi/2)|^2. \quad (13.82)$$

For a radially symmetric potential $V(r)$, we make use of a partial wave

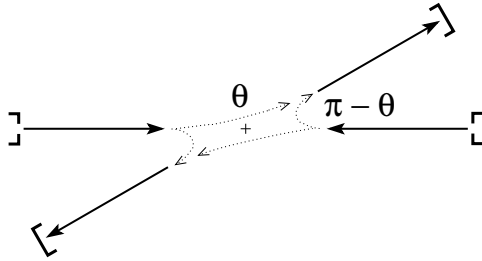


Fig. 13.5: The symmetrization of the scattering wavefunction generates two trajectories with scattering angles θ and $\pi - \theta$ which have to be added coherently.

representation,

$$f(\theta) = \sum_l i^l (2l + 1) f_l P_l(\cos \theta)$$

$$\begin{aligned} \downarrow P_l(\cos \theta) &= (-1)^l P_l(\cos(\pi - \theta)) \\ f(\theta) + f(\pi - \theta) &= 2 \sum_{l \text{ even}} i^l (2l + 1) f_l P_l(\cos \theta), \end{aligned} \quad (13.83)$$

and find that there are only even angular momenta l (for odd l the Legendre polynomial changes sign and the contributions cancel one another).

Spin-1/2 Fermion: In the case of spin-1/2 fermions, the scattering depends on the symmetry of the total spin wave function:

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

$$\psi(\vec{r}) = \psi(-\vec{r}) \quad (13.84)$$

is symmetric, hence, the cross section is equal to that of spin 0 bosons,

$$\left. \frac{d\sigma}{d\Omega} \right|_s = |f(\theta) + f(\pi - \theta)|^2. \quad (13.85)$$

2. The spin-triplet states

$$\chi_t = \begin{cases} |\uparrow\uparrow\rangle \\ [|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle]/\sqrt{2} \\ |\downarrow\downarrow\rangle \end{cases} \quad (13.86)$$

are symmetric and enforce an anti-symmetric orbital wave function $\psi(\vec{r}) = -\psi(-\vec{r})$; the scattering amplitude then takes the form $f(\theta) \rightarrow f(\theta) - f(\pi - \theta)$ which contains only odd angular momentum l . The cross section

$$\left. \frac{d\sigma}{d\Omega} \right|_t = |f(\theta) - f(\pi - \theta)|^2 \quad (= 0 \text{ for } \theta = \pi/2). \quad (13.87)$$

Note that polarized fermions scatter only in odd angular momentum channels. This has quite dramatic consequences in cold atomic physics. E.g., the scattering of cold atoms is usually described by a contact potential which results from the s -wave scattering. As a result, bosonic atoms interact through s -channel scattering, while cold polarized fermions interact only weakly since symmetry tells that they can only scatter in the p -channel.

Finally, for a statistical ensemble in an unpolarized beam of fermions, the weighted average results in a scattering cross section

$$\begin{aligned}\frac{d\sigma}{d\Omega} &= \frac{3}{4} \frac{d\sigma}{d\Omega} \Big|_t + \frac{1}{4} \frac{d\sigma}{d\Omega} \Big|_s \\ &= |f(\theta)|^2 + |f(\pi - \theta)|^2 - \text{Re}[f^*(\theta)f(\pi - \theta)].\end{aligned}\quad (13.88)$$

13.6.11 Molecular spectra

We consider the low-energy rotation spectra of a molecular cloud with energies $E_{\text{rot}} = \hbar^2 l(l+1)/2\Theta \ll E_{\text{elektron}} \sim eV$, Θ the moment of inertia; the time scale corresponding to these low rotation energies is slow and we can regard the electron shells as rigid (the electron shell should be understood as a whole object, not as a cloud of individual electrons). We consider two examples of molecules with bosonic and fermionic nuclei:

- $(C^{12})_2$ -Molecules, where the nuclei are spin = 0 bosons and therefore only even l are allowed.
- H_2 -Molecules, where the nuclei are spin-1/2 fermions and hence the spin wave functions are either singlet or triplet states,

$$\begin{aligned}\chi &= \chi_s : l = \text{even, para-hydrogen,} \\ \chi &= \chi_t : l = \text{odd, ortho-hydrogen.}\end{aligned}\quad (13.89)$$

The conversion of ortho-hydrogen to para-hydrogen is difficult (the nuclei are well protected) and one obtains two types of gases with different rotation spectra,

$$\begin{aligned}E_{\text{rot,para}}/\hbar^2 &= 0, \frac{3}{\Theta}, \frac{10}{\Theta}, \frac{21}{\Theta}, \dots, \\ E_{\text{rot,ortho}}/\hbar^2 &= \frac{1}{\Theta}, \frac{6}{\Theta}, \frac{15}{\Theta}, \dots.\end{aligned}\quad (13.90)$$

13.7 Spin and race (german ‘Rasse’)

The analysis of the $n \times$ spin-1/2 fermion problem is relevant in the discussion of atoms and very instructive with regard to the understanding the role of permutation symmetry. The topic is also a nice application of group theory, is fun, but also somewhat non-trivial. Here is an attempt to present the

problem in simple terms. The key 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.

13.7.1 Two spin-1/2 fermions, an illustration

We separate the total wave function into spin and orbital components, $\Psi(x_1, x_2) = \psi(\vec{r}_1, \vec{r}_2) \chi(s_1, s_2)$. The (total) spin determines the spin symmetry

$$\begin{aligned} S = 1 : \quad \chi_t &= \begin{cases} |\uparrow\uparrow\rangle \\ [|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle]/\sqrt{2}, & \text{symmetric under } S_2, \\ |\downarrow\downarrow\rangle \end{cases} \\ S = 0 : \quad \chi_s &= [|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle]/\sqrt{2}, \quad \text{anti-symm. under } S_2. \end{aligned} \quad (13.91)$$

The fermionic anti-symmetry then defines the orbital symmetry,

$$\begin{aligned} S = 1 &\rightarrow \chi_t \text{ symmetric} &\rightarrow \varphi_a \text{ anti-symmetric}, \\ S = 0 &\rightarrow \chi_s \text{ anti-symmetric} &\rightarrow \varphi_s \text{ symmetric}. \end{aligned}$$

The theorem of spin and race generalizes this statement for a system of n fermions. The novelty is that other irreducible representations of S_n are now appearing, different from the one-dimensional representations S and A . In the end, the total spin (under rotation symmetry $SO(3)/SU(2)$) determines the spin symmetry (under exchange S_n) and via Pauli principle the orbital symmetry (under S_n). The spin and orbital races are the ‘quantum numbers’ belonging to the irreducible representation of the permutation symmetry (same as l is the angular momentum quantum number belonging to the irreducible representation of the rotation symmetry $SO(3)/SU(2)$). The spin and orbital races thus determine under which irreducible representation of S_n the spin and orbital functions transform. We anticipate the result for n spin-1/2 fermions:

$$\begin{aligned} (\mathcal{H}_{1/2})^{\otimes n} &= \oplus_s \mu_s(n) \mathcal{H}_s, & s &= \text{spin quantum number}, \\ &= \oplus_s (2s+1) \mathcal{H}_{\mu_s(n)}. \end{aligned} \quad (13.92)$$

The first line describes the reduction of $(\mathcal{H}_{1/2})^{\otimes n}$ under the rotation group $SO(3)/SU(2)$ and tells that the $(2s+1)$ -dimensional irreducible representation living in the Hilbert space \mathcal{H}_s appears $\mu_s(n)$ times. E.g., $(\mathcal{H}_{1/2})^{\otimes 4} = (\mathcal{H}_0 \oplus \mathcal{H}_1) \otimes (\mathcal{H}_0 \oplus \mathcal{H}_1) = 2\mathcal{H}_0 \oplus 3\mathcal{H}_1 \oplus \mathcal{H}_2$ and hence $\mu_0(4) = 2$, $\mu_1(4) = 3$,

$\mu_2(4) = 1$. The $4^2 = 16$ -dimensional Hilbert space decomposes into 2 one-dimensional, 3 three-dimensional, and 1 five-dimensional Hilbert spaces (blocks), totalling up to $2 \times 1 + 3 \times 3 + 1 \times 5 = 16$ dimensions.

The second line is the reduction of $(\mathcal{H}_{1/2})^{\otimes n}$ under the permutation group S_n and tells, that the μ_s -dimensional irreducible representation of S_n that lives in $\mathcal{H}_{\mu_s(n)}$ appears exactly $(2s+1)$ times. E.g., $(\mathcal{H}_{1/2})^{\otimes 4} = \mathcal{H}_{2'} \oplus 3\mathcal{H}_{3'} \oplus 5\mathcal{H}_{1'}$ (the prime marking spaces belonging to the reduction under S_4), again totalling to $1 \times 2 + 3 \times 3 + 5 \times 1 = 16$ dimensions. The blocks associated with these two decompositions are mutually non-mixing, i.e., the block $2'$ contains only spin-0 blocks, the three $3'$ blocks contain only spin-1 blocks, and the five $1'$ blocks include only the spin-2 block, see (13.122).

From the second decomposition, we learn that the spin s determines the spin race μ_s , that is, it determines which irreducible representation of S_n describes the transformation of the spin wavefunction under permutation of its arguments. If the spin s occurs μ_s times in the decomposition of $(\mathcal{H}_{1/2})^{\otimes n}$, as told by the first line, the second line tells us that the μ_s -dimensional irreducible representation of S_n appears $(2s+1)$ times, i.e., the irreducible representation of S_n with race μ_s that belongs to a specific Young tableau occurs just $(2s+1)$ times.¹⁴ Furthermore, the spin race μ_s uniquely determines the orbital race μ_o via the condition of anti-symmetry for the total wavefunction Ψ . The races μ_s and μ_o are associated with conjugated Young tableaux.

13.8 Formalism of spin and race* (german ‘Rasse’)

This section explains and illustrates the decomposition of $(\mathcal{H}_{1/2})^{\otimes n}$ according to $SU(2)$ and S_n and the theorem of spin and race. As a byproduct, we will see how nature realizes only a small selection out of a multitude of possible permutation symmetries of the wavefunctions. In this discussion we need the:

¹⁴We can understand a Young tableau as the identifier for an irreducible representation of the permutation group S_n , as the angular momentum l is the identifier for \mathcal{D}^l , the $(2l+1)$ -dimensional representation of the rotation group $SO(3)$.

13.8.1 Irreducible representations of S_n

We consider the decomposition of permutations into cycles, for example

$$\begin{aligned} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} &= (123)(4) \\ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} &= (14)(23) \\ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix} &= (1423) \end{aligned} \quad (13.93)$$

$$\begin{aligned} S_5: \quad 5 &= 5 & \begin{array}{|c|c|c|c|c|} \hline \square & \square & \square & \square & \square \\ \hline \end{array} & [5] \\ &= 4+1 & \begin{array}{|c|c|c|c|} \hline \square & \square & \square & \square \\ \hline \square & & & \end{array} & [4,1] \\ &= 3+2 & \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \end{array} & [3,2] \\ &= 3+1+1 & \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & & \\ \hline \square & & \end{array} & [3,1^2] \\ &= 2+2+1 & \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \square & \end{array} & [2^2,1] \\ &= 2+1+1+1 & \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \\ \hline \square & \\ \hline \square & \end{array} & [2,1^3] \\ &= 1+1+1+1+1 & \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \square \\ \hline \square \end{array} & [1^5] \end{aligned}$$

Each decomposition of n defines an equivalence class of S_n . For example, for $n = 5$, we find 7 classes with the decompositions $5 = 5$, $4 + 1$, $3 + 2$, $3 + 1 + 1$, $2 + 2 + 1$, $2 + 1 + 1 + 1$, $1 + 1 + 1 + 1 + 1$. For each of these decompositions/partitions, we define a Young diagram with h rows and k columns as above.

Each partition or class creates an irreducible representation of S_n . Each partition has an associated partition,

$$\text{partition } \lambda \rightarrow \text{associated partition } \tilde{\lambda} \quad (13.94)$$

and an associated Young diagram,

$$\text{Young Diagram } Y \rightarrow \text{associated Young Diagram } \tilde{Y} \quad (13.95)$$

which results from swapping rows and columns, as illustrated with the following example with the corresponding mirror axis,

$$\begin{array}{ccc}
 \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & & \\ \hline \end{array} & \rightarrow & \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \square & \\ \hline \end{array} \\
 Y_\lambda & & \tilde{Y}_\lambda = Y_{\tilde{\lambda}} \\
 \lambda = [3, 2] & & \tilde{\lambda} = [2^2, 1]
 \end{array} \quad (13.96)$$

In addition to the Young diagrams Y , we define the Young tableau by filling in the diagrams with numbers $1, 2, \dots, n$:

$$\begin{array}{ccc}
 \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & & \\ \hline \end{array} & \rightarrow & \begin{array}{|c|c|c|} \hline 4 & 2 & 5 \\ \hline 1 & 3 & \\ \hline \end{array} \\
 & & \theta_\lambda \\
 & & \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 4 & 5 & \\ \hline \end{array} \\
 & & \rightarrow \begin{array}{|c|c|} \hline 4 & 1 \\ \hline 2 & 3 \\ \hline 5 & \\ \hline \end{array} \\
 & & \tilde{\theta}_\lambda \quad \text{associated tableaux}
 \end{array}$$

For each Young tableau, we define the line-symmetrizer (these are operators associated with the representations of S_n , actually U_{s_λ} and U_{a_λ} ; we simplify the notation $U_s \rightarrow s$, etc.)

$$s_\lambda = \frac{1}{\lambda_1! \dots \lambda_h!} \sum_{\pi} \pi_\lambda \quad (13.97)$$

and the anti-symmetrizer

$$a_\lambda = \frac{1}{\lambda_1! \dots \lambda_h!} \sum_{\pi} (-1)^\pi \pi_\lambda. \quad (13.98)$$

Here, π_λ only permutes the elements within but not between the rows, i.e., each cycle in π_λ contains only items from a single row. For better illustration, we look at the $\lambda = [3, 2]$ tableau,¹⁵

$$\begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 4 & 5 & \\ \hline \end{array}, \quad (13.99)$$

then s_λ is given by

$$\begin{aligned} s_\lambda = \frac{1}{3!2!} & \left[(1)(2)(3)(4)(5) + (12)(3)(4)(5) + (13)(2)(4)(5) \right. \\ & + (1)(23)(4)(5) + \\ & \left. + (123)(4)(5) + \text{all with } (45) \right]. \end{aligned} \quad (13.100)$$

In particular, we have $s_{\lambda=[n]} = s$, $a_{\lambda=[n]} = a$. The associated (column-) symmetrizer permutes elements in the column,

$$\begin{aligned} \tilde{s}_\lambda &= \frac{1}{\tilde{\lambda}_1! \cdots \tilde{\lambda}_k!} \sum_{\tau} \tau_\lambda, \\ \tilde{a}_\lambda &= \frac{1}{\tilde{\lambda}_1! \cdots \tilde{\lambda}_k!} \sum_{\tau} (-1)^\tau \tau_\lambda, \end{aligned} \quad (13.101)$$

with the permutation τ_λ leaving the columns of the Young tableaux θ_λ invariant, i.e., τ_λ permutes only elements within the columns. Taking the above example $\lambda = [3, 2]$, we have the conjugated partition $\tilde{\lambda} = [2^2, 1]$ and

$$\tilde{s}_\lambda = \frac{1}{2!2!1!} \left[(1)(2)(3)(4)(5) + (14)(2)(3)(5) + (14)(25)(3) + (1)(4)(25)(3) \right].$$

In particular, we find that $\tilde{s}_{\lambda=[1^n]} = s$, $\tilde{a}_{\lambda=[1^n]} = a$.

The irreducible symmetrizers

$$\begin{aligned} i_\lambda &= s_\lambda \tilde{a}_\lambda \quad (\text{irreducible symmetrizer}) \\ j_\lambda &= \tilde{a}_\lambda s_\lambda \quad (\text{irreducible anti-symmetrizer}) \end{aligned} \quad (13.102)$$

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 s and

¹⁵for the $\lambda = [3, 2]$ tableau, we have $\lambda_1 = 3$ and $\lambda_2 = 2$

a , which define the 1-dimensional representations S and A . The following relations hold true¹⁶

$$\begin{aligned}
 i_{\lambda=[n]} &= s_{\lambda=[n]} \tilde{a}_{\lambda=[n]} = s \cdot e = s, \\
 j_{\lambda=[n]} &= \tilde{a}_{\lambda=[n]} s_{\lambda=[n]} = e \cdot s = s, \\
 i_{\lambda=[1^n]} &= s_{\lambda=[1^n]} \tilde{a}_{\lambda=[1^n]} = e \cdot a = a, \\
 j_{\lambda=[1^n]} &= a.
 \end{aligned} \tag{13.103}$$

In order to find the dimension of the irreducible representation belonging to the diagram Y , we write down all standard tableaux to Y : in the standard tableau, we always fill in numbers from left to right, from top to bottom; again we consider S_5 as an example,

	1 2 3 4 5	dim = 1, s, S
	$ \begin{array}{cccc cccc} 1 & 2 & 3 & 4 & 1 & 2 & 3 & 5 \\ 5 & & & & 4 & & & \\ 1 & 2 & 4 & 5 & 1 & 3 & 4 & 5 \\ 3 & & & & 2 & & & \end{array} $	dim = 4
	$ \begin{array}{ccc ccc} 1 & 2 & 3 & 1 & 2 & 4 \\ 4 & 5 & & 3 & 5 & \\ 1 & 2 & 5 & 1 & 3 & 4 & 1 & 3 & 5 \\ 3 & 4 & & 2 & 5 & & 2 & 4 & \end{array} $	dim = 5
	$ \begin{array}{ccc ccc ccc} 1 & 2 & 3 & 1 & 2 & 4 & 1 & 2 & 5 \\ 4 & & & 3 & & & 3 & & \\ 5 & & & 5 & & & 4 & & \\ 1 & 3 & 4 & 1 & 3 & 5 & 1 & 4 & 5 \\ 2 & & & 2 & & & 2 & & \\ 5 & & & 4 & & & 3 & & \end{array} $	dim = 6
	$ \begin{array}{cc cc} 1 & 2 & 1 & 2 \\ 3 & 4 & 3 & 5 \\ 5 & & 4 & \end{array} $	

¹⁶ e is identity.

	$\begin{array}{c c c} 1 & 3 & 1 & 3 & 1 & 4 \\ 2 & 4 & 2 & 5 & 2 & 5 \\ 5 & & 4 & & 3 & \end{array}$	dim = 5
$\begin{array}{ c c } \hline \square & \square \\ \hline \square & \\ \hline \square & \\ \hline \end{array}$	$\begin{array}{c c c c} 1 & 2 & 1 & 3 & 1 & 4 & 1 & 5 \\ 3 & & 2 & & 2 & & 2 & \\ 4 & & 4 & & 3 & & 3 & \\ 5 & & 5 & & 5 & & 4 & \end{array}$	dim = 4
$\begin{array}{ c } \hline \square \\ \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array}$	$\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array}$	dim = 1, a, A

The first representation is always the so-called 'normal order tableau'. As expected, we find the usual relationship between the order of the group $n!$, the number of classes K , and the dimensions d_i of the irreducible representations,

$$n! = \sum_{i=1}^K d_i^2 : \quad 5! = 120 = 1 + 4^2 + 5^2 + 6^2 + 5^2 + 4^2 + 1. \quad (13.104)$$

Next, we use the above structures to define orbital and spin functions with well-defined symmetry under interchanges.

13.8.2 Orbital wavefunctions with defined symmetry

Let $u(x)$, $v(x)$, $w(x)$, ... be one-particle wavefunctions. We are looking for many-body wavefunctions of the type¹⁷

$$\Psi(1, \dots, n) = \sum \cdots u(\cdot) v(\cdot) w(\cdot) \quad (13.105)$$

with a well-defined race, that means, with a defined symmetry under permutation. We assume that the functions $u(x)$, $v(x)$, $w(x)$, ... are orthonormal. In the following, we consider the cases $n = 2$ and $n = 3$.

$n = 2$: Y is either $Y = \square\square$ thus S , or $Y = \begin{array}{|c|} \hline \square \\ \hline \end{array}$ and thus A . We obtain the basis vectors of the irreducible representation by applying the (anti)symmetrizers to the product state $u(1)v(2)$ (we fix the order of the wave

¹⁷Here, we discuss the case where $u \neq v \neq w \neq u \dots$

functions $u(\cdot)v(\cdot)w(\cdot)$ and permute the coordinates, we abbreviate $e = (1)(2)$,

$$\begin{aligned}
S : \square\square \quad i_{[2]} &= s_{[2]} = [e + (12)]/2 \\
\sqrt{2} i_{[2]} \varphi &= [u(1)v(2) + u(2)v(1)]/\sqrt{2} = \varphi_s \\
A : \square \quad i_{[1^2]} &= \tilde{a}_{[1^2]} = [e - (12)]/2 \\
\sqrt{2} i_{[1^2]} \varphi &= [u(1)v(2) - u(2)v(1)]/\sqrt{2} = \varphi_a. \quad (13.106)
\end{aligned}$$

$n = 3$: There are three possible tableaux Y , $Y = \square\square\square = S$, $Y = \begin{smallmatrix} \square \\ \square \end{smallmatrix} = A$, and the third variant $Y = \begin{smallmatrix} \square & \square \end{smallmatrix}$ which we denote by Z .¹⁸

$$\begin{aligned}
S : \begin{smallmatrix} \square & \square & \square \end{smallmatrix} \quad i_{[3]} &= [e + (12) + (23) + (13) + (123) + (132)]/6 \\
\sqrt{6} i_{[3]} \varphi &= [u(1)v(2)w(3) + u(2)v(1)w(3) + u(1)v(3)w(2) \\
&\quad + u(3)v(2)w(1) + u(3)v(1)w(2) + u(2)v(3)w(1)]/\sqrt{6} \\
&= \varphi_s, \\
A : \begin{smallmatrix} \square \\ \square \\ \square \end{smallmatrix} \quad i_{[1^3]} &= [e - (12) - (23) - (13) + (123) + (132)]/6 \\
\sqrt{6} i_{[1^3]} \varphi &= [u(1)v(2)w(3) - u(2)v(1)w(3) - u(1)v(3)w(2) \\
&\quad - u(3)v(2)w(1) + u(2)v(3)w(1) + u(3)v(1)w(2)]/\sqrt{6} \\
&= \varphi_a, \\
Z : \begin{smallmatrix} \square & \square \\ \square \end{smallmatrix} \quad i_{[21]} &= s_{[21]} \tilde{a}_{[21]} \\
&= \frac{1}{2!1!} [e + (12)] \frac{1}{2!1!} [e - (13)] \\
&= [e + (12) - (13) - (132)]/4 = e_z/4 \\
\sqrt{4} i_{[21]} \varphi &= [u(1)v(2)w(3) + u(2)v(1)w(3) \\
&\quad - u(3)v(2)w(1) - u(2)v(3)w(1)]/2 \\
&= \varphi_z, \\
\begin{smallmatrix} \square & \square \\ \square \end{smallmatrix} \quad i'_{[21]} &= [e + (13) - (12) - (123)]/4 = e'_z/4 \\
\sqrt{4} i'_{[21]} \varphi &= [u(1)v(2)w(3) + u(3)v(2)w(1) \\
&\quad - u(2)v(1)w(3) - u(3)v(1)w(2)]/2 \\
&= \varphi'_z. \quad (13.107)
\end{aligned}$$

¹⁸Note that $U_\pi \varphi(1, 2, 3) = \varphi(\pi_1^{-1}, \pi_2^{-1}, \pi_3^{-1})$ with $\varphi(1, 2, 3) = u(1)v(2)w(3)$ and the cycles (123) and (132) are mutually inverse.

Note the definition of the cycle notation in (13.26) and (13.29), in particular

$$(12)(13) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 3 & 1 & 2 \end{pmatrix} = (132). \quad (13.108)$$

Applying the group to φ_z creates a two-dimensional representation of S_3 ; we write $4i_{[21]} = e_z$ and find that

$$\begin{aligned} e e_z &= e_z, \\ (12)e_z &= e_z, \\ (23)e_z &= (23) + (132) - (123) - (12) \equiv \bar{e}_z, \\ (31)e_z &= -e_z - \bar{e}_z, \\ (123)e_z &= -e_z - \bar{e}_z, \\ (132)e_z &= \bar{e}_z. \end{aligned} \quad (13.109)$$

Thus $e_z = e + (12) - (13) - (123)$ and $\bar{e}_z = (23) + (132) - (123) - (12)$ or $\varphi_z = e_z \varphi / 2$ and $\bar{\varphi}_z = \bar{e}_z \varphi / 2$, respectively, span a 2-dimensional irreducible representation (we call it Z) of S_3 .

Moreover,

$$\begin{aligned} e'_z &= e + (13) - (12) - (123), \\ \bar{e}'_z &= (23)e'_z = (23) + (123) - (132) - (13) \end{aligned} \quad (13.110)$$

span a second 2-dimensional irreducible representation Z' of S_3 ; since $e'_z e_z = 0$, this representation is different from Z . The reduction of the 6-dimensional representation space spanned by the product wavefunctions $\varphi(\pi_1^{-1}, \pi_2^{-1}, \pi_3^{-1})$, $\pi \in S_3$, $\pi \in S_3$ thus generates the decomposition

$$S \oplus A \oplus Z \oplus Z'. \quad (13.111)$$

The standard tableaus define each a generating vector in every one of these irreducible representations.

13.8.3 Spin functions with defined symmetry

We consider the spin problem for n Spin = 1/2 fermions. The states χ_\uparrow and χ_\downarrow span the Hilbert space $\mathcal{H}_{1/2}$ and the product functions

$$\chi_{\sigma_1} \otimes \chi_{\sigma_2} \otimes \cdots \chi_{\sigma_n} \quad (13.112)$$

define a basis in the 2^n -dimensional product space $(\mathcal{H}_{1/2})^{\otimes n}$. The operators S^2 , S_z , $\pi \in S_n$, where $\vec{S} = \sum_1^n \vec{s}_i$ denotes the total spin, commute pairwise and can thus be ‘diagonalized’ simultaneously, hence, we can jointly reduce the representations of $SU(2)$ (with quantum numbers s for S^2 and m for S_z) and S_n (with quantum numbers given by standard Young tableaux).

SU(2): $\mathcal{H}_{1/2}^{\otimes n} = \oplus_s \mu_s(n) \mathcal{H}_s$ according to the rules of spin addition. For illustration, we again consider the cases $n = 2, \dots, 5$,

$$\begin{aligned}
 n = 2: \quad \left(\mathcal{H}_{\frac{1}{2}}\right)^{\otimes 2} &= \mathcal{H}_0 \oplus \mathcal{H}_1, \\
 &\mu_0(2) = \mu_1(2) = 1. \\
 n = 3: \quad \left(\mathcal{H}_{\frac{1}{2}}\right)^{\otimes 3} &= \left(\mathcal{H}_0 \oplus \mathcal{H}_1\right) \otimes \mathcal{H}_{\frac{1}{2}}, \\
 &= 2\mathcal{H}_{\frac{1}{2}} \oplus \mathcal{H}_{\frac{3}{2}}, \\
 &\mu_{\frac{1}{2}}(3) = 2, \quad \mu_{\frac{3}{2}}(3) = 1. \\
 n = 4: \quad \left(\mathcal{H}_{\frac{1}{2}}\right)^{\otimes 4} &= \left(\mathcal{H}_{\frac{1}{2}} \oplus \mathcal{H}_{\frac{1}{2}} \oplus \mathcal{H}_{\frac{3}{2}}\right) \otimes \mathcal{H}_{\frac{1}{2}}, \\
 &= 2\mathcal{H}_0 \oplus 3\mathcal{H}_1 \oplus \mathcal{H}_2, \\
 &\mu_0(4) = 2, \quad \mu_1(4) = 3, \quad \mu_2(4) = 1. \\
 n = 5: \quad \left(\mathcal{H}_{\frac{1}{2}}\right)^{\otimes 5} &\mu_{\frac{1}{2}}(5) = 5, \quad \mu_{\frac{3}{2}}(5) = 4, \quad \mu_{\frac{5}{2}}(5) = 1. \quad (13.113)
 \end{aligned}$$

Note that,

$$\sum_s \mu_s(n) \cdot (2s + 1) = 2^n, \quad \mu_{\frac{n}{2}}(n) = 1. \quad (13.114)$$

S_n: In the spin problem, there are only two basis vectors χ_{\uparrow} and χ_{\downarrow} . This means that the Young tableaux can have a maximum of 2 rows, otherwise the associated projector disappears on the product basis. For example, if we had 3 rows,

$$\begin{array}{|c|}
 \hline
 \uparrow \\
 \hline
 \downarrow \\
 \hline
 \downarrow \\
 \hline
 \end{array}$$

then the anti-symmetrization of the two spin \downarrow states would result in a 0. Furthermore, the state \uparrow must be paired with a state \downarrow in every column, otherwise there is again a vanishing result. For example, the result of the

operation $\begin{smallmatrix} 1 & 2 \\ 3 \end{smallmatrix}$ applied to $\chi_\uparrow\chi_\downarrow\chi_\uparrow$ with 1 & 3 both in the state \uparrow (we fix the coordinates 1, 2, 3 and permute the quantum numbers of the wavefunctions, here the spin variables $\uparrow \leftrightarrow +$, $\downarrow \leftrightarrow -$),

$$\begin{aligned} & [e + (12) - (13) - (132)][+, -, +] \\ & = [+ - +] + [- + +] - [+ - +] - [- + +] = 0. \end{aligned} \quad (13.115)$$

Finally, the representation space $\mathcal{H}_{1/2}^{\otimes n}$ contains the $\mu_s(n)$ -dimensional representation of S_n just $(2s + 1)$ times: the vectors

$$\xi_m = \chi_\uparrow(\pi_1) \cdots \chi_\uparrow(\pi_\nu) \chi_\downarrow(\pi_{\nu+1}) \cdots \chi_\downarrow(\pi_n) \quad (13.116)$$

with

$$S_z = m = \nu \cdot \frac{1}{2} - (n - \nu) \cdot \frac{1}{2} = \nu - \frac{n}{2} \quad (13.117)$$

span the $n!/\nu!(n - \nu)!$ -dimensional vector space of a fixed S_z .¹⁹ Within the space spanned by the ξ_m , m fixed, we can search for subspaces orthogonal to S^2 . Allowed values of s , which can contain m , are

$$s = \frac{n}{2}, \quad \frac{n}{2} - 1, \quad \dots, \quad |m| + 1, \quad |m|, \quad (13.118)$$

which are just $n/2 - |m| + 1$ values. The subspace belonging to a fixed pair s, m appears $\mu_s(n)$ times and defines a $\mu_s(n)$ -dimensional representation of S_n : $\pi \in S_n$ changes neither s nor m . The subspace in $\mathcal{H}_{1/2}^{\otimes n}$ for fixed s, m thus creates a representation space of S_n ; the representation belongs to the Young tableau $[n/2 + s, n/2 - s]$ with the dimension $\mu_s(n)$ and is therefore irreducible. By applying S_+ and S_- (both commute with $\pi \in S_n$), we find $2s + 1$ equivalent representations and thus $\mathcal{H}_{1/2}^{\otimes n}$ decomposes according to

$$\mathcal{H}_{\frac{1}{2}}^{\otimes n} = \bigoplus_s (2s + 1) \mathcal{H}_{\mu_s(n)}, \quad (13.119)$$

with $\mathcal{H}_{\mu_s(n)}$ the irreducible representation belonging to the Young diagram

$$\begin{array}{c} n/2 + s \\ \hline \boxed{} \\ \hline \boxed{} \\ n/2 - s \end{array}.$$

¹⁹Only permutations that occur outside of the \uparrow and \downarrow groups change ξ_m .

We check this result on the usual examples $n = 2, \dots, 5$,

$$\begin{aligned}
n = 2 : \quad & \begin{array}{|c|c|} \hline & \\ \hline \end{array} & \dim = 1, \quad s = 1, \text{ triplet } \begin{cases} s(\uparrow\uparrow), \\ s(\uparrow\downarrow), \\ s(\downarrow\downarrow), \end{cases} \\
& \begin{array}{|c|} \hline \\ \hline \end{array} & \dim = 1, \quad s = 0, \text{ singlet } \{ a(\uparrow\downarrow), \\
& & \mathcal{H}_{\frac{1}{2}}^{\otimes 2} = 3\mathcal{H}_{1'} \oplus \mathcal{H}_{1'}. \\
n = 3 : \quad & \begin{array}{|c|c|c|} \hline & & \\ \hline \end{array} & \dim = 1, \quad s = 3/2, \\
& \begin{array}{|c|c|} \hline & \\ \hline \end{array} & \dim = 2 = \mu_{\frac{1}{2}}(3), \quad s = 1/2, \\
& & \mathcal{H}_{\frac{1}{2}}^{\otimes 3} = 3\mathcal{H}_{1'} \oplus 2\mathcal{H}_{2'}. \\
n = 4 : \quad & \begin{array}{|c|c|c|c|} \hline & & & \\ \hline \end{array} & \dim = 1 = \mu_2(4), \quad s = 2, \\
& \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 4 & & \end{array} \begin{array}{|c|c|c|} \hline 1 & 3 & 4 \\ \hline 2 & & \end{array} \begin{array}{|c|c|c|} \hline 1 & 2 & 4 \\ \hline & & 3 \end{array} \\
& & \dim = 3 = \mu_1(4), \quad s = 1, \\
& \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 3 & 4 \end{array} \begin{array}{|c|c|} \hline 1 & 3 \\ \hline 2 & 4 \end{array} \\
& & \dim = 2 = \mu_0(4), \quad s = 0, \\
& & \mathcal{H}_{\frac{1}{2}}^{\otimes 4} = 5\mathcal{H}_{1'} \oplus 3\mathcal{H}_{3'} \oplus \mathcal{H}_{2'}. \\
n = 5 : \quad & \begin{array}{|c|c|c|c|c|} \hline & & & & \\ \hline \end{array} & \mu_{\frac{5}{2}}(5) = 1, \\
& \begin{array}{|c|c|c|c|} \hline & & & \\ \hline \end{array} & \mu_{\frac{3}{2}}(5) = 4, \\
& \begin{array}{|c|c|c|} \hline & & \\ \hline \end{array} & \mu_{\frac{1}{2}}(5) = 5, \\
& & \mathcal{H}_{\frac{1}{2}}^{\otimes 5} = 5\mathcal{H}_{1'} \oplus 3\mathcal{H}_{4'} \oplus \mathcal{H}_{5'}.
\end{aligned} \tag{13.120}$$

In this way, we arrive at the decomposition

$$\mathcal{H}_{\frac{1}{2}}^{\otimes n} = \oplus_s \mu_s(n) \mathcal{H}_s = \oplus_s (2s + 1) \mathcal{H}_{\mu_s(n)}. \tag{13.121}$$

The spin s determines the Young tableau and thus the spin symmetry. The representation space $\mathcal{H}_{\mu_s(n)}$ is spanned by $|s, m\rangle$ with the same s and m values.

Finally, we present the specific example for four fermions with spin 1/2,

$$\mathcal{H}_{\frac{1}{2}}^{\otimes 4} \stackrel{\text{SU}(2)}{=} 2\mathcal{H}_0 \oplus 3\mathcal{H}_1 \oplus \mathcal{H}_2 \quad (13.122)$$

$$\stackrel{\text{S}_4}{=} \mathcal{H}_{2'} + 3\mathcal{H}_{3'} + 5\mathcal{H}_{1'}$$

$$= \left\{ \begin{array}{l} \left\{ \begin{array}{l} \mathcal{H}_0 : |0, 0\rangle \text{ ————— } \mathcal{H}_{2'} \\ \mathcal{H}'_0 : |0', 0'\rangle \text{ ————— } \end{array} \right. \\ \\ \left\{ \begin{array}{l} \begin{array}{l} |1, 1\rangle \text{ ————— } \mathcal{H}_{3'} \\ \mathcal{H}_1 : |1, 0\rangle \text{ ————— } \mathcal{H}_{3'} \\ |1, -1\rangle \text{ ————— } \mathcal{H}_{3'} \end{array} \\ \\ \begin{array}{l} |1', 1'\rangle \text{ ————— } \\ \mathcal{H}'_1 : |1', 0'\rangle \text{ ————— } \\ |1', -1'\rangle \text{ ————— } \end{array} \\ \\ \begin{array}{l} |1'', 1''\rangle \text{ ————— } \\ \mathcal{H}''_1 : |1'', 0''\rangle \text{ ————— } \\ |1'', -1''\rangle \text{ ————— } \end{array} \end{array} \right. \\ \\ \left\{ \begin{array}{l} |2, 2\rangle \text{ ————— } \mathcal{H}_{1'} \\ |2, 1\rangle \text{ ————— } \mathcal{H}_{1'} \\ \mathcal{H}_2 : |2, 0\rangle \text{ ————— } \mathcal{H}_{1'} \\ |2, -1\rangle \text{ ————— } \mathcal{H}_{1'} \\ |2, -2\rangle \text{ ————— } \mathcal{H}_{1'} \end{array} \right. \end{array} \right.$$

13.8.4 Spin & Orbit

Consider a spin wavefunction with a symmetry characterized by the diagram Y_s that has at most two rows. Let the diagram Y_o describe the symmetry of the orbital function. Then following statement applies: The tensor product of the two representations contains the representation

$$S \Leftrightarrow Y_o = Y_s$$

$$A \Leftrightarrow Y_o = \tilde{Y}_s. \quad (13.123)$$

It follows that the spin race given by Y_s defines the orbital race Y_o . Let the spin of the n -particle state be given by s . Then the symmetries of the spin and orbital wavefunctions are given by associated tableaus as illustrated in figure 13.6.

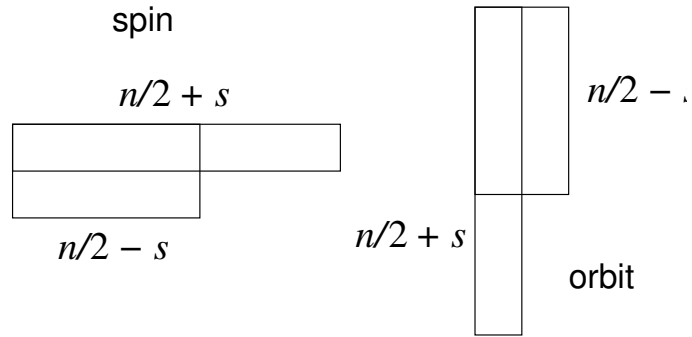


Fig. 13.6: Young tableaus associated with the spin and orbit function for n spin-1/2 fermions with total spin s .

This concludes the discussion of the theorem of spin and race.

Chapter 14

Atoms

The textbooks on quantum mechanics by Gordon Baym and Leonard Schiff cover the topics discussed in this chapter.

Here, we want to understand the structure of atoms, the periodic table, and the Grotrian (or term) diagram (i.e., the splitting of energy levels). We are particularly interested in the ground state configuration of a multi-electron atom, its spin S , orbital angular momentum L , and total angular momentum J ; knowing these three quantities provides us with the term symbol $^{2S+1}L_J$.

To get started, we consider the two-electron atom, Helium He with $Z = 2$ or the ions H^- ($Z = 1$) and Li^+ ($Z = 3$); these systems are described by the Hamiltonian

$$H = H_0 + H_C = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} - \frac{Ze^2}{r_1} - \frac{Ze^2}{r_2} + \frac{e^2}{|\vec{r}_1 - \vec{r}_2|}. \quad (14.1)$$

Due to the Coulomb interaction between the two electrons, $H_C = e^2/r_{12}$, $r_{12} = |\vec{r}_1 - \vec{r}_2|$, the Hamiltonian (14.1) is not exactly solvable. The simplest approximate solution handles this term within perturbation theory, that is, 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 to $H_0\Psi_0 = E_0\Psi_0$ are product wavefunctions $\Psi_0(\vec{r}_1, \vec{r}_2) = \varphi_{\alpha_1}(\vec{r}_1)\varphi_{\alpha_2}(\vec{r}_2)$, where $\varphi_{\alpha}(\vec{r})$ is a Hydrogen wavefunction for nuclear charge Z , see chapter 5, with $\alpha = nlm$. The associated energies are:

$$E_0 = E_{\alpha_1} + E_{\alpha_2} \quad (14.2)$$

with

$$E_\alpha = -\frac{Z^2}{n^2} E_R, \quad E_R = \frac{e^2}{2a_B} \approx 13.605 \text{ eV}. \quad (14.3)$$

Furthermore, we have to account for or Pauli exclusion for fermions, with the wave function

$$\Psi_0(x_1, x_2) = -\Psi_0(x_2, x_1) \quad (14.4)$$

antisymmetric in $x_i = (\vec{r}_i, s_i)$. The separation of spin and orbital parts allows to rewrite Ψ as a product

$$\Psi(x_1, x_2) = \Psi_0(\vec{r}_1, \vec{r}_2) \chi_0(s_1, s_2), \quad (14.5)$$

such that the anti-symmetry of the overall wave function can be constructed in two ways, either with an anti-symmetric spin function (spin singlet state)

$$\begin{cases} \Psi_0 & \text{symmetric,} \\ \chi_0 & \text{anti-symmetric, } S = 0, \text{ singlet} \end{cases} \quad (14.6)$$

or via

$$\begin{cases} \Psi_0 & \text{anti-symmetric,} \\ \chi_0 & \text{symmetric, } S = 1, \text{ triplet} \end{cases} \quad (14.7)$$

with a symmetric spin wave function (spin triplet state). For the triplet state, the orbital Ψ_0 must be antisymmetric, i.e., $\alpha_1 \neq \alpha_2$, which costs at least the additional energy

$$\Delta E = Z^2(1 - 1/2^2) E_R \approx Z^2 \cdot 10 \text{ eV}. \quad (14.8)$$

We thus choose the spin singlet state as our ground state candidate,

$$\begin{aligned} \Psi_0(x_1, x_2) &= \varphi_{100}(\vec{r}_1) \varphi_{100}(\vec{r}_2) \chi_{00}(s_1, s_2), \\ \varphi_{100}(r) &= (Z^3/\pi a_B^3)^{1/2} \exp(-Zr/a_B), \\ \chi_{00}(s_1, s_2) &= [\chi_\uparrow(s_1)\chi_\downarrow(s_2) - \chi_\downarrow(s_1)\chi_\uparrow(s_2)]/\sqrt{2}, \\ E_0 &= -2Z^2 E_R = -Z^2 e^2/a_B. \end{aligned} \quad (14.9)$$

The correction E_1 is obtained through a trivial integral, see Chap. 8 of QM I with $\alpha \rightarrow Z$,

$$\begin{aligned} E_1 &= \frac{Z^6}{a_B^6 \pi^2} \int d^3 r_1 d^3 r_2 \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} \exp[-2Z(r_1 + r_2)/a_B] \\ &= \frac{5}{8} \frac{Z e^2}{a_B}, \end{aligned} \quad (14.10)$$

and we find the following estimate for the ground state energy of the He atom ($Z = 2$)

$$E_G \approx -\left(Z - \frac{5}{8}\right) \frac{Ze^2}{a_B} \approx -5.5 \text{ Ry.} \quad (14.11)$$

The experimentally measured value is $E_G^{\text{He}} \approx -5.8 \text{ Ry}$, about 4 eV lower than our calculated value.

We can do better with the help of the variational method (we replace the nuclear charge Z in the wave function φ_{100} with the variation parameter α) to find the estimate (see Chap. 8)

$$\begin{aligned} E_G &\approx -\left(\alpha - 4 + \frac{5}{8}\right) \frac{\alpha e^2}{a_B} \Big|_{\alpha=27/16} = -\left(2 - \frac{5}{16}\right)^2 \frac{e^2}{a_B} \\ &\approx -5.7 \text{ Ry,} \end{aligned} \quad (14.12)$$

an improvement of $\sim 2.7\text{eV}$. Still, about 1.4eV out of a total of 79 eV are missing, a quite reasonable result though. The improvement due to replacing $Z = 2 \rightarrow \alpha = 27/16$ is a consequence of taking screening effects into account: each electron only ‘sees’ the reduced charge $2 - 5/16 = 27/16$ instead of $Z = 2$, as previously discussed in Chap. 8.

For the hydrogen ion H^- , we find (via the variational method) the result

$$E_G \approx -\left(\frac{11}{16}\right)^2 \frac{e^2}{a_B} \approx -12.86 \text{ eV.} \quad (14.13)$$

The configuration $\text{H} + e^-$ (hydrogen plus free electron) corresponds to the energy -13.6 eV , making H^- unstable; experimentally, however, H^- is stable, therefore our result deviates at least $3/4 \text{ eV}$ from the actual energy.

14.1 Hartree & Hartree-Fock Approximation

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

$$H = \sum_{i=1}^n \left[\frac{p_i^2}{2m} + U(\vec{r}_i) \right] + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\vec{r}_i - \vec{r}_j|} \quad (14.14)$$

(we neglect interactions involving spin). The potential $U(\vec{r})$ is a given (external) potential; for atoms of charge Z , the potential is $U_Z = -Ze^2/r$, for

molecules, we sum over the different nuclei in U_M (ions with charges Z_l and positions (\vec{R}_l)), and for solids, we sum over the positions \vec{R}_l in the Bravais lattice in U_G ,

$$U_M(\vec{r}) = - \sum_{l=1}^L \frac{Z_l e^2}{|\vec{r} - \vec{R}_l|}, \quad U_G(\vec{r}) = - \sum_l \frac{Z e^2}{|\vec{r} - \vec{R}_l|}. \quad (14.15)$$

Naively, we could again solve the one-particle problem first

$$h = \frac{p^2}{2m} + U(\vec{r}), \quad h\varphi_j(\vec{r}) = \mathcal{E}_j\varphi_j(\vec{r}), \quad (14.16)$$

define a many-electron wave function in the form of a suitable anti-symmetric product wavefunction (a Slater determinant (13.65)), and treat the interaction term in the Hamiltonian as a perturbation. The expected results will be mediocre, as we have seen from the example of the He atom: The electrons are not only subject to the potential $U(\vec{r})$, but also move in the (average) potential created by the other electrons, resulting in a screened nuclear potential U . This screening, in turn, depends on the orbits of the electrons, so we have to determine the orbits and the screening potential in a self-consistent manner. This is precisely the achievement of the Hartree & Hartree-Fock approximation: the Hartree approximation takes into account (only) the screening effects, the Hartree-Fock theory also includes the antisymmetry of the wave function.

14.1.1 Hartree Approximation

In the Hartree approximation, the task is to find the optimal product wave function

$$\Psi = \varphi_{\alpha_1}(x_1) \varphi_{\alpha_2}(x_2) \cdots \varphi_{\alpha_n}(x_n), \quad (14.17)$$

where we at least partially account for the Pauli exclusion by orthogonalizing the single-particle orbitals φ_{α_i} and avoiding double occupations. Mathematically, we are looking for one-particle functions φ_{α_i} such that the expected value of the energy $\langle \Psi | H | \Psi \rangle = \min$ is minimal; furthermore, the wave functions have to be normalized, $\| \varphi_{\alpha_i} \| = 1$. This constraint is enforced with the help of a Lagrange multiplier, such that we have to minimize the following functional¹

$$\mathcal{F}[\varphi_{\alpha_1}(x_1), \dots, \varphi_{\alpha_n}(x_n)] = \langle \Psi | H | \Psi \rangle - \sum_i \lambda_{\alpha_i} \langle \varphi_{\alpha_i} | \varphi_{\alpha_i} \rangle \quad (14.18)$$

¹we drop the unity in the constraint as we take derivatives.

$$\begin{aligned}
&= \int \prod_{i=1}^n dr_i \left\{ \varphi_{\alpha_1}^* \cdots \varphi_{\alpha_n}^* \left[\sum_i^n \frac{p_i^2}{2m} + U(\vec{r}_i) + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\vec{r}_i - \vec{r}_j|} \right] \varphi_{\alpha_1} \cdots \varphi_{\alpha_n} \right\} \\
&\quad - \sum_{i=1}^n \lambda_{\alpha_i} \int dr_i \varphi_{\alpha_i}^* \varphi_{\alpha_i} \\
&= \sum_{i=1}^n \int dr_i \varphi_{\alpha_i}^* \left[\frac{p_i^2}{2m} + U(\vec{r}_i) + \frac{1}{2} \sum_{j \neq i} \int dr_j \varphi_{\alpha_j}^* \frac{e^2}{|\vec{r}_i - \vec{r}_j|} \varphi_{\alpha_j} - \lambda_{\alpha_i} \right] \varphi_{\alpha_i}.
\end{aligned}$$

In (14.18) the spin of the electrons only appears in the index α_i but not explicitly; hence, the sum over the spin variables s_i is trivial, since the Hamiltonian does not contain any spin terms and $\chi_\sigma(s) = \delta_{\sigma s}$. At the minimum of \mathcal{F} , the variation $\delta\mathcal{F} = 0$ vanishes and we find the condition (from $\delta\mathcal{F}/\delta\varphi_{\alpha_i}^* = 0$)

$$\left[\frac{-\hbar^2}{2m} \nabla^2 + U(\vec{r}) + \int d^3r' \frac{e}{|\vec{r} - \vec{r}'|} \sum_j' e |\varphi_{\alpha_j}(\vec{r}')|^2 \right] \varphi_{\alpha_i}(\vec{r}) = \lambda_{\alpha_i} \varphi_{\alpha_i}(\vec{r}) \quad (14.19)$$

Here, the Lagrange parameters appear as eigenvalues λ_{α_i} that describe a one-particle energy \mathcal{E}_{α_i} (more precisely, roughly an ionization energy, see later). The sum over $\sum_j' e |\varphi_{\alpha_j}|^2$ describes the average charge distribution $\rho_i(\vec{r}')$ of the other electrons $j \neq i$. The system (14.19) defines the Hartree equations; with the optimization of the one-particle wave functions φ_{α_i} , these equations take into account the presence of the other electrons via the interaction term

$$\int d^3r' \frac{e \rho_i(\vec{r}')}{|\vec{r} - \vec{r}'|}. \quad (14.20)$$

Note that the φ_{α_i} are not orthogonal because $\rho_i(\vec{r})$ depends on the index i . We usually write $\rho_i(\vec{r}) \approx \rho(\vec{r})$ as a single electron is only a small perturbation of the $n \gg 1$ electron cloud. The density $\rho(\vec{r})$ then describes the atomic charge distribution as seen by an external test charge.

14.1.2 Hartree-Fock Approximation

The Hartree-Fock approximation goes one step further and minimizes the functional

$$\mathcal{F} = \langle \Psi | H | \Psi \rangle - \sum_i \lambda_{\alpha_i} \langle \varphi_{\alpha_i} | \varphi_{\alpha_i} \rangle \quad (14.21)$$

in the variational space spanned by the slater determinants

$$\Psi = \frac{1}{\sqrt{n!}} \begin{vmatrix} \varphi_{\alpha_1}(1) & \cdots & \varphi_{\alpha_1}(n) \\ \vdots & \ddots & \vdots \\ \varphi_{\alpha_n}(1) & \cdots & \varphi_{\alpha_n}(n) \end{vmatrix}. \quad (14.22)$$

The one-particle expectation values of p_i^2 and $U(\vec{r}_i)$ remain unchanged, while the interaction term $\langle V_{\text{int}} \rangle$ now generates two terms (the spin variables s_i appear explicitly since the slater determinant accounts for exchange correlations between particles with identical spins, while electrons with unequal spins show no such correlations),

$$\begin{aligned} \langle V_{\text{int}} \rangle = \frac{1}{2} \int d^3r d^3r' \frac{e^2}{|\vec{r} - \vec{r}'|} \sum_{i \neq j} \left[\varphi_{\alpha_i}^*(\vec{r}) \overbrace{\varphi_{\alpha_j}^*(\vec{r}') \varphi_{\alpha_j}(\vec{r}')}^{\rho(\vec{r}')/e} \varphi_{\alpha_i}(\vec{r}) \right. \\ \left. - \delta_{s_i s_j} \varphi_{\alpha_i}^*(\vec{r}) \varphi_{\alpha_j}^*(\vec{r}') \varphi_{\alpha_j}(\vec{r}) \varphi_{\alpha_i}(\vec{r}') \right]. \end{aligned} \quad (14.23)$$

The spin dependent term $\propto \delta_{s_i s_j} \varphi_{\alpha_i}^* \cdots$ is called the Fock- or exchange-term; its derivation involves the following steps: starting with the product state $\Psi_\pi = \prod_i \varphi_{\alpha_i}(x_i)$, we define the anti-symmetrized state $\Psi = \sqrt{N!} U_a \Psi_\pi = (1/\sqrt{N!}) \sum_p (-)^p p \Psi_\pi$ and calculate the matrix element²

$$\begin{aligned} \langle \Psi | V_{\text{int}} | \Psi \rangle &= N! \langle \Psi_\pi U_a V_{\text{int}} U_a \Psi_\pi \rangle \\ &\downarrow U_a V_{\text{int}} U_a = V_{\text{int}} U_a^2 = V_{\text{int}} U_a \\ &= \frac{1}{2} \sum_{j \neq i} \sum_p (-)^p \langle \Psi_\pi V_{\text{int}}^{ij} p \Psi_\pi \rangle \\ &= \frac{1}{2} \sum_{i \neq j} \langle \Psi_\pi | V_{\text{int}}^{ij} (1 - p_{ij}) | \Psi_\pi \rangle \end{aligned} \quad (14.24)$$

and with $\Psi_\pi = \psi_\pi \chi_\pi$ and the normalization $\langle \chi_\pi \chi_\pi \rangle = 1$, we obtain

$$\langle \chi_\pi p_{ij} \chi_\pi \rangle = \delta_{s_i s_j}. \quad (14.25)$$

The two terms in (14.24) then describe the Hartree and the exchange contributions.

²See (13.36) for the definition of the anti-symmetrization operator U_a ; here, we replace the permutations $\pi \in S_n$ with $p \in S_n$. In the sum over p , only the terms $p = e$ and p exchanging i and j survive.

As before, we vary $\delta\mathcal{F}/\delta\varphi_{\alpha_i}^* = 0$ and generate the Hartree-Fock integro-differential equations

$$\begin{aligned} \left[-\frac{\hbar^2}{2m}\vec{\nabla}^2 + U(\vec{r}) \int d^3r' \frac{e\rho_i(\vec{r}')}{|\vec{r}-\vec{r}'|} \right] \varphi_{\alpha_i}(\vec{r}) \\ - \sum_j \delta_{s_i s_j} \int d^3r' \left[\varphi_{\alpha_j}^*(\vec{r}') \frac{e^2}{|\vec{r}-\vec{r}'|} \varphi_{\alpha_i}(\vec{r}') \right] \varphi_{\alpha_j}(\vec{r}) = \lambda_{\alpha_i} \varphi_{\alpha_i}(\vec{r}). \end{aligned} \quad (14.26)$$

In the second line, the (unknown) eigenfunction φ_{α_i} is integrated over; the electron then is subject to a non-local potential.

The Hartree (14.19) and Hartree-Fock (14.26) equations are solved numerically by iteration. The interpretation of the ‘energy eigenvalues’ $\lambda_{\alpha_i} = \mathcal{E}_{\alpha_i}$ is non-trivial. Multiplication of (14.26) by $\varphi_{\alpha_i}^*(\vec{r})$ and integration over \vec{r} gives us the eigenvalue

$$\begin{aligned} \mathcal{E}_{\alpha_i} = \langle \varphi_{\alpha_i} | H_0 | \varphi_{\alpha_i} \rangle + \sum_{j \neq i} \int d^3r \int d^3r' \\ \times \varphi_{\alpha_i}^*(\vec{r}) \varphi_{\alpha_j}^*(\vec{r}') \frac{e^2}{|\vec{r}-\vec{r}'|} \left[\varphi_{\alpha_j}(\vec{r}') \varphi_{\alpha_i}(\vec{r}) - \delta_{s_i s_j} \varphi_{\alpha_j}(\vec{r}) \varphi_{\alpha_i}(\vec{r}') \right]. \end{aligned} \quad (14.27)$$

This expression describes the energy of the electron in the orbital φ_{α_i} and involves the one-particle energy $\langle \varphi_{\alpha_i} | H_0 | \varphi_{\alpha_i} \rangle$ and the interaction- and exchange energies with all other electrons. Thus $-\mathcal{E}_{\alpha_i}$ is just the ionization energy in an experiment where the electron is removed from the system in a fast³ process, e.g., in a photoemission experiment. On the other hand, one understands that the ionization energy is the energy difference between a system with $N-1$ electrons plus an electron at infinity and an N particle system, with both systems relaxed, hence

$$\mathcal{E}_{\alpha_i}^{\text{ion}} = \langle \Psi_i | H | \Psi_i \rangle - \langle \Psi | H | \Psi \rangle, \quad (14.28)$$

where Ψ_i describes the relaxed $N-1$ particle wave function without the i th electron. This is the content of Koopman’s Theorem, telling that $\mathcal{E}_{\alpha_i}^{\text{ion}} \approx -\mathcal{E}_{\alpha_i}$, i.e., the orbital energy \mathcal{E}_{α_i} corresponds approximately (except for relaxation) to the ionization energy. In the same spirit, the difference $\mathcal{E}_{\alpha_i} - \mathcal{E}_{\alpha_k}$ gives the approximate excitation energy if a particle from the state φ_{α_k} is excited to the state φ_{α_i} .

³In a fast process, the electron is knocked out of the atom/system without letting the other particles relax, leaving a ‘hole’ in the system.

It should also be noted that the sum of the eigenvalues \mathcal{E}_{α_i} is not equal to the (Hartree-Fock) expectation value of the energy $\langle \Psi | H | \Psi \rangle$. Rather (avoiding double-counting of the interaction energy, see (14.23)),

$$\langle \Psi | H | \Psi \rangle = \sum_i \mathcal{E}_{\alpha_i} - \langle V_{\text{int}} \rangle \neq \sum_i \mathcal{E}_{\alpha_i}. \quad (14.29)$$

Also, note that the Hartree-Fock approximation does not consider correlation effects⁴. Neither (14.19) nor (14.26) are exact: in the Hartree-Fock approximation, the total energy E of the system includes the terms

$$E = \underbrace{\underbrace{E_0}_{\text{single-particle-}} + E_C + E_{\text{exchange.}}}_{\text{Hartree-Fock energy}}. \quad (14.30)$$

In addition, there is a correlation energy $E_{\text{correlation}}$ which arises in the context of many-body physics/quantum field theory via a systematic expansion in the Coulomb interaction. For atoms, the Hartree-Fock approximation provides results that are correct up to typical (Coulomb) correlation energies of size $\sim 0.5 \text{ eV}$ per electron.

14.2 Thomas-Fermi Screening

We consider electrons in the (smooth, external) potential $U(\vec{r})$. Such a slowly varying (in space) potential can be treated as locally constant, see Fig. 14.1.

The electrons then locally form a Fermi gas of density⁵

$$n(\vec{r}) = \frac{k_F^3(\vec{r})}{3\pi^2}. \quad (14.33)$$

⁴Correlation energy: An electron at position \vec{r} reduces the likelihood of encountering another electron there. In particular, an electron with spin $|\uparrow\rangle$ creates an exchange and correlation hole for other spin $|\uparrow\rangle$ electrons. The HF theory takes into account the exchange part of this hole; every portion of the exchange hole that goes beyond the (mean-field) HF theory is labeled as a correlation effect.

⁵The relation between the Fermi wave vector k_F and the density n follows from (with the box quantization $\vec{k} = (2\pi/L) \vec{n}$)

$$N = \sum_{k < k_F} 2 = \left(\frac{L}{2\pi}\right)^3 \int d^3k \cdot 2 = \frac{2V}{8\pi^3} \underbrace{\int_{4\pi} d\Omega}_{k_F^3/3} \int_0^{k_F} k^2 dk = \frac{V k_F^3}{3\pi^2}. \quad (14.31)$$

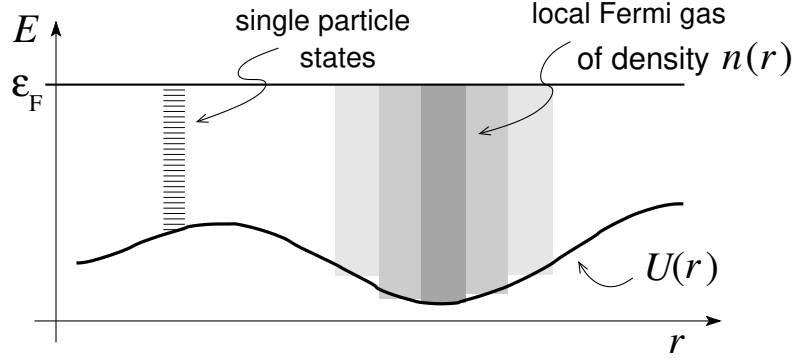


Fig. 14.1: Thomas-Fermi screening: The inhomogeneous electron system is approximated by a Fermi gas with local density $n(\vec{r}) = k_F^3(\vec{r})/3\pi^2$. The position-dependent Fermi vector $k_F(\vec{r})$ accounts for the presence of the external potential $U(\vec{r})$ as well as the screening effects by other electrons, see (14.35).

A test charge placed at the position \vec{r} experiences a local potential composed of the ‘bare’ potential $U(\vec{r})$ and the potential generated by the presence of neighboring charges constituting the inhomogeneous electron gas,

$$V(\vec{r}) = U(\vec{r}) + \int d^3r' \frac{e^2 n(\vec{r}')}{|\vec{r} - \vec{r}'|}. \quad (14.34)$$

The electrons in the Fermi gas distribute themselves such that the addition of the test electron costs the same energy ε_F at every location \vec{r} , hence,

$$\varepsilon_F = \frac{\hbar^2 k_F^2(\vec{r})}{2m} + V(\vec{r}). \quad (14.35)$$

The first term describes the kinetic energy of the test electron while the second term accounts for its potential energy $V(\vec{r})$. The combination of (14.33) and (14.35) provides us with the relationship between the local density $n(\vec{r})$ and the effective potential $V(\vec{r})$,

$$n(\vec{r}) = \frac{1}{3\pi^2} \left[\frac{2m}{\hbar^2} (\varepsilon_F - V[n(\vec{r}), \vec{r}]) \right]^{3/2}. \quad (14.36)$$

In 2 dimensions, the same analysis provides the relation (check out the 1D case yourself)

$$N = \frac{2V}{4\pi^2} \underbrace{\int_0^{k_F} k dk}_{k_F^2/2} = \frac{V k_F^2}{2\pi}. \quad (14.32)$$

Next, we transform the integral equation (14.36) for the density $n(\vec{r})$ into a differential equation for the effective potential $V(\vec{r})$: we apply the Laplacian Δ to (14.34) and make use of its Green's function, $\Delta |\vec{r} - \vec{r}'|^{-1} = -4\pi\delta(\vec{r} - \vec{r}')$, to find

$$\Delta V(\vec{r}) = \Delta U(\vec{r}) - 4\pi e^2 n(\vec{r}) \equiv -4\pi e^2 [n_{\text{ext}}(\vec{r}) + n(\vec{r})], \quad (14.37)$$

where $n_{\text{ext}}(\vec{r})$ denotes the charge density generated by the external potential $U(\vec{r})$ and $n(\vec{r})$ is the density of the screening charge. Equation (14.37) is the Poisson equation for the total potential V generated by the combination of external (n_{ext}) and screening (n) charge densities.

A side remark on the dielectric constant $\varepsilon(\vec{q})$. We consider a charge-neutral system, e.g., electrons on the positive background of the ions in a metal, and analyze the response of the system under a small perturbation $V(\vec{r})$ (the so-called linear response). With a charge density en_0 for the background and a charge density $-en[V(r)]$ for the electrons, we find for a small perturbation of the potential V an uncompensated charge density $-e\delta n = e[n_0 - n[V(r)]] \approx e[n_0 - (n(\varepsilon_F) - (\partial_{\varepsilon_F} n) V)]$, see (14.36). In a charge-neutral system, the uniform background charge density en_0 and $-en(\varepsilon_F)$ compensate each other, and we obtain a (small) density fluctuation $\delta n = (\partial_{\varepsilon_F} n)(-V)$ due to the perturbation V .

In electrostatics, the above situation is described with the help of the dielectric constant $\varepsilon = D/E$ (we consider an isotropic medium, with D the bare field (generated by external charges) and E the full field including screening charges. While ε is a constant in the case of a dielectric, in a metal with its mobile charge carriers, the dielectric response depends on the wave number \vec{q} of the perturbation $V(\vec{q})$ (the problem is conveniently analyzed in Fourier space as $\Delta V(\vec{r}) \leftrightarrow -q^2 V(\vec{q})$). Expressing the dielectric constant through the bare and full potentials U and V (instead of the fields D and E),

$$\varepsilon(q) \equiv \frac{U(q)}{V(q)} \quad (14.38)$$

we find that (using $q^2(V - U)$ from (14.37))

$$\varepsilon(\vec{q}) = \frac{q^2(U - V + V)}{q^2 V} = 1 - \frac{(V - U)q^2}{Vq^2} = 1 - \frac{4\pi e^2}{q^2} \frac{\delta n}{V}. \quad (14.39)$$

With $\delta n \approx (\partial_{\varepsilon_F} n)(-V)$, we arrive at the Thomas-Fermi result for the dielectric constant of a metal,

$$\varepsilon(\vec{q}) = 1 + \frac{4\pi e^2}{q^2} \frac{\partial n}{\partial \varepsilon_F} \equiv 1 + \frac{q_{\text{TF}}^2}{q^2}. \quad (14.40)$$

The Thomas-Fermi wavenumber q_{TF} is determined by the electron density n via k_{F} ,

$$q_{\text{TF}}^2 = \frac{1}{\lambda_{\text{TF}}^2} = 4\pi e^2 (\partial_{\varepsilon_{\text{F}}} n) = \frac{4}{\pi} \frac{k_{\text{F}}}{a_{\text{B}}}. \quad (14.41)$$

The combination of (14.37) and (14.40) provides us with the equation for the potential V of the form

$$(q^2 + \lambda_{\text{TF}}^{-2})V(\vec{q}) = 4\pi e^2 n_{\text{ext}}(\vec{q}), \quad (14.42)$$

and the inverse Fourier transformation to real space results in the screened Coulomb law (instead of the usual Poisson equation in a vacuum)

$$(\Delta - \lambda_{\text{TF}}^{-2})V(\vec{r}) = -4\pi e^2 n_{\text{ext}}(\vec{r}). \quad (14.43)$$

The Green's function for $\Delta - \lambda_{\text{TF}}^{-2}$ is of the Yukawa type, $G \propto e^{-(r/\lambda_{\text{TF}})}/r$. In a typical metal with $n \sim 10^{22} \text{ cm}^{-3}$, $k_{\text{F}} \approx 1 \text{ \AA}^{-1}$ and thus $\lambda_{\text{TF}} \sim 1 \text{ \AA}$; a charge in a metal is screened by other electrons within $\sim 1 \text{ \AA} = 10^{-10} \text{ m}$. Check out the analogous situation in a semiconductor where the electron density is much smaller, $n \sim 10^{18} \text{ cm}^{-3}$. \square

14.3 Thomas-Fermi Atom

We apply the idea of Thomas-Fermi screening to the discussion of atoms; in a first step, we determine the screened potential $V(\vec{r})$ seen by an individual electron in the atom with the help of the Poisson equation (14.37), with external charge $n_{\text{ext}}(\vec{r}) = Z\delta(\vec{r})$ generated by the nucleus and the screening charge $n(\vec{r})$ as given by (14.36). In a second step, we then use this potential in order to calculate the energy levels of individual electrons in the Thomas-Fermi atom by solving the Schrödinger equation for the Hamiltonian $H = p^2/2m + V(r)$ with $V(r)$ including the potentials of the nucleus and of all the other electrons.

We approximate $V(\vec{r})$ as spherically symmetric and have to solve the following equation away from the origin, i.e., $r > 0$,

$$r^{-2} \partial_r [r^2 \partial_r V(r)] = -\frac{4e^2}{3\pi} \left[\frac{2m}{\hbar^2} (\varepsilon_{\text{F}} - V(r)) \right]^{3/2}. \quad (14.44)$$

For $r \rightarrow 0$, the boundary condition for this equation reads $V(r) \approx -Ze^2/r$ as screening vanishes for $r \rightarrow 0$ and the electrons are exposed to the bare

potential of the nucleus. The atomic radius R is determined by the condition $n(R) = 0$. With the density determined via (14.36), the second boundary condition $V(R) = \varepsilon_F$ follows immediately. For a neutral atom, we have $V(r > R) = \text{const.} = 0$, while for a charged ion, $V(r > R) = -(Z - L)e^2/r$ with L the number of electrons, see Fig. 14.2.

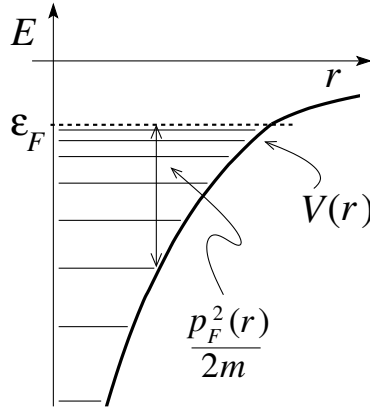


Fig. 14.2: Potential with bound states in the Thomas-Fermi atom. The states are filled up to the Fermi level ε_F ; their charge contributes, via screening, to the effective potential $V(r)$. For a neutral atom, $\varepsilon_F = 0$ and the electron shell extends to infinity. For a (positively charged) ion the electron cloud extends up to $R < \infty$ and $V(r) \propto 1/r$ for $r > R$.

In the following, we analyze the neutral case with $L = Z$ and define the dimensionless variables x and Φ for the length and energy (potential),

$$r = \left(\frac{3\pi}{4\sqrt{Z}} \right)^{2/3} \frac{a_B}{2} x = (0.8853/Z^{1/3}) a_B x, \quad V = -\frac{Ze^2}{r} \Phi(x). \quad (14.45)$$

Then the universal TF atom is characterized by a dimensionless ‘potential’ $\Phi(x)$ (independent of Z) that solves the differential equation

$$\sqrt{x} \partial_x^2 \Phi(x) = \Phi^{3/2}(x), \quad \Phi(0) = 1. \quad (14.46)$$

The dimensionless ‘potential’ satisfies the asymptotics⁶

$$\Phi(x) = \begin{cases} 1 - 1.59x, & x \rightarrow 0, \\ 144/x^3, & x \rightarrow \infty. \end{cases} \quad (14.47)$$

The full function $\Phi(x)$ is obtained via a numerical integration of (14.46). Note that the potential of the universal Thomas-Fermi atom extends to infinity since $V(r)$ decays like r^{-4} ($\Phi(x) \propto x^{-3}$); in reality we expect a finite expansion of the electron cloud and the potential. We find the following scaling laws for the Thomas-Fermi atom:

⁶Note that $V(r \rightarrow \infty) \propto 1/r^4$

- a radius $\sim a_B Z^{-1/3}$, heavy atoms are smaller,
- a potential $V \sim Z\Phi/r \sim Z^{4/3}\Phi(x)/x \sim Z^{4/3}$ for fixed x ,
- an e^- -density $n \sim V/r^2 \sim Z\Phi/r^3 \sim Z^2\Phi(x)/x^3 \sim Z^2$ at fixed x ,
- a kinetic energy $p^2/2m \sim n^{2/3} \sim Z^{4/3} \sim V$.

Once we have found the shielded potential $V(r)$, we can determine the individual electronic states in this potential by solving the associated Schrödinger equation $H\Psi = E\Psi$ with $H = p^2/2m + V(r)$. Note that $V(r)$ is no longer $\propto 1/r$ and thus the dynamic symmetry of the hydrogen atom is broken. However, the rotational symmetry is still preserved and we can find states with a fixed angular momentum $\hbar l$. The electrons within the TF atom arrange themselves in shells: The Thomas-Fermi atom binds a state with angular momentum l if the effective potential (including the centrifugal potential $\hbar^2 l(l+1)/2mr^2$)

$$V_{\text{eff}}(r) = V(r) + \frac{\hbar^2 l(l+1)}{2mr^2} \quad (14.48)$$

$$= \frac{Z^{2/3}e^2}{0.8853 a_B} \left[-Z^{2/3} \frac{\Phi(x)}{x} + \frac{l(l+1)}{1.77 x^2} \right]$$

has a local minimum with $V_{\text{eff}} < 0$, see Fig. 14.3.

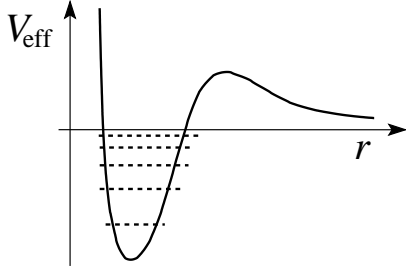


Fig. 14.3: Effective potential of the Thomas-Fermi atom including the centrifugal barrier. The TF atom binds a state with angular momentum l if V_{eff} has a minimum with negative energy.

With the TF potential $\Phi(x)$ as given by (14.46), one finds that the TF atom binds shells within the l and Z ranges as listed in table 14.1:

Note that with $V(r) \neq Ze^2/r$ the Hartree, Hartree-Fock, and Thomas-Fermi approximations give single-particle energies \mathcal{E}_{α_i} that depend on n and l . The degeneracy of the energy eigenvalue $\mathcal{E}_{\alpha_i} = \mathcal{E}_{nl}$ as determined by $\text{SO}(3)$ symmetry (orbit and spin) is given by $(2l+1)(2s+1) = 4l+2$. Each

	$l = 0$	$l \leq 1$	$l \leq 2$	$l \leq 3$
TF-Atom	$Z \leq 4$ only s	$5 \leq Z \leq 19$ $s \& p$	$20 \leq Z \leq 53$ $s, p \& d$	$54 \leq Z$ $s, p, d \& f$
Experimental		p -electrons for $Z \geq 5$	d - e^- for $Z \geq 21$	f - e^- i for $Z \geq 58$

Table 14.1: Shells in the Thomas-Fermi atom. Larger angular momenta l raise the minimum in $V_{\text{eff}}(r)$ and require a larger nuclear charge Z in order to bind another stable shell. Boron has $Z = 5$, Scandium has $Z = 21$, and Cerium has $Z = 58$.

eigenvalue \mathcal{E}_{nl} defines a shell with the following electron filling,

$$\begin{aligned}
 s\text{-shells:} & \quad 2e^-, \\
 p\text{-shells:} & \quad 6e^-, \\
 d\text{-shells:} & \quad 10e^-, \\
 f\text{-shells:} & \quad 14e^-.
 \end{aligned} \tag{14.49}$$

The determination of the shell sequence (increase of l or of n) requires a numerical analysis that provides the result

$$\begin{aligned}
 1s < 2s < 2p < 3s < 3p < 4s < 3d < 4p < 5s < 4d < 5p \\
 & < 6s < 4f < 5d < 6p < 7s < 5f
 \end{aligned} \tag{14.50}$$

This sequence can be memorized with the help of the scheme in the sketch 14.4.

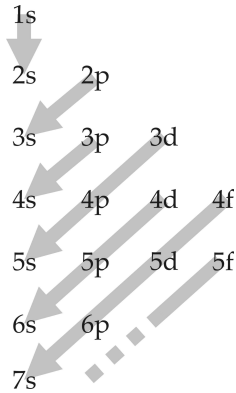


Fig. 14.4: Sequence of shells nl for the energy of the last added electron in the Thomas-Fermi atom. n = main quantum number, l = angular momentum quantum number. Note that the low energy/inner electrons have a different sequence (namely the trivial one): The sequence of the so-called X-ray shells is $1s$ (the 1K shell), $2s, 2p$ (2L shell), $3s, 3p, 3d$ (3M shell), $4s, 4p, 4d, 4f$ (4N shell), etc.

We go through the period table with some comments:

- H, He: the $1s$ orbital/shell is filled
- Li, Be: the $2s$ orbital is filled. B–Ne: the $2p$ -orbital is filled. $2s$ and $2p$ have different energies, the orbitals with small l are closer to the core and therefore their energies are smaller, $\mathcal{E}_{2s} < \mathcal{E}_{2p}$.
- Na–Ar: as before, first the $3s$ -, then the $3p$ -shell is filled.
- K–Kr: The energies $4s \approx 3d$, thus the sequence for filling the s - and d -shells is

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

The p -shell is filled with Ga–Kr. The $3d$ -shells are compact, and therefore chemically less important; but they provide the magnetic moments which produce the ferro- or antiferromagnetism of the elements Cr–Ni.

- Rb–Xe: essentially a repetition of the sequence K–Kr. Again, the s -orbitals have about the same energy as the d -orbitals of the previous shell, which results in the irregular filling

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

The p -shells are filled with In–Xe.

- Cs–Rn: Cs and Ba fill the $6s$ -orbital. La fills the first $5d^1$ orbital, followed by the lanthanides (rare earths) La–Lu filling up the $4f$ shells which are energetically close to $5d$. Again there are shifts in the occupations,

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

The compact f -orbitals are passive in chemical bonds, but result in large magnetic moments, especially in the case of Ce, Pr, Nd, Sm–Yb. Following the filling of the f -shells, the $5d$ and $6p$ shells are filled; Pt and Au show another reorganization with $5s^1$.

- Fr–Rg: Fr and Ra fill the $7s$ -orbital, Ac has the configuration $6d^1$, then the actinides fill the f -orbitals. The energies of $6d$ and $5f$ are close together and correspondingly there are once again reorganizations with $6d^{0,1}$ to Lr and $6d^{2,3}$ for Rf and Db. There are no stable elements with filled $6d$, $7p$ orbitals; Lead ($Z = 82$) is the last stable (non-radioactive) element, Bismuth ($Z = 83$) lives extremely long. Te (43) and Pm (61) are unstable, so there are 80 stable natural elements.

The outer electrons are relevant for the chemical properties of the elements. Accordingly, elements in a group or column of the periodic table have similar chemical properties. For example,

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

14.4 Ground state configuration of atoms

We lay out the problem with a specific example: Consider carbon C with the electron configuration $1s^2 2s^2 2p^2$. In the p -orbital there are 6 one-electron states available, 3 orbitals p_x , p_y , p_z and 2 spin states \uparrow and \downarrow ,

$$\Psi_p = \varphi_p \chi_s = \left\{ \begin{array}{c} \varphi_{p_x} \\ \varphi_{p_y} \\ \varphi_{p_z} \end{array} \right\} \otimes \left\{ \begin{array}{c} \chi_{\uparrow} \\ \chi_{\downarrow} \end{array} \right\}. \quad (14.51)$$

The first electron is chosen from 6 states, the second from the remaining 5 states; this provides us with $5 \cdot 6/2 = 15$ possible states with equal single-particle energies⁸ associated with the Hamiltonian

$$H_0 = \sum_i \frac{p_i^2}{2m} + V(r_i). \quad (14.52)$$

The more complete Hamiltonian

$$H = H_0 + H_C + H_{\text{SO}}, \quad (14.53)$$

⁸The potential $V(r)$ takes into account the potential of the nucleus and the mean interaction potential with other electrons (screening).

however, accounts for the residual Coulomb interaction beyond mean-field,

$$H_C = \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\vec{r}_i - \vec{r}_j|} - \sum_i \left(\frac{Ze^2}{r_i} + V(r_i) \right), \quad (14.54)$$

as well as the spin-orbit interaction

$$H_{\text{SO}} = \frac{1}{2m^2c^2} \frac{1}{r} \frac{dV}{dr} \sum_i \vec{l}_i \cdot \vec{s}_i, \quad (14.55)$$

with \vec{l}_i and \vec{s}_i the orbital angular momentum and spin of the i -th electron; we note that further correction terms exist that may be taken into account, such as the interaction between the spins of the electrons and the nucleus, as well as relativistic and field-theoretic corrections. Taking these energies into account, the 15 configurational states are split and it is our next goal to find the atomic groundstate including the terms in (14.53) and its remaining degeneracy.

To solve this task, we treat the terms H_C and H_{SO} with the help of perturbation theory. The Hamiltonian $H_0 + H_C$ is still SO(3)-symmetrical and spin-independent. Thus $\vec{L} = \sum_i \vec{l}_i$ and $\vec{S} = \sum_i \vec{s}_i$ define good quantum numbers.⁹ The spin-orbit coupling destroys this symmetry and \vec{L} , \vec{S} are no longer individually conserved. Therefore one has to consider the total angular momentum \vec{J} , with

$$\vec{J} = \sum_i \vec{j}_i, \quad \vec{j}_i = \vec{l}_i + \vec{s}_i; \quad (14.56)$$

this total angular momentum is conserved.

14.4.1 Structure of perturbation theory

A rough estimate of H_C and H_{SO} provide the following estimates for the energies (per particle)¹⁰

$$h_C \sim 0.1 \frac{e^2}{a_B} \sqrt{Z} \sim \sqrt{Z} \text{ eV}, \quad (14.57)$$

⁹Note that the total wave function Ψ must be antisymmetric, that is, the mixture of spin wave functions causes the individual spins \vec{s}_i to be no longer good quantum numbers. For example, for $n = 2$ the singlet spin-function $\chi_S = [|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle] / \sqrt{2}$ and $S_z\chi_S = 0$ but $s_{1z}\chi_S \neq s\chi_S$.

¹⁰The numerical pre-factor in (14.57) can be estimated with the help of the results for Helium, $Z = 2$; the correction scales like \sqrt{Z} because it is a fluctuation term.

$$\begin{aligned}
h_{\text{SO}} &= \frac{1}{2m^2c^2} \underbrace{\frac{1}{r} \frac{dV}{dr}}_{Ze^2/a^3} \underbrace{\vec{l} \cdot \vec{s}}_{\hbar^2} \quad [\text{with } a = a_{\text{B}}/Z^{1/3}] \\
&\sim \underbrace{\frac{\hbar^2}{2ma_{\text{B}}^2}}_{\sim \text{Ry}} \frac{1}{mc^2} \underbrace{\frac{e^2}{a_{\text{B}}}}_{\sim \text{Ry}} Z^2 \sim Z^2 10^{-3} \text{ eV}.
\end{aligned} \tag{14.58}$$

This allows us to estimate the relative magnitude of the two corrections,

$$\begin{aligned}
h_{\text{C}} &> h_{\text{SO}}, & \text{for } Z < 10^2, \\
h_{\text{C}} &< h_{\text{SO}}, & \text{for } Z > 10^2.
\end{aligned} \tag{14.59}$$

A more precise estimate provides the intersection of the two energies at $Z \sim 80$ which corresponds to the atomic number of lead (Pb). We analyze the case of light atoms with $H_{\text{C}} > H_{\text{SO}}$ in some detail and then comment on the heavy ones.

14.4.2 Atoms lighter than Pb

We first deal with the problem $H_0 + H_{\text{C}}$ and then consider the spin-orbit coupling H_{SO} as a perturbation. The strategy is as follows: without H_{SO} , L and S are still good quantum numbers and we find the values that are allowed by the configuration and the combinations that are compatible with the Pauli exclusion principle. The highly degenerate configuration level is then split into LS -multiplets through the interaction energy H_{C} that affects the different L, S combinations in a different manner; these are still $(2L+1)(2S+1)$ -fold degenerate. The further splitting of the LS -multiplets by H_{SO} will then be discussed, which leads to the fine structure with levels that are $(2J+1)$ -fold degenerate. Finally, the remaining $(2J+1)$ -fold degeneracy in the fine structure due to the $SO(3)$ symmetry can be lifted by a magnetic field.

The symmetry in the problem is $SO(3)$. We combine the degenerate configurational states into states with sharp angular momentum and spin L^2 , L_z , S^2 , S_z and find the state with the lowest energy. This state is still $(2L+1)(2S+1)$ -fold degenerate. In our specific example of carbon C with the configuration $1s^2 2s^2 2p^2$ the $2p$ orbitals can be combined into states with (we can ignore the full shells $1s^2 2s^2$)

$$\begin{aligned}
L_{\text{max}} &= 2 & \text{to} & & L_{\text{min}} &= 0 & (H_1 \otimes H_1 &= H_0 \oplus H_1 \oplus H_2), \\
S_{\text{max}} &= 1 & \text{to} & & S_{\text{min}} &= 0 & (H_{\frac{1}{2}} \otimes H_{\frac{1}{2}} &= H_0 \oplus H_1).
\end{aligned} \tag{14.60}$$

The Pauli principle then allows for the combinations¹¹

$$\begin{aligned}
 L = 2 \quad \text{symmetric} &\rightarrow S = 0 \quad \text{antisymm.} \rightarrow J = 2, \\
 L = 0 \quad \text{symmetric} &\rightarrow S = 0 \quad \text{antisymm.} \rightarrow J = 0, \\
 L = 1 \quad \text{antisymm.} &\rightarrow S = 1 \quad \text{symmetric} \rightarrow J = 0, 1, 2,
 \end{aligned} \tag{14.61}$$

the so-called LS -multiplets. For their characterization, we make use of the term symbol defined as

$$^{2S+1}L_J, \tag{14.62}$$

where 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.

The configurational states then split under H_C , with different LS -multiplets assuming different energies (in the 0.1 eV range). In figure 14.6 this splitting is illustrated on the example of Carbon (C). The sequence of this splitting

configuration	LS –multiplets	degeneracy ($2L+1$) ($2S+1$)
C		
15	$L=0 \quad S=0 \quad {}^1S_0$	1
$1s^2 2s^2 2p^2$	$L=2 \quad S=0 \quad {}^1D_2$	5
	$L=1 \quad S=1 \quad {}^3P_0 {}^3P_1 {}^3P_2$	$1+3+5=9$
		15

Fig. 14.6: Configuration of the carbon atom C and its splitting into LS -multiplets and their degeneracy. The splitting of the highly degenerate configurational level involves the part of the Coulomb energy that is not already taken into account in the one-particle energies, see (14.54).

¹¹For two electrons with the same angular momentum $l_1 = l_2 = l$, it can be shown that their total angular momentum L being even (odd) implies a symmetric (anti-symmetric) orbital wave function (e.g., calculate the relation between the CGC $\langle llm_1m_2|lll_{\text{tot}}m_1+m_2\rangle$ and $\langle llm_2m_1|lll_{\text{tot}}m_1+m_2\rangle$) (this no longer applies if we have more than two electrons or if $l_1 \neq l_2$). For $l_1 = l_2 = 1$ and $L = 2$, the symmetry of the state follows trivially since both electrons have the same orbital states, the antisymmetry of the $L = 1$ state can be shown via angular momentum addition (CGC). Note also the decomposition $(H_0^s \oplus H_1^a \oplus H_2^s)_L \otimes (H_0^a \oplus H_1^s)_S = (H_0^a) \oplus (H_1^s) \oplus (H_1^s) \oplus (H_0^a \oplus H_1^a \oplus H_2^a) \oplus (H_2^s) \oplus (H_1^s \oplus H_2^s \oplus H_3^s)$. Of these sectors, only those that are anti-symmetric are allowed, $H_0^a \oplus H_0^a \oplus H_1^a \oplus H_2^a \oplus H_2^a$, has dimension 15.

is not trivial—the L , S , J quantum numbers for the atomic ground state are given by Hund’s rules discussed below, see section 14.4.3.

In the next step, we consider the spin-orbit coupling H_{so} . Within an LS -multiplet, one can use the Wigner-Eckart theorem to show (see 14.5) that $H_{\text{so}} \propto \sum_i \vec{l}_i \cdot \vec{s}_i \propto \vec{L} \cdot \vec{S}$. The spin-orbit coupling H_{so} therefore couples the momenta \vec{L} and \vec{S} and only the total angular momentum \vec{J} remains a good quantum number. In the $(2L+1)(2S+1)$ -dimensional LS -multiplet, we can go over to the \vec{J} basis (the ‘quantum number’ K symbolizes the quantum numbers of the electron configuration, which is not described in further detail)

$$|K L S M_L M_S\rangle \rightarrow |K L S J M\rangle \quad (14.63)$$

and represent the value of $\vec{L} \cdot \vec{S}$ through the quantum numbers L , S , and J ,

$$\begin{aligned} \vec{L} \cdot \vec{S} &= \frac{1}{2}(J^2 - L^2 - S^2) \\ &= \frac{1}{2}[J(J+1) - \underbrace{L(L+1) - S(S+1)}_{\text{constant in the } LS\text{-multiplet}}]. \end{aligned} \quad (14.64)$$

The term H_{so} will separate the LS -multiplets into $2J+1$ levels with energies $\sim \text{Ry}^2/(mc^2)Z^2 \cdot J$ in the 1–10 meV range. The term diagram for carbon with the LS -splitting and the fine structure of the spin-orbit coupling is sketched in Figure 14.7.

configuration	LS –multiplets	fine structure	degeneracy ($2J+1$)
C 15 $1s^2 2s^2 2p^2$	$L=0 \quad S=0$	1S_0	1
	$L=2 \quad S=0$	1D_2	5
	$L=1 \quad S=1$	3P_2	5
		3P_1	3
		3P_0	1
			<hr/> 15

Fig. 14.7: Configuration of the carbon atom C with its splitting into LS -multiplets, followed by a further level splitting constituting the fine structure (splitting in the range 10 meV). The latter is induced by the (relativistic) spin-orbit coupling which mixes the orbital and spin angular momenta. The total angular momentum J remains a good quantum number.

The remaining $(2J+1)$ degeneracy can be lifted by a (weak) magnetic field,

see the discussion in section 14.7. The resulting Zeeman splitting for the carbon atom is sketched in figure 14.8.

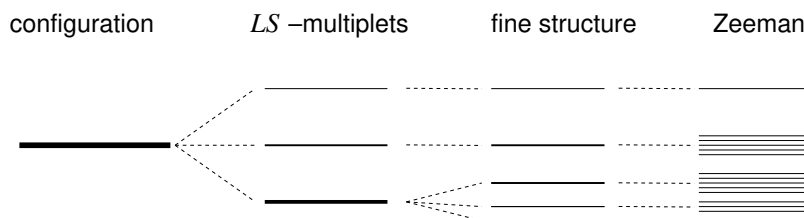


Fig. 14.8: Configuration of the carbon atom C with splitting into LS -multiplets, their fine structure and the Zeeman splitting in the magnetic field. The associated Hamiltonian with the various terms is given by $H_0 + H_C + H_{\text{SO}} + \sum_i \mu_B \vec{H} \cdot (\vec{l}_i + 2\vec{s}_i)/\hbar$. In addition to the splitting in the external magnetic field (Zeeman), the nuclear spin also creates an effective magnetic field for the electrons, which leads to the hyperfine structure (not shown).

The exact calculation of the term diagram and the ground state $^{2s+1}L_J$ involves a complicated calculation as is outlined in the scheme 14.9 for the case of light atoms. On the other hand, the task of finding the ground state configurations (i.e., the corresponding term symbols $^{2s+1}L_J$) for the light atoms can be accomplished easily by using the empirical Hund's rules.

14.4.3 Hund's rules

1. Full shells do not contribute to the orbit/spin angular momenta L and S .
2. The LS -multiplet with the largest S has the smallest energy.
3. Among different L with the same S , the largest L has the smallest energy.
4. If the shell is less than half full or half full, then

$$J = |L - S| \text{ is minimal.}$$

If the shell is more than half full, then

$$J = L + S \text{ is maximal.}$$

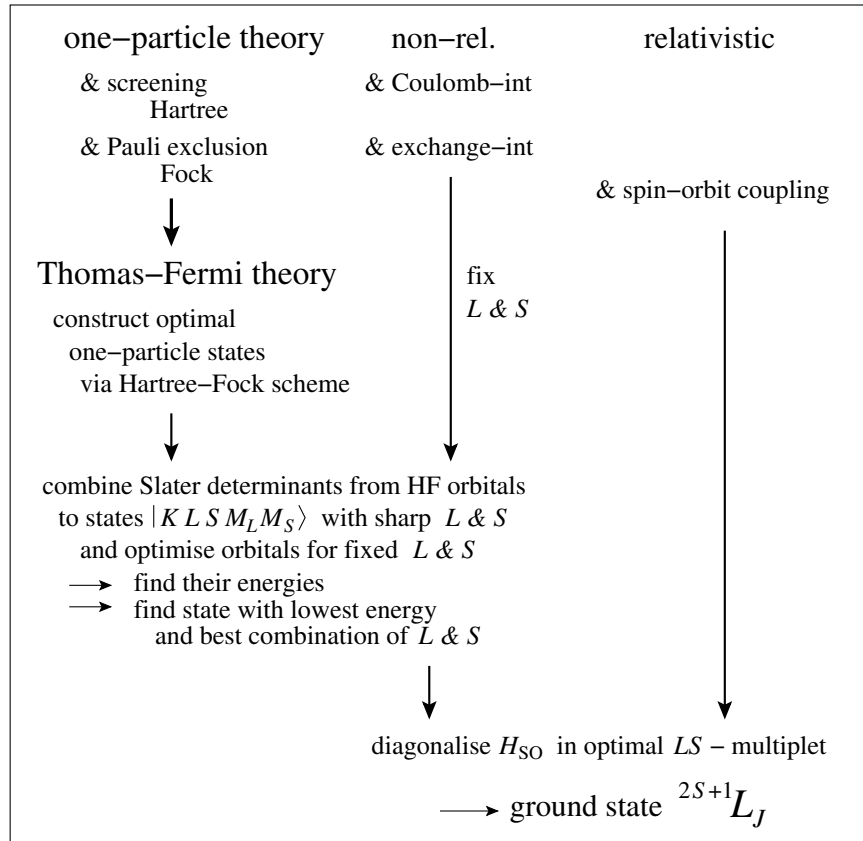


Fig. 14.9: Schematic representation for finding the ground state configuration (term diagram) $^{2S+1}L_J$ of an atom starting from electron states in the bare potential $U(\vec{r}) = -Ze^2/r$ of the atomic nucleus with charge Z .

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

1. In full shells the angular momenta cancel out and the ground state configuration corresponds to the term symbol 1S_0 .
2. If the total spin S is maximal, then the spin wave function χ_{spin} is symmetric, the orbital function φ_{orbit} must be anti-symmetric and the electrons avoid each other due to Pauli exclusion. This configuration results in the smallest Coulomb energy.
3. If the orbital angular momentum L is maximal, then the radii of the

electron orbitals are large and the electron density is small, which in turn minimizes the Coulomb energy.

4. The spin-orbit coupling $\xi \vec{L} \cdot \vec{S}$ encourages (anti-) parallel orbital- and spin-angular momenta if the coefficient ξ is (positive) negative. For (less-) more-than-half-filled shells ξ is (positive) negative, see the analysis in section 14.6.

Examples for the application of Hund's rules, term symbols

H: $^2S_{1/2}$, trivial.

He: filled $1s$ -shell, 1S_0 .

Li: Li has a filled shell (\rightarrow no contribution to L or S) plus an electron in the $2s$ -shell, and thus has a ground state configuration equivalent to H, $^2S_{1/2}$.

Be: filled $2s$ -shell 1S_0 like He.

B: The configuration is $1s^2 2s^2 2p$; of key relevance is the electron in the $2p$ -shell with $L = 1$ and $S = 1/2$ (6 states) with possible angular momenta $J = 3/2$ (4-fold degenerate) or $J = 1/2$ (2-fold degenerate). The fourth Hund rule selects the minimum total momentum $J = 1/2$ and we obtain the term symbol $^2P_{1/2}$.

C: The configuration $1s^2 2s^2 2p^2$ (15 states) admits the following LS -multiplets:

- $L_{\max} = 2 \Rightarrow \varphi_{\text{orbit}}$ symmetric with anti-symmetric singlet state $S = 0 \Rightarrow J = 2$; term symbol 1D_2 , 5-fold degenerate, 5 states.
- $L = 1 \Rightarrow \varphi_{\text{orbit}}$ antisymmetric and symmetric spin-triplet $S = 1 \Rightarrow J = 0, 1, 2$; term symbols $^3P_0, ^3P_1, ^3P_2$: 1, 3, 5-fold degenerate, 9 states.
- $L = 0 \Rightarrow \varphi_{\text{orbit}}$ symmetric and $S = 0 \Rightarrow J = 0$; term symbol 1S_0 : not degenerate, 1 state.

Adding up the degeneracies gives the 15 states. The second Hund rule selects the configurations $^3P_0, ^3P_1, ^3P_2$, all with $L = 1$ (no selection based on Hund's third rule). The p shell is less than half full and the fourth Hund rule determines the total angular momentum $J = 0$, so the correct term symbol for the ground state is 3P_0 .

N: The configuration $1s^2 2s^2 2p^3$ allows for $6 \cdot 5 \cdot 4 / (3 \cdot 2 \cdot 1) = 20$ states. The Pauli exclusion admits the following combinations of L and S :

- $L_{\max} = 3 \Rightarrow \varphi_{\text{orbit}}$ symmetric, thus χ_{spin} must be antisymmetric, but there is no antisymmetric spin function for 3 particles, so we must have $L < 3$. For 3 electrons the analysis of spin and race (see section 13.7) provides the following allowed combinations:
- $L = 2 \Rightarrow S = 1/2, {}^2D_{3/2}, {}^2D_{5/2}$, total of 10 states.
- $L = 1 \Rightarrow S = 1/2, {}^2P_{1/2}, {}^2P_{3/2}$, 6 states.
- $L = 0 \Rightarrow S = 3/2, {}^4S_{3/2}$, 4 states.

The analysis correctly gives the expected 20 states. The ground state is ${}^4S_{3/2}$ according to Hund's second rule.

O: $1s^2 2s^2 2p^4 \hat{=} 2p^{-2}$, i.e., oxygen is handled like carbon, 15 states with ${}^1S_0, {}^3P_0, {}^3P_1, {}^3P_2, {}^1D_2$ but with a more than half filled shell. Hund's rules select:

H2: ${}^3P_0, {}^3P_1, {}^3P_2$.

H3: all $L = 1$ orbitals are equivalent, no selection.

H4: 3P_2 .

F: like B, but ${}^2P_{3/2}$ because the shell is 5/6 full.

Ne: full shell, like He, 1S_0 .

Furthermore, there are usually (i.e., for the s and p electrons) the same configuration for elements of the same column, e.g., B, Al, In, Tl, have all a ground state with ${}^2P_{1/2}$, C, Si, Ge, Sn, Pb, all with 3P_0 , etc. However, there are many deviations for the d and f electrons.

In the following, we want to analyze the spin-orbit coupling and the Zeeman effect in more detail. We need the Wigner-Eckart theorem as a mathematical tool, which we first derive and then discuss along a few specific examples. Before that, however, a brief remark about the heavy atoms.

14.4.4 Atoms heavier than Pb

For heavy atoms, we first analyze $H_0 + H_{\text{so}}$ and then consider H_C as a perturbation. We combine the one-particle orbitals $|n, l, m, s\rangle = \varphi_{nlm} \chi_s = R_{nl} Y_{lm} \chi_s$ with fixed l and s in such a way as to have a fixed \vec{j} , $j = l \pm 1/2$,

and find the new orbitals $|n, l, j, m_j\rangle$. These orbitals take the spin-orbit interaction into account already at the one-particle level.

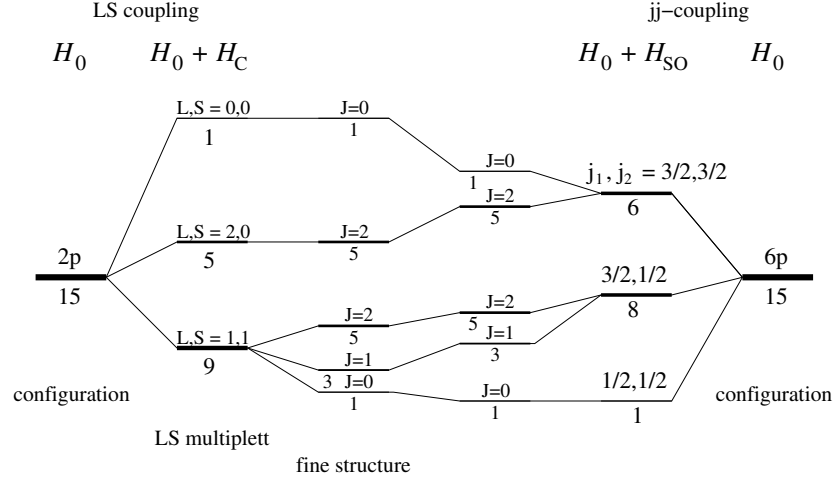


Fig. 14.10: Evolution of the term scheme for two p electrons, e.g., from Carbon C ($2p$ electrons) to lead Pb ($6p$ electrons). The $2p$ configuration is 15-fold degenerate and splits (from left to right) for light atoms into the LS -multiplets due to the Coulomb interaction. The latter split again into the fine structure due to a small spin-orbit coupling. For heavy atoms, we start again from the 15-fold degenerate configuration of the $6p$ electrons, which split according to the jj -coupling due to the spin-orbit interaction. By adding the small Coulomb interaction, the J levels split again.

As an example, let us again consider a two-electron atom with an outer p shell (e.g., lead with $n = 6$) then we find the two j -states $j = 1/2$ and $j = 3/2$. In the next step, we construct the configurations for the many-body problem $H_0 + H_{SO}$. Accordingly, we find Slater determinants at fixed J instead of fixed L & S . According to the Pauli exclusion principle, we are not allowed to have 2 electrons in the same j, j_z state. In our example of lead, the possible combinations are $(j_1, j_2) = (3/2, 3/2)$, $(3/2, 1/2)$ and $(1/2, 1/2)$. For the choice $(j_1, j_2) = (3/2, 3/2)$, we can construct six (that is, $(4 \times 4 - 4)/2$ different pairings) (m_1, m_2) states, $(m_1, m_2) = (3/2, 1/2)$, $(3/2, -1/2)$, $(3/2, -3/2)$, $(1/2, -1/2)$, $(1/2, -3/2)$, $(-1/2, -3/2)$. This gives $M = m_1 + m_2$ values in $\{-2, \dots, 2\}$ which correspond to the allowed values $J = 0, 2$. For $j_1 = 3/2, j_2 = 1/2$, we find 8 ($= 4 \times 2$) states $(m_1, m_2) = (3/2, 1/2)$, $(3/2, -1/2)$, $(1/2, 1/2)$, $(1/2, -1/2)$, $(-1/2, 1/2)$, $(-1/2, -1/2)$,

$(-3/2, 1/2)$, $(-3/2, -1/2)$ (note that all $|j, j_z\rangle$ states are different, so all 4×2 combinations count). This again gives $M = m_1 + m_2$ values in $\{-2, \dots, 2\}$ which this time correspond to the permitted values $J = 1, 2$. Finally $j_1 = 1/2 = j_2$ gives a state to $J = 0$ and thus we have found a total of 15 states. Compared to carbon, we have classified these 15 states differently.

Finally, we take into account the remaining interaction between the electrons by calculating the expectation value of the interaction energy H_C for states with fixed total angular momentum \vec{J} , resulting in a $(2J+1)$ -fold degenerate state for every J with sharp values of j_i^2, J^2, J_z .

The two schemes for describing light and heavy atoms are called

Russel-Saunders- or LS-coupling	jj-coupling
$Z \leq 80$	$Z \geq 80$

- | | |
|--|---|
| <ol style="list-style-type: none"> 1. With $\vec{L} = \sum_i \vec{l}_i, \vec{S} = \sum_i \vec{s}_i$ the LS-multiplets, split by H_C, are constructed. 2. With $\vec{J} = \vec{L} + \vec{S}$, the spin-orbit coupling H_{SO} is diagonalized and the fine structure is determined. | <ol style="list-style-type: none"> 1. With $\vec{j}_i = \vec{l}_i + \vec{s}_i$, the spin-orbit coupling H_{SO} is accounted for at the one-particle level. 2. For the different combinations of j_i with fixed J, the interaction H_C is calculated. |
|--|---|

The transition from LS -coupling to jj -coupling is sketched in 14.10 with np electrons, $n = 2, 6$ in our example.

14.5 Wigner-Eckart theorem

The Wigner-Eckart Theorem is a statement about the matrix elements of tensor operators; it facilitates their calculation by expressing their dependence on the magnetic quantum numbers through the Clebsch-Gordan coefficients.

Formulation: Let T_q^k be a tensor operator of order k , see chapter 4 of QM I for details. α and α' denote the quantum numbers that are not associated

with the rotational symmetry. Then¹²

$$\langle \alpha' j' m' | T_q^k | \alpha j m \rangle = \frac{\langle \alpha' j' || T^k || \alpha j \rangle}{\sqrt{2j'+1}} \underbrace{\langle k j q m | k j j' m' \rangle}_{\text{CGC}}. \quad (14.65)$$

The (reduced) matrix element $\langle \alpha' j' || T^k || \alpha j \rangle$ depends only on T^k , j , j' and the remaining quantum numbers α, α' ; it is independent of q, m, m' . The Clebsch-Gordon coefficients (CGC) then contribute the dependence on q, m, m' .

Proof*: Consider the Clebsch-Gordan combined angular momentum states $(|kq\rangle \otimes |j_1 m_1\rangle \rightarrow |k j_1 j m\rangle)$

$$|\tilde{\alpha} j m\rangle = \sum_{q, m_1} T_q^k |\alpha j_1 m_1\rangle \langle k j_1 q m_1 | k j_1 j m\rangle. \quad (14.66)$$

The state $|\tilde{\alpha} j m\rangle$ is an eigenstate of J^2 and J_z with eigenvalues $j(j+1)$ and m ; this follows from the fact that $|\tilde{\alpha} j m\rangle$ transforms under rotations $\mathbf{R}_{\vec{\omega}}$ according to the irreducible representation \mathcal{D}^j (the proof is following),

$$U_{\vec{\omega}} |\tilde{\alpha} j m\rangle = \sum_{m'_2} d_{m'_2 m}^j |\tilde{\alpha} j m'_2\rangle. \quad (14.67)$$

Then consider the scalar product of $|\tilde{\alpha} j m\rangle$ with $\langle \alpha' j' m' |$; the matrix element $\langle \alpha' j' m' | \tilde{\alpha} j m \rangle$ is only $\neq 0$, if $j' = j$ and $m' = m$. It is also independent of m because

$$\begin{aligned} \langle \alpha' j' m' | \tilde{\alpha} j m \rangle &= \int d\omega \langle \alpha' j' m' | U_{\vec{\omega}}^{-1} U_{\vec{\omega}} | \tilde{\alpha} j m \rangle \\ &\stackrel{(14.67)}{=} \sum_{\bar{m}, \bar{m}'} \underbrace{\int d\omega d_{\bar{m}' m'}^{j'*} d_{\bar{m} m}^j}_{\delta_{j' j} \delta_{\bar{m}' \bar{m}} \delta_{m' m} / (2j+1)} \langle \alpha' j' \bar{m}' | \tilde{\alpha} j \bar{m} \rangle \\ &= \delta_{j j'} \delta_{m m'} \underbrace{\frac{1}{2j+1} \sum_{\bar{m}} \langle \alpha' j \bar{m} | \tilde{\alpha} j \bar{m} \rangle}_{\text{average over } \bar{m}}. \end{aligned} \quad (14.68)$$

Multiplication with $\langle k j_1 j m | k j_1 q' m'_1 \rangle$ and subsequent summation over j, m exploiting the orthonormality relation of CGC, gives the factors $\delta_{q q'} \delta_{m_1 m'_1}$

¹²The CGC add the angular momenta $|k, q\rangle$ and $|j, m\rangle$ to $|j', m'\rangle$.

and we find the result

$$\begin{aligned}\langle \alpha' j' m' | T_q^k | \alpha j m \rangle &= \sum_{\bar{m}} \frac{\langle \alpha' j' \bar{m} | \tilde{\alpha} j' \bar{m} \rangle}{2j' + 1} \langle k j q m | k j j' m' \rangle \\ &= \frac{\langle \alpha' j' \| T^k \| \alpha j \rangle}{\sqrt{2j' + 1}} \langle k j q m | k j j' m' \rangle.\end{aligned}\quad \square$$

To prove (14.67), we rotate the state $|\tilde{\alpha} j m\rangle$ with $U_{\tilde{\omega}}$

$$\begin{aligned}U_{\tilde{\omega}} |\tilde{\alpha} j m\rangle &= \sum_{q, m_1} \underbrace{U_{\tilde{\omega}} T_q^k U_{\tilde{\omega}}^{-1}}_{\sum_{q'} d_{q'q}^k T_{q'}^k} \underbrace{U_{\tilde{\omega}} |\alpha j_1 m_1\rangle}_{\sum_{m'_1} d_{m'_1 m_1}^{j_1} |\alpha j_1 m'_1\rangle} \langle k j_1 q m_1 | k j_1 j m \rangle \\ &= \sum_{q', m'_1} T_{q'}^k |\alpha j_1 m'_1\rangle \sum_{q, m_1} \underbrace{d_{q'q}^k d_{m'_1 m_1}^{j_1}}_{\mathcal{D}^k \otimes \mathcal{D}^{j_1} = \oplus_{j_2} \mathcal{D}^{j_2}} \langle k j_1 q m_1 | k j_1 j m \rangle.\end{aligned}$$

where we have used the rotation properties of the tensor T^k and of states $|l, m\rangle$. The product $d_{q'q}^k d_{m'_1 m_1}^{j_1}$ can be expressed through the rotation matrices $d_{m'_2 m_2}^{j_2}$ (see later for the proof),

$$d_{q'q}^k d_{m'_1 m_1}^{j_1} = \sum_{j_2, m_2, m'_2} \langle k j_1 q' m'_1 | k j_1 j_2 m'_2 \rangle \langle k j_1 q m_1 | k j_1 j_2 m_2 \rangle d_{m'_2 m_2}^{j_2}. \quad (14.69)$$

This relation corresponds to the decomposition of $\mathcal{D}^k \otimes \mathcal{D}^{j_1} = \oplus_{j_2} \mathcal{D}^{j_2}$ in Clebsch-Gordan notation, see Eq. (14.72). In the next step, we sum over q and m_1 and obtain

$$\begin{aligned}&\sum_{q, m_1} \left[\sum_{j_2, m_2, m'_2} \langle k j_1 q' m'_1 | k j_1 j_2 m'_2 \rangle \langle k j_1 q m_1 | k j_1 j_2 m_2 \rangle d_{m'_2 m_2}^{j_2} \right] \\ &\quad \sum_{q, m_1} \text{underlined} \rightarrow \delta_{j j_2} \delta_{m m_2} \times \langle k j_1 q m_1 | k j_1 j m \rangle \\ &= \sum_{m'_2} \langle k j_1 q' m'_1 | k j_1 j m'_2 \rangle d_{m'_2 m}^j.\end{aligned}$$

Thus, we obtain the result (14.67) [compare to (14.66)]

$$\begin{aligned}U_{\tilde{\omega}} |\tilde{\alpha} j m\rangle &= \sum_{q', m'_1, m'_2} T_{q'}^k |\alpha j_1 m'_1\rangle \langle k j_1 q' m'_1 | k j_1 j m'_2 \rangle d_{m'_2 m}^j \\ &= \sum_{m'_2} d_{m'_2 m}^j |\tilde{\alpha} j m'_2\rangle.\end{aligned}\quad (14.70)$$

To decompose $\mathcal{D}^k \otimes \mathcal{D}^{j_1} = \oplus_{j_2} \mathcal{D}^{j_2}$, Eq. (14.69), we consider the product state $|j_1 m_1\rangle \otimes |j_2 m_2\rangle = |j_1 j_2 m_1 m_2\rangle$ under the rotation $\exp(-i\vec{\omega} \cdot \vec{J}) = \exp(-i\vec{\omega} \cdot \vec{J}_1) \exp(-i\vec{\omega} \cdot \vec{J}_2)$ (using again the rotation properties of the states $|l, m\rangle$)

$$e^{-i\vec{\omega} \cdot \vec{J}} |j_1 j_2 m_1 m_2\rangle = \sum_{m'_1, m'_2} d_{m'_1 m_1}^{j_1} d_{m'_2 m_2}^{j_2} |j_1 j_2 m'_1 m'_2\rangle.$$

The matrix element with $\langle j_1 j_2 m'_1 m'_2 |$ gives

$$\langle j_1 j_2 m'_1 m'_2 | e^{-i\vec{\omega} \cdot \vec{J}} | j_1 j_2 m_1 m_2 \rangle = d_{m'_1 m_1}^{j_1} d_{m'_2 m_2}^{j_2}. \quad (14.71)$$

Inserting identities provides the following expression for the left-hand side

$$\begin{aligned} & \sum_{j, m, m'} \langle j_1 j_2 m'_1 m'_2 | \underbrace{|j_1 j_2 j m'\rangle \langle j_1 j_2 j m'|}_{\mathbb{1}} e^{-i\vec{\omega} \cdot \vec{J}} \underbrace{|j_1 j_2 j m\rangle \langle j_1 j_2 j m|}_{\mathbb{1}} | j_1 j_2 m_1 m_2 \rangle \\ &= \sum_{j, m, m'} \langle j_1 j_2 m'_1 m'_2 | j_1 j_2 j m' \rangle \underbrace{\langle j_1 j_2 j m' | e^{-i\vec{\omega} \cdot \vec{J}} | j_1 j_2 j m \rangle}_{d_{m' m}^j} \langle j_1 j_2 j m | j_1 j_2 m_1 m_2 \rangle \\ &= d_{m'_1 m_1}^{j_1} d_{m'_2 m_2}^{j_2}. \end{aligned} \quad (14.72)$$

(14.72) is precisely the relation used in (14.69) (with the substitution $j_1, j_2, m'_1, m'_2, j, m', m, m_1, m_2 \rightarrow k, j_1, q', m'_1, j_2, m'_2, m_2, q, m_1$).

14.5.1 Applications of the WET

Let \vec{A} be a vector operator, then A_q^1 is defined via

$$\begin{aligned} A_{-1}^1 &= (A_x - iA_y)/\sqrt{2}, \\ A_0^1 &= A_z, \\ A_1^1 &= -(A_x + iA_y)/\sqrt{2}, \end{aligned} \quad (14.73)$$

an irreducible spherical tensor of 1. order.

i) Let $\vec{A} = \vec{J}$, then

$$\underbrace{\langle \alpha' j j | J_z | \alpha j j \rangle}_{\hbar j \delta_{\alpha\alpha'}} = \frac{1}{\sqrt{2j+1}} \langle \alpha' j || J^1 || \alpha j \rangle \underbrace{\langle 1j0j | 1j j j \rangle}_{-\sqrt{j/(j+1)}} \quad (14.74)$$

and we find the reduced matrix element in the form

$$\langle \alpha' j' || J^1 || \alpha j \rangle = -\sqrt{j(j+1)(2j+1)} \hbar \delta_{\alpha\alpha'} \delta_{jj'}. \quad (14.75)$$

- ii) The m, m' dependence of the matrix elements of a vector operator A_q^1 is given by that of J_q^1 ,

$$\begin{aligned}\langle \alpha j m' | A_q^1 | \alpha j m \rangle &= \langle \alpha j m' | J_q^1 | \alpha j m \rangle \frac{\langle \alpha j \| A^1 \| \alpha j \rangle}{\langle \alpha j \| J^1 \| \alpha j \rangle} \\ &= \langle \alpha j m' | J_q^1 | \alpha j m \rangle \frac{\langle \alpha j m | \vec{J} \cdot \vec{A} | \alpha j m \rangle}{\hbar^2 j(j+1)},\end{aligned}\quad (14.76)$$

and analogously with $A_q^1 \rightarrow A_i$ and $J_q^1 \rightarrow J_i$. That is, the matrix elements of \vec{A} are equal to the matrix elements of the projection of \vec{A} on \vec{J} , $\langle \vec{A} \rangle \sim \langle \vec{J} \rangle \langle \vec{J} \cdot \vec{A} \rangle / \langle J^2 \rangle$.

ii.a) We have:

$$\frac{\langle \alpha j m' | A_q^1 | \alpha j m \rangle}{\langle \alpha j m' | J_q^1 | \alpha j m \rangle} = \frac{\langle \alpha j \| A^1 \| \alpha j \rangle}{\langle \alpha j \| J^1 \| \alpha j \rangle}. \quad (14.77)$$

ii.b)

$$\begin{aligned}\langle \alpha j m | \vec{J} \cdot \vec{A} | \alpha j m \rangle &= \sum_M \langle \alpha j m | \vec{J} | \alpha j M \rangle \langle \alpha j M | \vec{A} | \alpha j m \rangle \\ &= C \langle \alpha j \| J^1 \| \alpha j \rangle \langle \alpha j \| A^1 \| \alpha j \rangle.\end{aligned}\quad (14.78)$$

The constant C with the Clebsch-Gordon coefficients follows from the special case $\vec{A} = \vec{J}$,

$$\langle \alpha j m | J^2 | \alpha j m \rangle = C \langle \alpha j \| J^1 \| \alpha j \rangle^2, \quad (14.79)$$

and we find the relationship between the irreducible matrix elements

$$\frac{\langle \alpha j \| A^1 \| \alpha j \rangle}{\langle \alpha j \| J^1 \| \alpha j \rangle} = \frac{\langle \alpha j m | \vec{J} \cdot \vec{A} | \alpha j m \rangle}{\langle \alpha j m | J^2 | \alpha j m \rangle}. \quad (14.80)$$

Inserting in a) gives the desired result,

$$\langle \alpha j m' | A_q^1 | \alpha j m \rangle = \langle \alpha j m' | J_q^1 | \alpha j m \rangle \frac{\langle \alpha j m | \vec{J} \cdot \vec{A} | \alpha j m \rangle}{\hbar^2 j(j+1)}. \quad (14.81)$$

The third application takes us back to the Grotrian diagrams of the atoms.

- iii) We show that the matrix elements of H_{so} evaluated within an LS -multiplet can be expressed through the operator $\vec{L} \cdot \vec{S}$

$$\begin{aligned} \langle KLSM_L M_S | H_{\text{so}} | KLSM'_L M'_S \rangle &\equiv \langle H_{\text{so}} \rangle' \\ &= \xi(KLS) \langle KLSM_L M_S | \vec{L} \cdot \vec{S} | KLSM'_L M'_S \rangle. \end{aligned} \quad (14.82)$$

Here, $|KLSM_L M_S\rangle$ defines a basis in the multiplet with fixed L and S .

- iii.a) The spin-orbit coupling is defined via $H_{\text{so}} = \sum_i \vec{X}_i \cdot \vec{s}_i$ with

$$\vec{X}_i = -\frac{|e|\hbar}{2mc^2} \vec{v}_i \wedge \vec{E}_i, \quad (14.83)$$

\vec{v}_i are the velocities of the electrons and \vec{E}_i is the electric field evaluated at the location of the i th particle. The vector operator \vec{X}_i acts on the orbital component of $\Psi = |KLSM_L M_S\rangle$. If $V = V(r)$ is a spherically symmetric potential, \vec{X}_i can be written as

$$\vec{X}_i = \frac{1}{2m^2 c^2} \frac{1}{r_i} \frac{\partial V}{\partial r_i} \vec{l}_i, \quad (14.84)$$

and

$$H_{\text{so}} = \frac{1}{2m^2 c^2} \sum_i \frac{1}{r_i} \frac{\partial V}{\partial r_i} (\vec{l}_i \cdot \vec{s}_i). \quad (14.85)$$

- iii.b) We construct a product basis in \mathcal{H}_{LS} ,

$$|KLSM_L M_S\rangle = \sum_{\alpha} c_{\alpha} |\alpha KLM_L\rangle |\alpha SM_S\rangle \quad (14.86)$$

such that $|KLSM_L M_S\rangle$ is anti-symmetric.

- iii.c) The matrix elements $\langle H_{\text{so}} \rangle'$ can be written as

$$\begin{aligned} \langle KLSM_L M_S | H_{\text{so}} | KLSM'_L M'_S \rangle & \quad (14.87) \\ &= \sum_{i\alpha\alpha'} c_{\alpha}^* c_{\alpha'} \langle \alpha KLM_L | \vec{X}_i | \alpha' KLM'_L \rangle \langle \alpha SM_S | \vec{s}_i | \alpha' SM'_S \rangle. \end{aligned} \quad (14.88)$$

\vec{X}_i and \vec{s}_i are vector operators and are evaluated for total angular momentum and total spin states. Then according to the WET (14.76), we can use that

$$\begin{aligned} \langle \alpha KLM_L | \vec{X}_i | \alpha' KLM'_L \rangle &\propto \langle LM_L | \vec{L} | LM'_L \rangle, \\ \langle \alpha SM_S | \vec{s}_i | \alpha' SM'_S \rangle &\propto \langle SM_S | \vec{S} | SM'_S \rangle, \end{aligned} \quad (14.89)$$

i.e., the dependence on the magnetic quantum numbers M_L , M'_L and M_S , M'_S is determined by the matrix elements of \vec{L} and \vec{S} . Hence, the claim follows immediately from writing

$$\begin{aligned}\langle H_{\text{so}} \rangle' &= \sum_i \langle KLSM_L M_S | \vec{X}_i \cdot \vec{s}_i | KLSM'_L M'_S \rangle \\ &= \sum_i \sum_{\alpha\alpha'} \text{factor}(KLS i \alpha \alpha') \langle LSM_L M_S | \vec{L} \cdot \vec{S} | LSM'_L M'_S \rangle, \\ &= \xi(KLS) \langle KLSM_L M_S | \vec{L} \cdot \vec{S} | KLSM'_L M'_S \rangle.\end{aligned}\quad (14.90)$$

Given the matrix elements (14.90), we immediately find the fine structure of the LS -multiplet by transforming from the $|KLSM_L M_S\rangle$ basis to the basis $|KLSJM\rangle$ of the total angular momentum. We use the relation

$$2\vec{L} \cdot \vec{S} = J^2 - L^2 - S^2 \quad (14.91)$$

and find that¹³

$$\begin{aligned}\langle KLSJM | H_{\text{so}} | KLSJ'M' \rangle & \quad (14.92) \\ &= \delta_{JJ'} \delta_{MM'} [\xi(KLS)/2] [J(J+1) - S(S+1) - L(L+1)],\end{aligned}$$

resulting in a level splitting of the fine structure of magnitude

$$E_{KLS}(J) - E_{KLS}(J-1) = J \xi(KLS). \quad (14.93)$$

If $\xi > 0$ (one speaks of a regular multiplet), J is minimal in the ground state, if $\xi < 0$ (inverted multiplet), J is maximal in the ground state. In the following we show, using the Hartree-Fock approximation, that

$\xi > 0$	regular multiplet	if the shell is	$\leq 1/2$ filled,	hence proving
$\xi < 0$	inverted multiplet	if the shell is	$> 1/2$ filled,	

ing the fourth Hund's rule within the Hartree-Fock Scheme.

14.6 Spin-Orbit coupling in HF Theory*

In this section we calculate the coefficient $\xi(KLS)$ within Hartree-Fock theory. Let the potential $V(\vec{r}) = V(r)$ be spherically symmetrical, then the Hamiltonian H_{so} can be written as

$$H_{\text{so}} = \frac{1}{2m^2c^2} \sum_i \frac{1}{r_i} \frac{\partial V}{\partial r_i} \vec{l}_i \cdot \vec{s}_i, \quad (14.94)$$

¹³Note that $H_{\text{so}} = \xi(KLS)(\vec{L} \cdot \vec{S})$ in the LS -multiplet.

with \vec{l}_i and \vec{s}_i the orbital and spin angular momenta of the i th particle. With $M_L = M'_L$ and $M_S = M'_S$, we obtain for the factor $\xi(KLS)$ in (14.90) (the terms $L_x S_x = 0$ and $L_y S_y = 0$ do not contribute to the diagonal),

$$\xi(KLS) = \frac{1}{M_L M_S} \langle KLSM_L M_S | H_{\text{so}} | KLSM_L M_S \rangle. \quad (14.95)$$

We write $|KLSM_L M_S\rangle$ as a superposition of Slater determinants $|K\beta\rangle$, e.g., for carbon C in the configuration $K = 1s^2 2s^2 2p^2$, we have 15 Slater determinants¹⁴ $|K\beta\rangle$. The states $|K\beta\rangle$ are eigenfunctions of L_z and S_z with eigenvalues M_L and M_S . We are looking for the matrix element $\langle KLSM_L M_S | H_{\text{so}} | KLSM_L M_S \rangle$ in order to determine $\xi(KLS)$

$$\begin{aligned} \langle KLSM_L M_S | H_{\text{so}} | KLSM_L M_S \rangle &= \sum_{\beta, \beta'} c_{\beta}^* c_{\beta'} \langle K\beta | \frac{1}{2m^2 c^2} \sum_i [\partial_{r_i} V / r_i] \vec{l}_i \cdot \vec{s}_i | K\beta' \rangle \\ &= \sum_{\beta, \beta', i} c_{\beta}^* c_{\beta'} \frac{1}{2m^2 c^2} \langle K\beta | [\partial_{r_i} V / r_i] \vec{l}_i \cdot \vec{s}_i | K\beta' \rangle. \end{aligned} \quad (14.96)$$

In the matrix element $\langle K\beta | [\partial_{r_i} V / r_i] \vec{l}_i \cdot \vec{s}_i | K\beta' \rangle$ all orbitals in β and β' must be the same, except for the particle i . Furthermore, L_z and S_z are fixed. Therefore the orbitals for the i th particle must also be the same and only the diagonals survive, $\langle K\beta | \dots | K\beta' \rangle \propto \delta_{\beta\beta'}$. The summation over the occupied one-particle states in $|K\beta\rangle$ provides the following expression for the β term in $\langle H_{\text{so}} \rangle$ (note that ξ_{nl} does not depend on m_l because of the $\text{SO}(3)$ symmetry)

$$\begin{aligned} \langle H_{\text{so}} \rangle_{\beta} &= \sum_{m_l, m_s} \int d^3 r \varphi_{nlm_l}^*(\vec{r}) \frac{1}{2m^2 c^2} [\partial_r V / r] \vec{l} \varphi_{nlm_l}(\vec{r}) \cdot \underbrace{\langle m_s | \vec{s} | m_s \rangle}_{\text{only } \langle s_z \rangle = m_s \neq 0} \\ &= \frac{1}{2m^2 c^2} \sum_{m_l, m_s} m_l m_s \underbrace{\int d^3 r |\varphi_{nlm_l}|^2 [\partial_r V / r]}_{\substack{\text{SO}(3) \text{ invar.} \\ = \xi_{nl} > 0 \text{ for } \partial_r V > 0}} \\ &= \frac{1}{2m^2 c^2} \xi_{nl} \sum_{m_l, m_s} m_l m_s. \end{aligned} \quad (14.97)$$

¹⁴We choose K associated with the occupation of the last shell; as we will see, full shells do not contribute, therefore it suffices to use only Slater determinants describing electrons in the last partially-filled shell. The Slater determinants $|K\beta\rangle$ then involve only one radial wave function $\varphi_{nl}(r)$ and various angular and spin functions Y_{lm} and χ_s ; for carbon C the 15 Slater determinants are $|K\beta\rangle$.

With $\xi_{nl} \sim e^2/(Z^{-1/3}a_B)^3$, we have $\xi \sim (Z/m^2c^2a_B^2)(e^2/2a_B) = Z\alpha^2E_R \sim Z$ meV and the sign of $\sum m_l m_s$ determines the sign of ξ . We differentiate three cases:

1. **full shells:**

$$\sum_{m_l, m_s} m_l m_s = \sum_{m_l} m_l \sum_{m_s} m_s = 0, \quad (14.98)$$

which implies that full shells do not contribute (note that the sums over m_l and m_s disappear individually).

2. **shells $\leq 1/2$ filled:** The 2nd Hund's rule requires $S = \max$. Without loss of generality¹⁵ let us choose $M_S = S$. Then all occupied states have $m_s = 1/2$ and

$$\sum m_l m_s = \sum m_l S = M_L S. \quad (14.99)$$

Since all β terms in $\langle H_{SO} \rangle$ have the same quantum numbers M_L and M_S , we find the result

$$\langle H_{SO} \rangle = \xi(KLS) M_L S \quad \text{with} \quad \xi(KLS) = \frac{\xi_{nl}}{2m^2c^2} > 0. \quad (14.100)$$

3. **shells $> 1/2$ filled:** The 2nd Hund's rule requires $S = \max$ and we choose $M_S = S$. From $M_S = S = \max$, we conclude that all unoccupied states have a spin component $m_s = -1/2$. With

$$\sum_{\text{occupied}} m_l m_s + \sum_{\text{unoccupied}} m_l m_s = 0 \quad (14.101)$$

we find

$$\sum_{\text{occupied}} m_l m_s = - \sum_{\text{unoccupied}} m_l m_s = -(-M_L)(-S) = -M_L S, \quad (14.102)$$

and thus $\xi(KLS) = -\xi_{nl}/2m^2c^2 < 0$.

¹⁵ $\xi(KLS)$ is independent of M_S .

14.7 Zeeman splitting of the fine structure in a magnetic field

We consider atoms in a weak ($< 10^5$ Gauss) magnetic field $\vec{H} = (0, 0, H_z)$; the term generated by the magnetic field (we have $H = -\vec{\mu} \cdot \vec{H}$, $\vec{\mu} = -(e/2mc)\vec{L}$, with the charge $e > 0$)

$$H_z = \frac{e}{2mc}(\vec{L} + 2\vec{S}) \cdot \vec{H} \quad (14.103)$$

is handled within perturbation theory. The $(2J+1)$ -fold degeneracy remaining in the fine structure is lifted by the magnetic field. The term $\vec{L} + \vec{S} = \vec{J}$ in (14.103) contributes the amount $e\hbar MH_z/2mc = \mu_B MH_z$, with M the magnetic quantum number for J_z . We still have to calculate the remaining contribution

$$\begin{aligned} \langle KLSJM|S_z|KLSJM\rangle &\stackrel{(14.76)}{=} \\ \langle KLSJM|J_z|KLSJM\rangle &\frac{\langle KLSJM|\vec{J} \cdot \vec{S}|KLSJM\rangle}{\hbar^2 J(J+1)}. \end{aligned} \quad (14.104)$$

With $\vec{J} \cdot \vec{S} = \vec{S}^2 + \vec{L} \cdot \vec{S} = (J^2 - L^2 + S^2)/2$, we obtain

$$\langle KLSJM|S_z|KLSJM\rangle = \hbar M \frac{J(J+1) - L(L+1) + S(S+1)}{2J(J+1)} \quad (14.105)$$

and for the Zeeman splitting we find

$$\langle H_z \rangle = \frac{e\hbar}{2mc} g M H_z = g \mu_B M H_z \quad (14.106)$$

with the gyromagnetic factor g [also Landé- g -Factor]

$$g = 1 + \frac{J(J+1) - L(L+1) + S(S+1)}{2J(J+1)}. \quad (14.107)$$

The elementary moment of the atom then is

$$g\mu_B = g \frac{e\hbar}{2mc}. \quad (14.108)$$

For $S = 0$, we have $L = J$ and $g = 1$; conversely, for $L = 0$ the total 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. The LSJ multiplet splits into $2J + 1$ sub-levels with

$$\Delta E_Z = g\mu_B H_z. \quad (14.109)$$

Since the g factor depends on the quantum numbers LSJ , the level splitting in the spectrum can assume quite a complicated structure. One differentiates between the normal and anomalous Zeeman effects, with the anomalous one being the norm. The normal Zeeman effect occurs in the special situation that $S = 0$ and thus $g = 1$ for both levels involved; in this situation, an optical line splits three times and one finds a 'Zeeman triplet'. In the anomalous Zeeman effect, the g factors of the two levels are different, with orbital and spin angular momenta both finite and mixed differently. This results in more complicated line splittings. In figure 14.11, the two cases are shown for the Na D₁ transition [$^2S_{1/2}$ ($g = 2$) – $^2P_{1/2}$ ($g = 2/3$)], an example of an anomalous Zeeman effect, and for the Cd transition 1P_1 ($g = 1$) – 1D_2 ($g = 1$), a normal Zeeman effect.

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

$$\langle H_Z \rangle = \frac{e\hbar}{2mc}(M_L + 2M_S) H_z. \quad (14.110)$$

The perturbation term H_{SO} generates the correction

$$\langle H_{SO} \rangle = \xi(KLS)M_L M_S, \quad (14.111)$$

with $\xi(KLS) \sim Z\alpha^2 E_R$, $\alpha = e^2/\hbar c \approx 1/137$ the fine structure constant, see (14.97), and the combined results provide the total splitting

$$\langle H_Z + H_{SO} \rangle \approx \mu_B(M_L + 2M_S)H_z + \xi(KLS)M_L M_S, \quad (14.112)$$

where $\xi > 0$ (< 0) for \leq ($>$) half-filled shells. This is the Paschen-Back effect for the splitting of levels in strong magnetic fields; the mixing of the angular momenta \vec{L} and \vec{S} by the spin-orbit coupling H_{SO} is suppressed by the strong magnetic field and both angular momenta behave classically¹⁶ (except for the $g = 2$ factor of the spin). Both normal and anomalous Zeeman effects transform into the Paschen-Back effect at high magnetic fields. Conversely, the nontrivial anomalous Zeeman effect at weak fields results from the mixing of \vec{S} and \vec{L} to \vec{J} by H_{SO} .

¹⁶The angular momenta \vec{L} and \vec{S} precess independently of each other in the field.

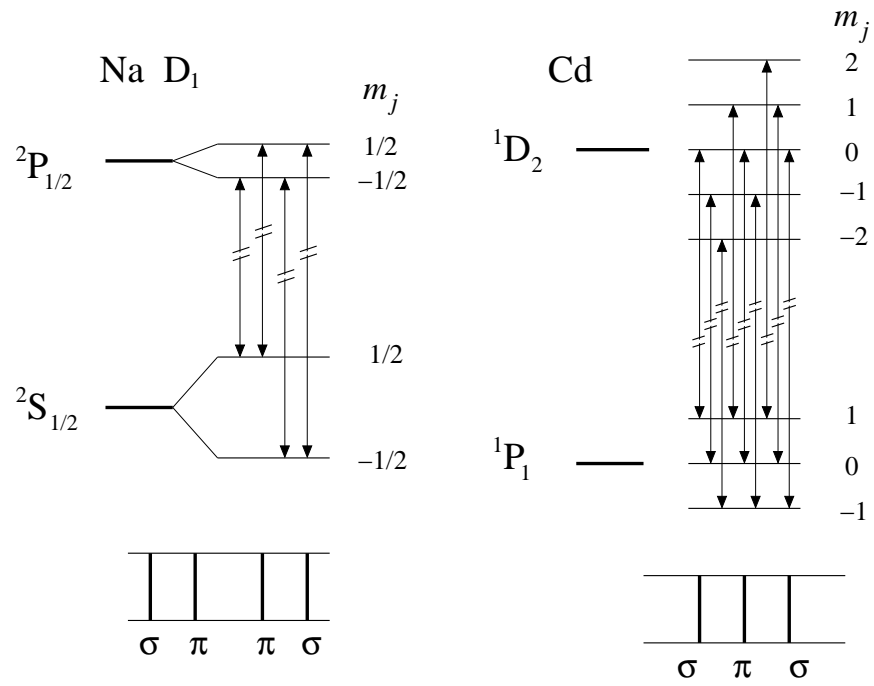


Fig. 14.11: Anomalous and normal Zeeman effect for the Na D₁ transition [$^2S_{1/2}$ ($g = 2$) – $^2P_{1/2}$ ($g = 2/3$)] and for the Cd [1P_1 ($g = 1$) – 1D_2 ($g = 1$)] transition. The π and σ transitions correspond to the selection rules $\Delta m = 0$ and $\Delta m = \pm 1$.

Chapter 15

Molecules

Material on this topic can be found in the textbooks on quantum mechanics by Gordon Baym and Leonard Schiff.

The description of molecules is a complex problem that can be tackled with approximative and numerical methods. We will approach the problem in three steps:

- We first discuss some qualitative estimates and identify two energy E scales in the problem, that of the electrons, E_e , and those of the nuclei E_{vib} and E_{rot} . The latter are smaller by the small parameter m_e/M , where m_e and M denote the electron and nuclear masses, from which follows, via Heisenberg's relation $\tau \sim \hbar/E$, that the dynamics of the nuclei is slow as compared to the one of the electrons.
- In the second step, we therefore assume the nuclei to be static and solve the problem of electrons, taking the positions \vec{R}_i of the nuclei as parameters. This calculation provides us with the electron energies $E_n^e(\vec{R}_i)$, where n denotes the electronic quantum numbers. Together with the nuclear potential $V_N(\vec{R}_i)$, these two components form the effective potential $\mathcal{E}_n(\vec{R}_i) = V_N + E_n^e$ for the motion of the nuclei.
- In the third and final step, we then examine the dynamics of the nuclei.

In pursuing this scheme, we assume that the electrons can always follow the nuclei instantaneously, an (adiabatic) approximation which is justified due to the smallness of the parameter m_e/M , which is said to constitute the small parameter in the problem. The separation into an electronic and

a nuclear problem is known as the Born-Oppenheimer Approximation. For an illustration of this scheme, we solve the H_2^+ and that H_2 problem (hydrogen molecules in the ionized and neutral state) and discuss several aspects of basic chemistry, covalent bonds, hybridization, and some other more complex molecules.

15.1 Born-Oppenheimer approximation

In this chapter, we are interested in the valence electrons that participate in the chemical bonds. The core electrons are hardly influenced by the presence of a binding partner. In contrast, valence electrons spread over the entire molecule (of size a a few Ångström) and have typical energies

$$E_{e-} \cong \frac{\hbar^2}{m_e a^2} \stackrel{a \sim 1\text{\AA}}{\sim} 1 - 10 \text{ eV}, \quad (15.1)$$

hence the electronic energy scale is of order of a few eV. We denote typical nuclear distances with R . Oscillations of the nuclei around their equilibrium separation R_0 define the energy scale $\hbar\omega$ of the core motion. To find $\hbar\omega$, we consider the potential energy $M\omega^2(R - R_0)^2$ of the nuclear vibration modes, where $R_0 \sim a$. For $R - R_0 \sim a$, the electron configuration is changed appreciably, e.g., the hydrogen molecule H_2 turns into a helium atom, and thus costs an electronic energy $\sim \hbar^2/m_e a^2$. Relating these two energies, $M\omega^2 a^2 \sim \hbar^2/m_e a^2$, we find the vibrational energy scale

$$E_{\text{vib}} \sim \hbar\omega \sim \sqrt{\frac{m_e}{M}} \frac{\hbar^2}{m_e a^2} \sim 0.1 \text{ eV}. \quad (15.2)$$

The nuclei not only vibrate but also rotate. Rotations cost an energy L^2/Θ , where $\Theta \sim Ma^2$ denotes the moment of inertia; with $L^2 \sim \hbar^2 l(l+1) \sim \hbar^2$, we find

$$E_{\text{rot}} \simeq L^2/\Theta \sim \frac{m_e}{M} \frac{\hbar^2}{m_e a^2} \sim 1 \text{ meV}. \quad (15.3)$$

The relation (15.2) also defines an important length scale, the amplitude δ of the vibration modes. The estimate $P^2/M \sim \hbar\omega$ provides us with the momentum scale $P \sim (M/m_e)^{1/4} \hbar/a$ and the velocity turns out to be $V = P/M \sim (M/m_e)^{1/4} \hbar/Ma = (m_e/M)^{3/4} v$, with $v \sim \hbar/ma$ the electron velocity. For the oscillation amplitude δ , we find

$$\delta \sim VT \sim V/\omega \sim a \left(\frac{m_e}{M} \right)^{1/4}, \quad (15.4)$$

hence typical vibrations are of order 1/10 of the molecular size. The same result follows from the estimate

$$M\omega^2\delta^2 \sim \hbar\omega \quad \rightarrow \quad \delta \sim \left(\frac{m_e}{M}\right)^{1/4} a. \quad (15.5)$$

The above energy scales define the associated time scales via $\tau \sim \hbar/E$.

electronic few eV	vibrational ~ 0.1 eV	rotational ~ 1 meV
\downarrow	\downarrow	\downarrow
10^{-16} s	10^{-14} s	10^{-12} s
$\underbrace{\hspace{10em}}$ \sim phonons in solid		

Next, we consider the Hamiltonian describing a molecule with nuclei of mass M_ν at positions \vec{R}_ν with momenta \vec{P}_ν and electrons with the coordinates \vec{r}_i and momenta \vec{p}_i (we summarize these coordinates under $\vec{r} = \vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_n$ and $\vec{R} = \vec{R}_1, \dots, \vec{R}_\nu, \dots, \vec{R}_m$, we only consider the orbital part of the energy and drop spin-orbit coupling), i

$$\begin{aligned}
 H = & \underbrace{\sum_\nu \frac{P_\nu^2}{2M_\nu}}_{T_N} + \underbrace{\sum_i \frac{p_i^2}{2m_e}}_{T_e} + \underbrace{\frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\vec{r}_i - \vec{r}_j|}}_{V_{ee}} \\
 & + \underbrace{\frac{1}{2} \sum_{\nu \neq \mu} \frac{Z_\nu Z_\mu e^2}{|\vec{R}_\nu - \vec{R}_\mu|}}_{V_{NN}} - \underbrace{\sum_{i, \mu} \frac{Z_\mu e^2}{|\vec{r}_i - \vec{R}_\mu|}}_{V_{eN}}, \quad (15.6)
 \end{aligned}$$

with the goal to find its eigenstates and eigenenergies from a solution of the Schrödinger equation

$$H\Psi(\vec{r}, \vec{R}) = E\Psi(\vec{r}, \vec{R}). \quad (15.7)$$

In applying the Born-Oppenheimer approximation, we treat the kinetic term T_N of the nuclei as a small perturbation $\propto m_e/M$. Neglecting T_N in the Hamiltonian, the variable \vec{R} loses its dynamics and becomes a parameter. The resulting electronic problem can be written as

$$\begin{aligned}
 [T_e + V_{ee}(\vec{r}) + V_{eN}(\vec{r}; \vec{R})] \varphi_n(\vec{r}; \vec{R}) &= \underbrace{[\mathcal{E}_n(\vec{R}) - V_{NN}(\vec{R})]}_{E_n^e(\vec{R})} \varphi_n(\vec{r}; \vec{R}). \quad (15.8) \\
 E_n^e(\vec{R}) &= \text{electronic energy}
 \end{aligned}$$

from which we obtain the electronic part $E_n^e(\vec{R})$ of the nuclear potential.

In order to formulate the nuclear problem, we go back to the full Schrödinger equation (15.7). The eigenfunctions $\varphi_n(\vec{r}; \vec{R})$ of the electronic problem form a CONS (complete orthonormal system) for each parameter set \vec{R} , and we can expand the total wave function Ψ ,

$$\Psi(\vec{r}; \vec{R}) = \sum_n \phi_n(\vec{R}) \varphi_n(\vec{r}; \vec{R}). \quad (15.9)$$

Substituting (15.9) into the SE (15.7) and making use of the electronic solution (15.8) results in the eigenvalue equation for the nuclear problem with dynamical coordinates \vec{R} ,

$$\sum_m [T_N + \mathcal{E}_m(\vec{R})] \phi_m(\vec{R}) \varphi_m(\vec{r}; \vec{R}) = E \sum_m \phi_m(\vec{R}) \varphi_m(\vec{r}; \vec{R}). \quad (15.10)$$

Integrating over the electronic degrees of freedom by constructing matrix elements with $\varphi_n^*(\vec{r}; \vec{R})$, we arrive at the system of equations describing the nuclear motion,

$$[T_N + \mathcal{E}_n(\vec{R})] \phi_n(\vec{R}) = E \phi_n(\vec{R}) - \sum_m A_{nm} \phi_m(\vec{R}), \quad (15.11)$$

with the operators¹

$$\begin{aligned} A_{nm} \phi_m(\vec{R}) = & - \sum_\nu \frac{\hbar^2}{2M_\nu} \int d\vec{r}^{3n} \varphi_n^*(\vec{r}; \vec{R}) \left[2\partial_{\vec{R}} \phi_m(\vec{R}) \cdot \partial_{\vec{R}} \varphi_m(\vec{r}; \vec{R}) \right. \\ & \left. + \phi_m(\vec{R}) \partial_{\vec{R}}^2 \varphi_m(\vec{r}; \vec{R}) \right]. \end{aligned} \quad (15.12)$$

Fortunately, the operator A_{nm} is small by a factor $(m_e/M)^{1/2}$ as compared with typical electronic energies ΔE^e . Here are some estimates: In the worst case,² we assume $\varphi_m(\vec{r}; \vec{R}) \approx \varphi_m(\vec{r} - \vec{R})$, such that $(1/M)\partial_{\vec{R}}^2 \varphi_m \leq (m_e/M)\partial_{\vec{r}}^2 \varphi_m/m_e$ and the second term of A_{nm} is of order

$$A_{nm}^{(2)} \phi_m \sim \frac{m_e}{M} \frac{\hbar^2}{2m_e a^2} \phi_m \sim \frac{m_e}{M} \Delta E^e \phi_m, \quad (15.13)$$

¹Note that

$$T_N \phi \varphi \propto \partial_{\vec{R}}^2 \phi \varphi \sim \underbrace{\varphi \partial_{\vec{R}}^2 \phi}_{T_N \phi_n} + \underbrace{2\partial_{\vec{R}} \phi \cdot \partial_{\vec{R}} \varphi + \phi \partial_{\vec{R}}^2 \varphi}_{A_{nm} \phi_m}.$$

²In general, the \vec{R} dependence is weaker than that on \vec{r} as the two are not rigidly connected.

small by the factor m_e/M . To estimate $\partial_{\vec{R}}\phi$, we approximate ϕ by an oscillator wavefunction $\phi(\vec{R}) \sim \exp[-M\omega(R - R_0)^2/2\hbar]$ and find that (with the vibration amplitude $\vec{\delta} = \vec{R} - \vec{R}_0$)

$$\begin{aligned}\partial_{\vec{R}}\phi \cdot \partial_{\vec{R}}\varphi &\sim \frac{M\omega}{\hbar}\phi(\vec{R})(\vec{R} - \vec{R}_0) \cdot \partial_{\vec{R}}\varphi(\vec{r}; \vec{R}) \\ &\sim \frac{M\omega}{\hbar}\phi(\vec{R})[\varphi(\vec{r}; \vec{R} + \vec{\delta}) - \varphi(\vec{r}; \vec{R})]\end{aligned}\quad (15.14)$$

and the first term in A_{nm} contributes with

$$A_{nm}^{(1)}\phi_m \sim \frac{\hbar^2}{2M} \frac{M\omega}{\hbar}\phi_m \sim \hbar\omega\phi_m \sim \left(\frac{m_e}{M}\right)^{1/2} \Delta E^e \phi_m, \quad (15.15)$$

which is small by the factor $(m_e/M)^{1/2}$.

Going back to the coupled eigenvalue problem (15.11), we note that different n -components involve electron energies E_n^e that reside on the scale $\hbar^2/2m_e a^2$. In addition, the term $A_{nm}\phi_m$ mixes different n -components, but only if the mixing ‘amplitudes’ can match the electronic scale, $A_{nm} \sim E^e$. With our estimates (15.13) and (15.15), the energies associated with A_{nm} are not sufficient to mix different levels and we obtain the following simple Schrödinger equation for the nuclear problem (in the adiabatic approximation, we set $A_{nm} \sim 0$ and exclude ‘high-energy’ electronic transitions),

$$[T_N + \mathcal{E}_n(\vec{R})] \phi_n(\vec{R}) \approx E \phi_n(\vec{R}). \quad (15.16)$$

We will see, that favorable (for molecular bindings) electronic configurations form an attractive nuclear potential as sketched in Fig. 15.1, otherwise binding into a molecule is not possible. The eigenstates in the Born-Oppenheimer approximation are given by simple product functions

$$\Psi(\vec{r}; \vec{R}) \approx \phi_n(\vec{R}) \varphi_n(\vec{r}; \vec{R}), \quad (15.17)$$

where n is the electronic level index. The nuclear problem (15.16) describes the vibrational and rotational excitations, $\phi_n \rightarrow \phi_{n,m,l}$ and $E \rightarrow E_{m,l}$ with m and l the vibrational and rotational quantum numbers; the corresponding energies $\hbar\omega_n$ and $\hbar^2/2\Theta$ follow from the curvature $\partial_{\vec{R}}^2 \mathcal{E}_n|_{\vec{R}_{n0}} = M\omega_n^2$ of the electronic potential $\mathcal{E}_n(\vec{R})$ and the moment of inertia $\Theta \sim MR_{n0}^2$.

The electronic energy is $E_n^e(\vec{R}_{n0}) = \mathcal{E}_n(\vec{R}_{n0}) - V_{NN}(\vec{R}_{n0})$ and the total energy is $E \approx \mathcal{E}_n(\vec{R}_{n0}) + \hbar\omega_n/2 + E_{\text{vib}} + E_{\text{rot}} = E_n^e(\vec{R}_{n0}) + V_{NN}(\vec{R}_{n0}) + \hbar\omega_n/2 + E_{\text{vib}} + E_{\text{rot}}$. In the following, we will first focus on the electron problem of simple

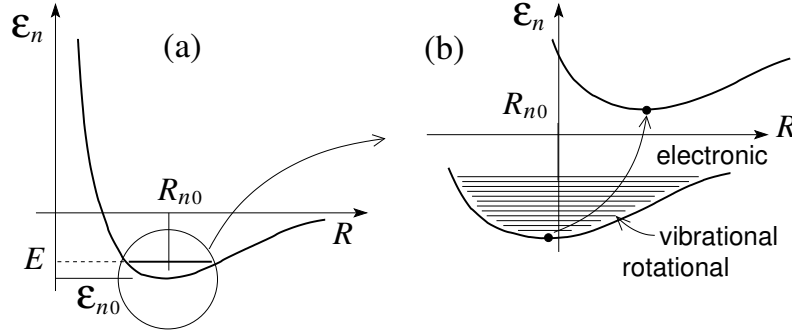


Fig. 15.1: (a) Electronic energy $\mathcal{E}_n = E_n^e + V_{\text{NN}}$ as a function of the distance R between the nuclei. The equilibrium distance \vec{R}_{n0} of the nuclei is defined by the minimum in \mathcal{E}_n , $\partial_{\vec{R}} \mathcal{E}_n|_{\vec{R}_{n0}} = 0$. (b) The curvature $\partial_{\vec{R}}^2 \mathcal{E}_n|_{\vec{R}_{n0}} = M\omega_n^2$ defines the vibration spectrum, the moment of inertia Θ (in general a tensor) defines the rotational bands of the molecules. The electronic excitation energies exceed by far the rotational and vibrational excitations of the molecule.

molecules (such as H_2^+ and H_2 , where we calculate the energies $\mathcal{E}_n(\vec{R})$) and then on the nuclear problem (with its vibrational and rotational excitations).

Note that highly excited vibrational states that deform the molecule on a scale a can convert their excitation energy into electronic excitations as shown in 15.1. This makes transitions between different \mathcal{E}_n via A_{nm} possible and the adiabatic approximation breaks down. The same applies to highly-excited rotational states.

15.2 The electronic problem, H_2^+ ion

The H_2^+ ion is described by the Hamiltonian

$$H = -\frac{\hbar^2}{2m_e} \nabla^2 - \frac{e^2}{|\vec{r} - \vec{R}_A|} - \frac{e^2}{|\vec{r} - \vec{R}_B|} + \frac{e^2}{|\vec{R}_A - \vec{R}_B|} \quad (15.18)$$

see Fig. 15.2 for the definition of the nuclear and electronic coordinates. In the Hamilton (15.18), we have neglected the kinetic part of the nuclei, so that \vec{R}_A and \vec{R}_B are only parameters of the electronic problem. Our goal is to calculate the ground state configuration (i.e., the spacing R between nuclei and the energy) of the H_2^+ molecular ion; in order to do so, we determine a variational energy $\langle H \rangle(R)$ as function of spacing R .

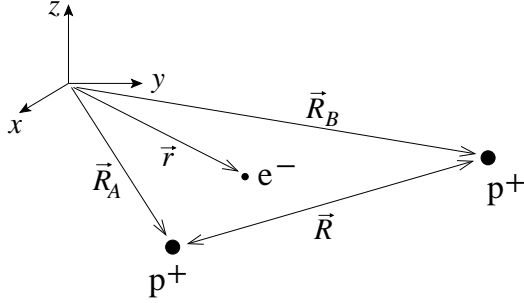


Fig. 15.2: Coordinates used in the description of the H_2^+ molecular ion.

A reasonable variational ansatz involves the linear combination of hydrogen $1s$ -wavefunctions centered in \vec{R}_A and \vec{R}_B ,

$$\begin{aligned}\varphi(\vec{r}, \vec{R}) &= \alpha\varphi_A(\vec{r}) + \beta\varphi_B(\vec{r}), \\ \varphi_X(\vec{r}) &= \sqrt{\frac{1}{\pi a_B^3}} e^{-|\vec{r}-\vec{R}_X|/a_B}.\end{aligned}\quad (15.19)$$

The problem has an inversion point, with the symmetry $\vec{r} \rightarrow -\vec{r}$ with respect to the center of mass $(\vec{R}_A + \vec{R}_B)/2$, and we can search for eigenfunctions with defined parity, that is $\alpha = \beta$ for even and $\alpha = -\beta$ for odd parity,

$$\varphi_{\pm} = N_{\pm}[\varphi_A(\vec{r}) \pm \varphi_B(\vec{r})]. \quad (15.20)$$

Normalizing the wave functions to 1 defines the normalization factors N_{\pm}

$$\begin{aligned}1 \equiv \int d^3r |\varphi_{\pm}|^2 &= N_{\pm}^2 \left[\underbrace{2 \int d^3r |\varphi_A|^2}_1 \pm 2 \underbrace{\int d^3r \varphi_A(\vec{r}) \varphi_B(\vec{r})}_S \right], \\ \Rightarrow N_{\pm} &= \frac{1}{\sqrt{2} \sqrt{1 \pm S(R)}}.\end{aligned}\quad (15.21)$$

The overlap integral $S = \int d^3r \varphi_A \varphi_B$ depends on the distance R and is conveniently calculated in elliptical coordinates (see Fig. 15.3). With the definitions

$$\begin{aligned}\mu &= \frac{1}{R} (|\vec{r} - \vec{R}_A| + |\vec{r} - \vec{R}_B|), \\ \nu &= \frac{1}{R} (|\vec{r} - \vec{R}_A| - |\vec{r} - \vec{R}_B|), \\ \int d^3r &= \int_1^{\infty} d\mu \int_{-1}^1 d\nu \int_0^{2\pi} d\varphi \frac{R^3}{8} (\mu^2 - \nu^2),\end{aligned}\quad (15.22)$$

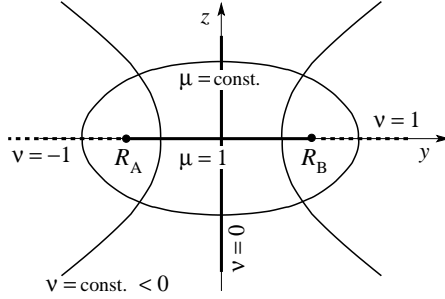


Fig. 15.3: Elliptical coordinates $\mu = (|\vec{r} - \vec{R}_A| + |\vec{r} - \vec{R}_B|)/R$ and $\nu = (|\vec{r} - \vec{R}_A| - |\vec{r} - \vec{R}_B|)/R$ for the calculation of the overlap integral S in H_2^+ problem.

we find the expression for the overlap integral (with $x = R/a_B$)

$$\begin{aligned}
 S &= \int d^3r \varphi_A \varphi_B = \int d\mu d\nu d\varphi \frac{R^3}{8} (\mu^2 - \nu^2) \frac{1}{\pi a_B^3} e^{-R\mu/a_B} \\
 &= \frac{x^3}{4} \int d\mu d\nu [\partial_x^2 - \nu^2] e^{-\mu x} = \frac{x^3}{4} \left[2\partial_x^2 \left(\frac{e^{-x}}{x} \right) - \frac{2}{3} \frac{e^{-x}}{x} \right] \\
 &= \left(1 + \frac{R}{a_B} + \frac{R^2}{3a_B^2} \right) e^{-R/a_B}. \tag{15.23}
 \end{aligned}$$

The expectation value $\langle H \rangle_{\pm} = \mathcal{E}_{\pm}(R)$ can be written as a combination of the matrix elements

$$\begin{aligned}
 \langle A|H|A \rangle &= \int d^3r \varphi_A^* H \varphi_A \quad \text{and} \quad \langle A|H|B \rangle = \int d^3r \varphi_A^* H \varphi_B, \\
 \mathcal{E}_{\pm}(R) &= \frac{\langle A|H|A \rangle \pm \langle A|H|B \rangle}{1 \pm S}. \tag{15.24}
 \end{aligned}$$

With the ground state energy \mathcal{E}_1 of the H atom, we obtain the following results for the matrix elements

$$\begin{aligned}
 \langle A|H|A \rangle &= \mathcal{E}_1 + \frac{e^2}{R} - \underbrace{\int d^3r \frac{e^2 \varphi_A^2}{|\vec{r} - \vec{R}_B|}}_{\frac{e^2 x^3}{2R} \int d\mu d\nu e^{-x(\mu+\nu)} (\mu+\nu) = \frac{e^2}{R} [1 - (1+R/a_B) e^{-2R/a_B}]} \\
 &= \mathcal{E}_1 + \frac{e^2}{R} \left(1 + \frac{R}{a_B} \right) e^{-2R/a_B} \tag{15.25}
 \end{aligned}$$

and

$$\begin{aligned}
 \langle A|H|B \rangle &= \left(\mathcal{E}_1 + \frac{e^2}{R} \right) S - \underbrace{\int d^3r \frac{e^2 \varphi_A \varphi_B}{|\vec{r} - \vec{R}_B|}}_{\frac{e^2 x^3}{2R} \int d\mu d\nu e^{-\mu x} (\mu+\nu)} \tag{exchange integral}
 \end{aligned}$$

$$= \left(\mathcal{E}_1 + \frac{e^2}{R} \right) S - \frac{e^2}{a_B} \left(1 + \frac{R}{a_B} \right) e^{-R/a_B}. \quad (15.26)$$

Finally, if we plot the energies \mathcal{E}_\pm as a function of the distance R between the two protons, see Fig. 15.4, we find for φ_+ a bound state with $R_0 \approx 1.3$ Å (instead of the experimental value 1.06 Å) and $\mathcal{E}_0 \approx -1.76$ eV (instead of -2.8 eV in the experiment). The even wavefunction φ_+ is finite in the area between the protons and allows the system to efficiently gain energy from the Coulomb interaction.

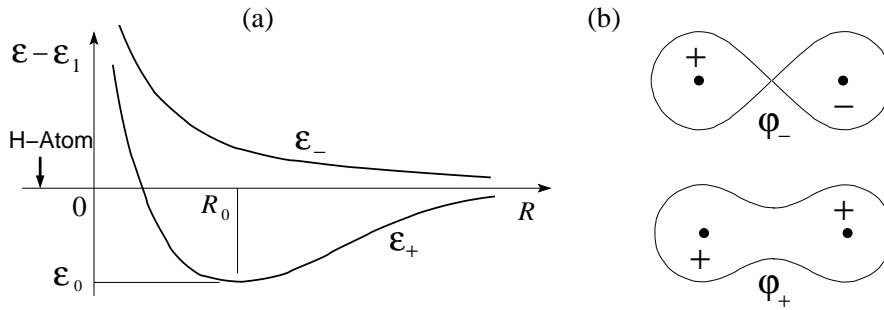


Fig. 15.4: (a) Energy \mathcal{E}_\pm as a function of R for one-electron states φ_\pm with even (+) and odd (−) parity, see (b). The state φ_+ with \mathcal{E}_+ forms a minimum as a function of the spacing R between the protons that results in a bound state. The protons remain unbound for the odd state φ_- .

The odd state φ_- , on the other hand, cannot bind the protons; with the negative (−) parity, $\varphi_- = 0$ in the center between the protons and the loss of negative Coulomb energy prevents binding. Thus, we find that

φ_+ produces a bonding orbital,
and φ_- produces an anti-bonding orbital.

Finally, we analyze the different limits $R \rightarrow 0$ and $R \rightarrow \infty$:

φ_+ : correctly describes the limit $R \rightarrow \infty$, with the electron delocalized between the two protons and the tunnel amplitude $\exp(-R/a_B) \rightarrow 0$. The region of overlapping wave functions at $R \rightarrow 0$ is poorly described because $\text{H}_2^+(R \rightarrow 0)$ corresponds to He^+ but the wave function $\varphi_+(R \rightarrow 0) \rightarrow \varphi_{\text{H-atom}}$. Accordingly, the resulting electronic binding energy $E_0^e(R_0 \rightarrow 0) = \mathcal{E}_+ - e^2/R \rightarrow -3$ is too small compared to the He^+ ion with $E_0^e[\text{He}] = -4$ Ry.

φ_- : for $R \rightarrow 0$ behaves like $\varphi_- \sim \cos \theta e^{-r/a_B}$, i.e., like an excited p state.

Note also the analogy between the reactions $\text{H}_2^+ = p + p + e^- \rightleftharpoons (\text{HD})^+ = p + d + \mu^-$. In the muon-catalyzed fusion of a proton and a deuterium, the muon plays the role of a catalyst for the reaction $p + d \rightarrow {}^3\text{He}^{2+} + 5.4 \text{ MeV}$. The heavy (compared to the electron) muon creates a much smaller p - d molecule.

15.3 H_2 molecule

The H_2 molecule is described by the Hamiltonian

$$H = -\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2) + \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} + \frac{e^2}{|\vec{R}_A - \vec{R}_B|} \quad (15.27)$$

$$- \left[\frac{e^2}{|\vec{r}_1 - \vec{R}_A|} + \frac{e^2}{|\vec{r}_1 - \vec{R}_B|} + \frac{e^2}{|\vec{r}_2 - \vec{R}_A|} + \frac{e^2}{|\vec{r}_2 - \vec{R}_B|} \right]$$

with the coordinates for the electrons ($\vec{r}_{1,2}$) and for the protons ($\vec{R}_{A,B}$) as sketched in Fig. 15.5.

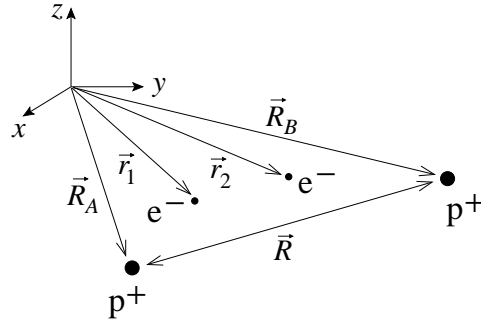


Fig. 15.5: Coordinates used in the description of the H_2 molecule.

In our analysis of the stationary Schrödinger problem, we examine two variational functions: molecular orbitals (due to Hund and Mullikan) and valence orbitals (Heitler-London).

Molecular orbitals (Hund-Mullikan)

$$\Psi_s^M = \frac{1}{2(1+S)} [\varphi_A(\vec{r}_1) + \varphi_B(\vec{r}_1)] [\varphi_A(\vec{r}_2) + \varphi_B(\vec{r}_2)] \chi_s$$

$$= \underbrace{\varphi_+(\vec{r}_1)\varphi_+(\vec{r}_2)}_{\text{symmetric}} \underbrace{\chi_s}_{\text{antisymm.}}$$

$$\Psi_t^M = \frac{1}{\sqrt{2}} \underbrace{[\varphi_+(\vec{r}_1)\varphi_-(\vec{r}_2) - \varphi_-(\vec{r}_1)\varphi_+(\vec{r}_2)]}_{\text{antisymmetric}} \underbrace{\chi_t}_{\text{symm.}} \quad (15.28)$$

Valence orbitals (Heitler-London)

$$\begin{aligned} \Psi_s^V &= \frac{1}{\sqrt{2(1+S^2)}} [\varphi_A(\vec{r}_1)\varphi_B(\vec{r}_2) + \varphi_B(\vec{r}_1)\varphi_A(\vec{r}_2)] \chi_s, \\ \Psi_t^V &= \frac{1}{\sqrt{2(1-S^2)}} [\varphi_A(\vec{r}_1)\varphi_B(\vec{r}_2) - \varphi_B(\vec{r}_1)\varphi_A(\vec{r}_2)] \chi_t. \end{aligned} \quad (15.29)$$

These functions show the following properties/behavior in the limits $R \rightarrow 0, \infty$. First of all, Ψ_t^M involves the odd wavefunction φ_- and thus an anti-bonding orbital; accordingly Ψ_t^M is energetically unfavorable compared to Ψ_s^M . In the limit $R \rightarrow \infty$, we find the components

$$\begin{aligned} \Psi_s^M &\sim \varphi_A(1)\varphi_A(2) + \varphi_A(1)\varphi_B(2) \\ &\quad + \underbrace{\varphi_B(1)\varphi_B(2)}_{\text{unfavorable}} + \underbrace{\varphi_B(1)\varphi_A(2)}_{\text{correct}} \end{aligned} \quad (15.30)$$

unfavorable \rightarrow $p + \text{H}^-$ $\text{H} + \text{H}$ \leftarrow correct

for the molecular orbital Ψ_s^M . On the other hand, the valence orbitals $\Psi_{s,t}^V$ exhibit the correct $R \rightarrow \infty$ behavior, corresponding to two hydrogen atoms $\text{H} + \text{H}$. In the limit $R \rightarrow 0$, we find that

$$\begin{aligned} \Psi_s^M &\sim 1s^2 \text{ with } Z = 1 \text{ instead of } Z = 2, \\ \Psi_s^V &\sim 1s^2 \text{ with } Z = 1 \text{ instead of } Z = 2, \\ \Psi_t^V &\sim \frac{\varphi_A(\vec{r}_1) - \varphi_B(\vec{r}_1)}{\sqrt{1-S}} \frac{\varphi_A(\vec{r}_2) + \varphi_B(\vec{r}_2)}{\sqrt{1+S}}, \\ &\sim \cos\theta_1 \exp[-(r_1 + r_2)/a_B] \sim sp\text{-state}. \end{aligned} \quad (15.31)$$

The sp -state (electron 1 in p , electron 2 in s) involves the angular momentum $l = 1$ and thus leads to higher energies.

Overall, the Heitler-London functions $\Psi_{s,t}^V$ are somewhat better asymptotically. Both types (HM and HL) give reasonable (and numerically similar) variational results.³ We continue with the HL functions and calculate the variational expectation values $\mathcal{E}_\pm(R)$ from the relevant matrix elements (with the indices $+/- \hat{=} s/t$)

$$\mathcal{E}_\pm(R) = \langle H \rangle_\pm = \frac{\langle AB|H|AB \rangle \pm \langle BA|H|AB \rangle}{1 \pm S^2}, \quad (15.32)$$

³The limit $R \rightarrow \infty$ is not numerically relevant for the final result, since $R_0 \sim a_B$ is finite.

where we used that $\langle AB|H|AB\rangle = \langle BA|H|BA\rangle$ and $\langle AB|H|BA\rangle = \langle BA|H|AB\rangle$. For the matrix elements, we find

$$\begin{aligned}
\langle AB|H|AB\rangle &= 2\mathcal{E}_1 + \frac{e^2}{R} \\
&\quad + \underbrace{\int d^3r \int d^3r' \varphi_A^2(\vec{r}) \varphi_B^2(\vec{r}') \left[\frac{e^2}{|\vec{r} - \vec{r}'|} - \frac{e^2}{|\vec{r}' - \vec{R}_A|} - \frac{e^2}{|\vec{r} - \vec{R}_B|} \right]}_{V_C(R)} \\
\langle BA|H|AB\rangle &= \left[2\mathcal{E}_1 + \frac{e^2}{R} \right] S^2 \\
&\quad + \underbrace{\int d^3r \int d^3r' \varphi_A(\vec{r}) \varphi_B(\vec{r}) \varphi_A(\vec{r}') \varphi_B(\vec{r}') \left[\frac{e^2}{|\vec{r} - \vec{r}'|} - \frac{e^2}{|\vec{r}' - \vec{R}_A|} - \frac{e^2}{|\vec{r} - \vec{R}_B|} \right]}_{V_{\text{exc}}(R) \propto S^2}.
\end{aligned} \tag{15.33}$$

We identify the contribution $V_C(R)$ with the remaining Coulomb energy, where the terms $e^2/|\vec{r} - R_A|$ and $e^2/|\vec{r}' - R_B|$ are already taken into account in $2\mathcal{E}_1$, while the mixed terms $e^2/|\vec{r}' - R_A|$ and $e^2/|\vec{r} - R_B|$ and the $e-e$ interaction go into V_C . The contribution $V_{\text{exc}}(R)$ is the exchange term belonging to V_C ⁴ and is present due to the antisymmetrization of the wavefunction. This term is critical for the existence of the bound state.

The calculation of V_C and V_{exc} is a bit tedious⁵; here, we simply write

$$\mathcal{E}_{\pm}(R) = 2\mathcal{E}_1 + \frac{[V_C(R) + e^2/R] \pm [V_{\text{exc}}(R) + S^2 e^2/R]}{1 \pm S^2}, \tag{15.34}$$

with $V_C(R) + e^2/R > 0$; the exchange term is usually negative, $V_{\text{exc}}(R) + S^2 e^2/R < 0$. Therefore, the singlet state is more favorable in energy, since it makes better use of the exchange term, $\mathcal{E}_+ < \mathcal{E}_-$. It turns out that $\mathcal{E}_+(R) - 2\mathcal{E}_1$ has a 3 eV deep minimum at $R_0 = 1.5 a_B$, whereas $\mathcal{E}_-(R)$ does not produce a minimum and thus there is no bound state in the triplet channel, see the sketch in Fig. 15.6. Again, the bonding nature of the singlet state⁶ is due to the finite amplitude of the wavefunction in the middle between the nuclei. This result should actually lead to some irritation, because we have often argued that the Pauli principle helps to separate the electrons from each other and thereby *reduce* the repulsive energy (> 0)

⁴described by the so-called ‘exchange’ integral.

⁵See Y. Sugiura, Z. Phys. **45**, 484 (1927).

⁶The triplet state does not bind.

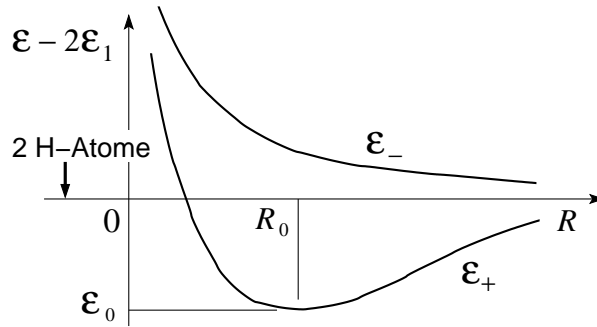


Fig. 15.6: The singlet state with the symmetric orbital function produces the bound state of the hydrogen molecule at $\epsilon_0 \approx 3$ eV and $R_0 \approx 0.8$ Å (the experimentally measured value is $R_0 \approx 0.74$ Å). The binding energy is produced by the exchange term $V_{\text{exc}} \propto S^2$ and involves the overlap integral S .

between the electrons. For example, the 2nd Hund rule says that S is maximal. Accordingly, the triplet state of $1s2s$ in the He atom is lower than the singlet combination.⁷ Based on these arguments, Ψ_t^V should win. But: In the Ψ_s^V state, we have to pay for some Coulomb energy due to the $e^- - e^-$ interaction, but we gain even more Coulomb energy from the e^- -proton interactions if the electrons are in the plane of symmetry (orthogonal to the axis), that is an overcompensation effect.

Finally, we consider the limit cases $R \rightarrow 0$ and $R \rightarrow \infty$ again. For $R \rightarrow 0$ we find H orbitals instead of He orbitals with $Z = 2$. Accordingly, the energy of the valence functions is too high, which means that the actual binding energy is larger and the distance between the protons is smaller, in agreement with the experimentally measured smaller distance (0.74 Å instead of 0.8 Å). In the limit $R \rightarrow \infty$, the interaction actually does not decrease exponentially (as calculated here) but rather algebraically $\propto 1/R^6$, an effect of the Van der Waals attraction between neutral atoms due to fluctuating (induced) dipoles. The dipole fluctuations of the two H atoms involve excited states of the H atoms and can be considered in perturbation theory. Here, we have only considered the ground state of the H atoms.

⁷The ground state of He is $1s^2$ and must be a singlet, see (15.30).

15.4 The nuclear problem

The nuclear problem involves vibration and rotation modes. With $\vec{R} = \vec{R}_A - \vec{R}_B$, the Born-Oppenheimer equation (15.16) reads

$$\left[-\frac{\hbar^2}{2\mu} \vec{\nabla}^2 + \mathcal{E}(R) \right] \phi(\vec{R}) = E \phi(\vec{R}), \quad (15.35)$$

with the reduced mass⁸ $\mu = M_1 M_2 / (M_1 + M_2)$ and the potential $\mathcal{E}(R) = \mathcal{E}_+(R)$ as sketched in Fig. 15.6. For the rotationally symmetric problem (15.35), we make the ansatz

$$\phi(\vec{R}) = R_{nl}(R) Y_{lm}(\theta, \varphi) \quad (15.36)$$

and arrive at the radial problem

$$\left[-\frac{\hbar^2}{2\mu} \left(\partial_R^2 + \frac{2}{R} \partial_R \right) + \mathcal{E}(R) + \frac{\hbar^2 l(l+1)}{2\mu R^2} \right] R_{nl}(R) = E R_{nl}(R). \quad (15.37)$$

With $R_{nl} = u_{nl}/R$, we find the 1D problem

$$\left[-\frac{\hbar^2}{2\mu} \partial_R^2 + V_{\text{eff}}(R) \right] u_{nl} = E u_{nl} \quad (15.38)$$

with

$$V_{\text{eff}}(R) = \mathcal{E}(R) + \frac{\hbar^2 l(l+1)}{2\mu R^2} \approx \overbrace{V_{\text{eff}}(R_0^l)}^{\text{minimum of } V_{\text{eff}}} + \overbrace{\frac{\mu\omega^2}{2} (R - R_0^l)^2}^{\mu\omega^2 = \partial_R^2 V_{\text{eff}}}. \quad (15.39)$$

The transformation to $x = R - R_0^l$ results in the harmonic oscillator problem

$$\left[-\frac{\hbar^2}{2\mu} \partial_x^2 + \mathcal{E}(R_0^l) + \frac{\hbar^2 l(l+1)}{2\mu (R_0^l)^2} + \frac{\mu\omega^2}{2} x^2 \right] u_{nl} = E_{nl} u_{nl} \quad (15.40)$$

with energy eigenvalues

$$E_{nl} = \overbrace{\mathcal{E}(R_0^l)}^{E_e + V_{\text{NN}}} + \underbrace{\frac{\hbar^2 l(l+1)}{2\mu (R_0^l)^2}}_{\text{rotations}} + \underbrace{\hbar\omega \left(n + \frac{1}{2} \right)}_{\text{vibrations}}. \quad (15.41)$$

⁸After separation into center of mass and relative problems, see chapter 5 of quantum mechanics I course, we focus on the latter.

In molecules with identical nuclei, we have to take into account the symmetry of the nuclear wave function under the exchange of the nuclei.⁹ For nuclei with an integer (half-integer) spin S , the wavefunction must be symmetric (antisymmetric).¹⁰ The exchange of the nuclei just corresponds to the parity operation $\vec{R} \rightarrow -\vec{R}$. We separate the nuclear wavefunction Ψ into an orbital component $\Phi(\vec{R})$ and a spin component χ . The symmetry of the orbital wavefunction is determined by the angular momentum quantum number l ,¹¹ $P = (-1)^l$. Then the following relationships apply,

integer nuclear spin	half-integer nuclear spin	
symmetric wavefunction	antisymmetric wavefunction	
$\chi_{\text{symm}}, \quad l \text{ even}$	$\chi_{\text{antisym}}, \quad l \text{ even}$	
$\chi_{\text{antisym}}, \quad l \text{ odd}$	$\chi_{\text{symm}}, \quad l \text{ odd}$	(15.42)

With S the spin of a nucleus, there are $(2S + 1)^2$ spin states that separate into $(S + 1)(2S + 1)$ symmetric and $S(2S + 1)$ antisymmetric states (e.g., check with singlets and triplets for $S = 1/2$). In a gas of such molecules, the ratio of the number of molecules with even l ($\# l \text{ even}$) to those with odd l ($\# l \text{ odd}$) is given by (with NS = nuclear spin, int = integer, h-int = half-integer)

$$\text{int NS} \rightarrow \frac{S + 1}{S} = \frac{\# l \text{ even}}{\# l \text{ odd}} = \frac{S}{S + 1} \leftarrow \text{h-int NS.} \quad (15.43)$$

The investigation of the relative intensities of the rotational bands of a gas in equilibrium¹² then allows for the determination of the statistics of the nuclei and their spin S .

15.5 Chemical bonds

We examine the nature of the chemical bond and its saturation, different types of bonds (σ and π bonds), the phenomenon of hybridization, and

⁹See also the analysis of diatomic molecules in 13.6.11

¹⁰The spin of the nucleus involves both orbital and spin parts. Cores with each an even number of neutrons and of protons have nuclear spin $I = 0$; cores with each an odd number of neutrons and of protons have an integer nuclear spin. The remaining cores with odd/even number of constituents have half-integer spin.

¹¹The spherical harmonics Y_{lm} has parity $(-1)^l$.

¹²We assume that the temperature is higher than the excitation energy of the rotational bands.

provide a number of illustrative examples.

Consider the simplest case of two H atoms; we found, see Fig. 15.6, that the singlet state produces a bound state, whereas the triplet state does not produce a bond,

$$\begin{aligned}\chi_s: \quad \uparrow &\xleftrightarrow{\text{attr.}} \downarrow \Rightarrow \text{bond, attraction,} \\ \chi_t: \quad \uparrow &\xleftrightarrow{\text{rep.}} \uparrow \Rightarrow \text{no bond, repulsion.}\end{aligned}$$

The spin state determines the spin and orbital symmetries and accordingly an attractive or repulsive exchange term is generated. The situation for the H + He bond can be discussed analogously. The He atom has two electrons: the potentially attractive exchange in the singlet channel produces a (high-energy) triplet state for the He atom (forbidden in $1s^2$, only allowed in the excited state $1s2p$), the repulsive exchange in the triplet channel is allowed but of no use in the binding,

$$\uparrow \xleftrightarrow{\text{attr.}} \begin{pmatrix} \uparrow \\ \downarrow \end{pmatrix} \text{ singlet exchange results in } \downarrow \xleftrightarrow{\text{rep.}} \begin{pmatrix} \uparrow \\ \uparrow \end{pmatrix} \text{ high energy.}$$

$$\uparrow \xleftrightarrow{\text{rep.}} \begin{pmatrix} \uparrow \\ \downarrow \end{pmatrix} \text{ triplet exchange results in } \uparrow \xleftrightarrow{\text{rep.}} \begin{pmatrix} \uparrow \\ \downarrow \end{pmatrix} \text{ useless.}$$

Together, the system H+He is repulsive and there is no bond. In generalizing these findings, we can make the following statements:

1. The exchange of an electron from atom A with paired electrons from atom B leads to a repulsive interaction.
2. Closed shells are chemically inert¹³ because all electrons are paired.
3. The unpaired electrons in incomplete shells, i.e., the valence electrons, typically s - or p -orbitals, are responsible for chemical bonds.
4. The binding of two atoms via an attractive exchange of two electrons leads to the pairing of these electrons in a singlet state. This state is repulsive with respect to a third electron and the chemical bond saturates. The same applies to Cooper pairs, paired electrons in the superconductor.¹⁴

¹³Inert = no bonds

¹⁴Are there stable bosonic pairs?

5. Incompletely filled d and f shells do not participate in the binding process because they have small and isolated orbitals as sketched in Fig. 15.7. Accordingly, these (compact) inner shells are usually chemically irrelevant; they only give the molecule a (often large) spin and thus large magnetic moments that are technologically useful.

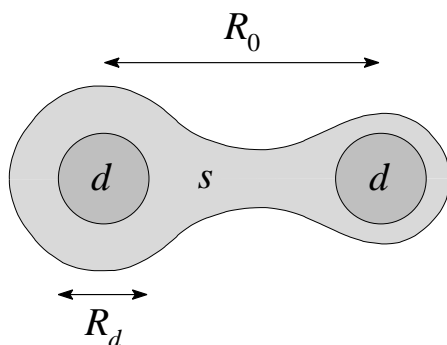


Fig. 15.7: The s orbitals determine the distance $R_0 > R_d$ between the nuclei in the molecule. The compact d orbitals with radius R_d produce magnetic moments.

15.5.1 Bonding types

In this section, we consider different atoms $X = \text{Be}, \text{C}, \text{N}, \text{O}$ and discuss the formation of the diatomic molecule X_2 . We start with

Be, Beryllium, with the configuration $1s^2 2s^2$. The electrons are paired and therefore chemically inert, hence one may expect no binding in a diatomic molecule. This naive expectation is wrong. The transition $2s \rightarrow 2p$ costs only little energy, such that a Be atom can generate two valence electrons by adopting a new $n = 2$ -shell configuration $2s 2p$. These valence electrons then are available for chemical bonds, e.g., in the formation of a Be_2 molecule (the inner $1s^2$ electrons remain chemically inert). A maximum exchange energy is found when the p orbitals (e.g., the p_z orbitals along the z -axis) maximally overlap.¹⁵ The p -orbitals have the form

$$|p_x\rangle = \frac{1}{\sqrt{2}}(Y_{1,1} + Y_{1,-1}) = \sqrt{\frac{3}{4\pi}} \frac{x}{r},$$

$$|p_y\rangle = \frac{1}{\sqrt{2}}(Y_{1,1} - Y_{1,-1}) = \sqrt{\frac{3}{4\pi}} \frac{y}{r},$$

¹⁵We choose, w.l.o.g., $m = 0$ with respect to the bonding axis.

$$|p_z\rangle = Y_{1,0} = \sqrt{\frac{3}{4\pi}} \frac{z}{r}. \quad (15.44)$$

and are shown in Fig. 15.8. In the formation of the Be_2 molecule,

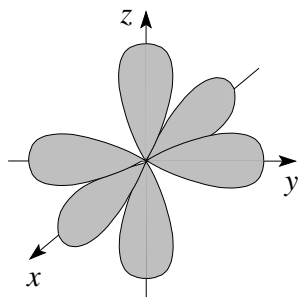


Fig. 15.8: The p -orbitals are aligned along the x , y and z axes.

both orbitals $2s$ and $2p_z$ simultaneously contribute to the (σ -type) binding. Since $2s$ is not directional, the axis is determined by the $2p$ orbital. Note that the Be_2 molecule is a tricky object and there is a vast literature on it; here, we have only discussed a very simple picture.

Next, we take a look at Carbon,

C, Carbon has the configuration $2s^2 2p^2$. Two carbon atoms form two bonds, a σ -bond with $m = 0$ with respect to the z -axis and a π -bond with $m = 1$ with respect to z . The bonds are sketched in 15.9.

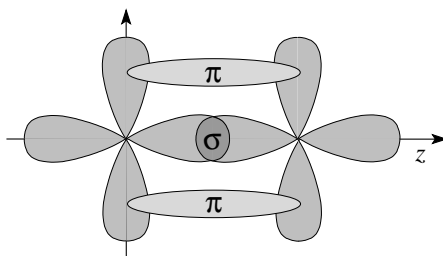


Fig. 15.9: σ - and π -bonds of p -orbitals. The σ -bond is rotationally symmetric with respect to z -axis, i.e., a $m = 0$ configuration ($\sigma \hat{=} s$ because $m = 0$). The π -bond has the $m = 1$ symmetry of a dipole ($\pi \hat{=} p$ because $m = 1$).

N, Nitrogen, with the configuration $2s^2 2p^3$, forms three bonds $\sigma + \pi + \pi$ in the diatomic molecule.

O, Oxygen with the configuration $2s^2 2p^4$ forms three bonds $\sigma + 2 \times \pi$ when forming the O_2 molecule, similar to nitrogen. In addition, there remains one more p -electron per atom. These electrons must be housed

in the $|p_x\rangle$ and $|p_y\rangle$ anti-bonding orbitals (ABO). In order to minimize the (repulsive) overlap in the ABO, the two electrons go into a triplet state and occupy both orbitals $|p_x\rangle$ and $|p_y\rangle$, see Fig. 15.10. The result is that oxygen O_2 is paramagnetic with spin 1.¹⁶

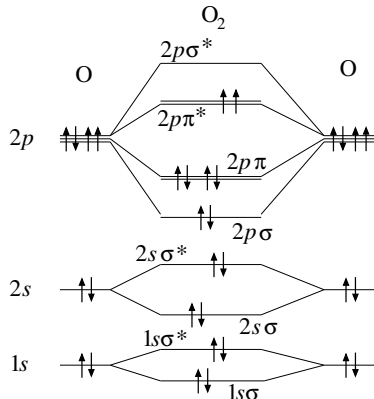


Fig. 15.10: Orbitals of an oxygen molecule with a σ and two π bonds (10 of 12 electrons in the $n = 2$ shell in a $S = 0$ state). The last two electrons in the $2p\pi^*$ level (the $*$ marks an anti-bonding state) arrange themselves in a triplet state (2nd Hund's rule) in order to minimize the Coulomb energy. This occupation results in paramagnetic oxygen molecules with total spin $S = 1$. The orbitals $2s\sigma^*$ and $2p\sigma^*$ result from the hybridized states $2s$ (dominant) plus $2p_z$ and $2p_z$ (dominant) plus $2s$.

15.5.2 Hybridization

The hybridization of orbitals serves to maximize the overlap and optimize the binding energy. We introduce the topic using a simple example, the Li_2 molecule. The energetic proximity of the $2s$ and $2p$ orbitals allows two alternative bonding options, either via an undirected σ -bond based on the $2s$ orbital, see Fig. 15.11(a), or alternatively, via a directed σ -bond with the $2p$ -orbital, see Fig. 15.11(b). But only a combination results in an optimal σ -bond, see Fig. 15.11(c).

¹⁶Otherwise, molecules without f electrons usually have a singlet configuration with $S = 0$.

$2s$	\rightarrow	σ -bond
$2p$	\rightarrow	directed σ -bond

optimization: $|s\rangle + \lambda|p\rangle$.

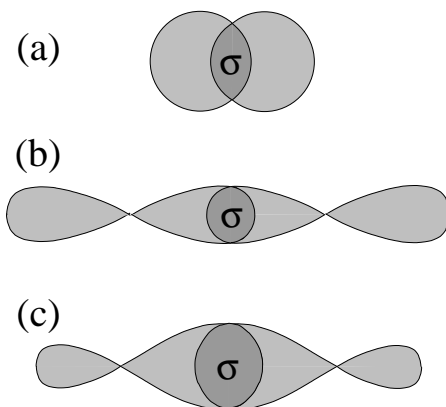


Fig. 15.11: Optimization of binding through hybridization. (a) simple σ -bonding of $2s$ orbitals. (b) Occupation of the $2p$ orbitals (costs energy) in order to achieve a stronger (directional) σ -bond, which should compensate for the energy used. (c) Optimization of (a) and (b) by hybridization.

The hybridization parameter λ results from the following considerations:

- $\lambda = 0$, no excitation energy due to the transfer $s \rightarrow p$, undirected bond.
- $\lambda > 0$, the transfer $s \rightarrow p$ costs energy, but the stronger, directed bond compensates for it.
- The exact calculation gives the value $\lambda_{\text{optimal}} \approx 0.3$ with the weights $|\Psi_s|^2 \sim 0.9$ and $|\Psi_p|^2 \sim 0.1$ for the $2s$ and $2p$ wavefunctions.

Other interesting examples of hybridized orbitals occur in the formation of water and a number of carbon compounds.

H_2O , water. Oxygen O has a $2p^4$ configuration, with orbitals along the axes. Without hybridization, the O – H bonds enclose a 90° angle, see Fig. 15.12.

In the formation of water molecules, the oxygen orbitals hybridize according to

$$|1\rangle = \frac{1}{2}(|s\rangle + |p_x\rangle + |p_y\rangle + |p_z\rangle), \quad |3\rangle = \frac{1}{2}(|s\rangle - |p_x\rangle + |p_y\rangle - |p_z\rangle),$$

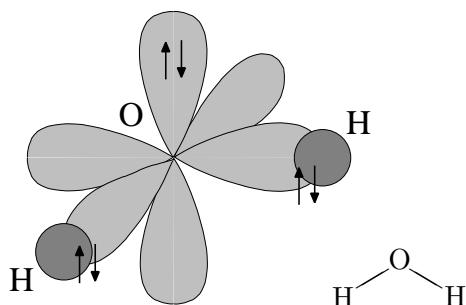


Fig. 15.12: Water H_2O , without hybridization.

$$|2\rangle = \frac{1}{2}(|s\rangle + |p_x\rangle - |p_y\rangle - |p_z\rangle), \quad |4\rangle = \frac{1}{2}(|s\rangle - |p_x\rangle - |p_y\rangle + |p_z\rangle), \quad (15.45)$$

and results in the favorable sp^3 hybridization with the tetrahedron geometry¹⁷, as indicated in Fig. 15.13. Experimentally, the angle

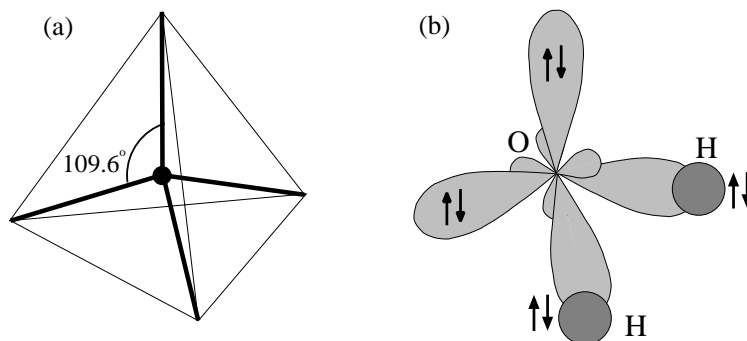


Fig. 15.13: Water H_2O , with sp^3 hybridization. The ideal tetrahedron angle is 109.6° , experimentally it is found to be 105° in water.

between the H-atoms is about 105° ; the hybridization is therefore incomplete.

NH_3 , Ammonia, has structurally the same tetrahedral configuration as H_2O with a 107° angle between the N – H bonds, see Fig. 15.14(a). Ammonia is of outstanding importance for agriculture/food (fertilizers: urea,

¹⁷Note: the $2s$ orbital is lower than $2p$. Since $2s$ is no longer doubly occupied in this configuration, some excitation energy must be paid.

ammonium nitrate, ammonium sulfate, ammonium phosphates), but also explosives, and is produced in the famous Haber-Bosch process (catalytic synthesis of atmospheric nitrogen and hydrogen at ~ 200 bar and $\sim 400^\circ\text{C}$; need for a good catalyst and for hydrogen). At the beginning of the 20th century, the project ‘Bread from Air’ turned away an impending famine in Europe.

A special feature is the freedom of the nitrogen atom to tunnel along q (see Fig. 15.14(a)) through the plane defined by the three H atoms (this is not a rotation but a additional degree of freedom). The potential

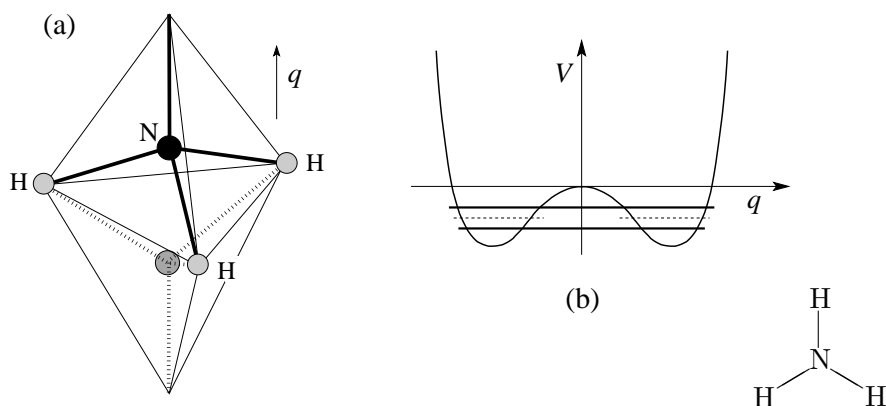


Fig. 15.14: (a) Ammonia molecule with two states. (b) Effective double well potential for the ammonia molecule. Tunnel processes between the two (semi-classical) states (dashed) allow energy to be gained by delocalizing the wave function. The tunnel splitting of the true delocalized eigenstates (solid lines) defines the maser frequency (24 GHz).

landscape for the tunnel process is sketched in Fig. 15.14(b). The two semi-classical states (dashed lines) mix and lead to an energy reduction due to the delocalization of the nitrogen atom; this mixing results in a (lower energy) ‘bonding state’ and an ‘anti-bonding state’ at higher energy (solid lines). The tunnelling gap, which energetically separates the ‘bonding’ and ‘anti-bonding’ states, defines the ‘maser frequency’ (micro-wave amplification through stimulated emission of radiation (about $\nu \approx 24$ GHz or $\lambda \approx 13$ mm), instead of the better known laser frequency in the visible range).

Carbon, C, appears in various forms.

CH_4 , Methane, assumes a perfectly tetrahedral symmetry with angles 109.6° .

C_2H_4 , Ethylene, has a planar structure with $sp^2 + p_z$ hybridization, see Fig. 15.15,

$$\begin{aligned} sp^2 : \quad |1\rangle &= \sqrt{1/3}|s\rangle + \sqrt{2/3}|p_x\rangle, \\ |2\rangle &= \sqrt{1/3}|s\rangle - \sqrt{1/6}|p_x\rangle + \sqrt{1/2}|p_y\rangle, \\ |3\rangle &= \sqrt{1/3}|s\rangle - \sqrt{1/6}|p_x\rangle - \sqrt{1/2}|p_y\rangle, \\ p_z : \quad |4\rangle &= |p_z\rangle. \end{aligned} \quad (15.46)$$

The hybridization allows for a double-bond between the C via a σ - and a π -bond as shown in Fig. 15.15. The π -bond, which is realized by the p_z orbitals, defines the plane for the C – H bonds.

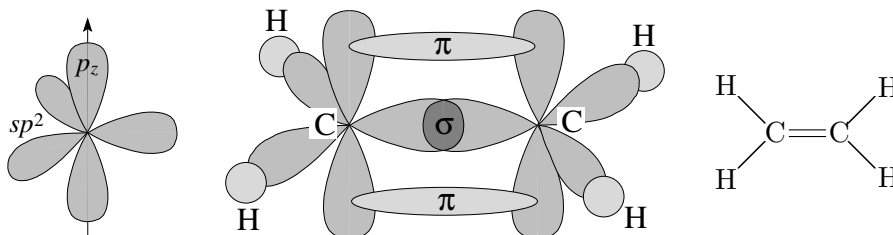


Fig. 15.15: Left, the $sp^2 + p_z$ orbitals of carbon atoms occur in ethylene and in benzene. Middle and right, ethylene C_2H_4 molecule with σ and π double-bonds.

C_2H_2 , Acetylene, has a linear structure $\text{H}-\text{C}\equiv\text{C}-\text{H}$ with a hybridization $sp + p_x + p_y$ for the C atom,

$$\begin{aligned} |1\rangle &= (|s\rangle + |p_z\rangle)/2, & |2\rangle &= (|s\rangle - |p_z\rangle)/2, \\ |3\rangle &= |p_x\rangle, & |4\rangle &= |p_y\rangle. \end{aligned} \quad (15.47)$$

The triple-bond between the carbon atoms involves one σ - and two π -bonds, as sketched in Fig. 15.16.

C_6H_6 , in benzene, as in ethylene, the carbon orbitals adopt an $sp^2 + p_z$ hybridization, see Fig. 15.17. In benzene there is a twofold degeneracy with complementary double-bonds; chemists speak of a mesomerizing double-bond, physicists call the same phenomenon a resonating valence bond (RVB = resonating valence bond): the coherent tunneling

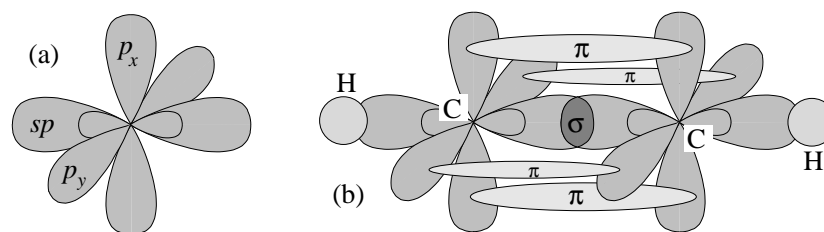


Fig. 15.16: (a) $sp + p_x + p_y$ hybridization of carbon atoms. (b) Acetylene C_2H_2 with a $\sigma\pi\pi$ triple-bond between the C atoms, $\text{H-C}\equiv\text{C-H}$.

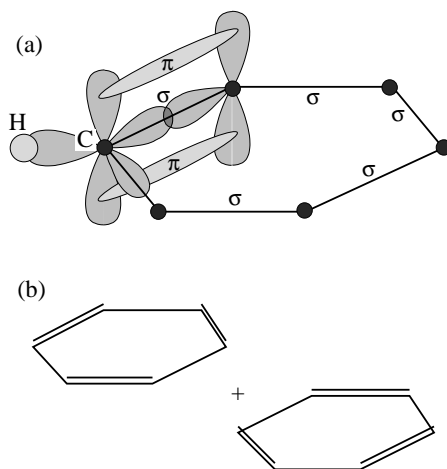


Fig. 15.17: (a) Benzene ring with $sp^2 + p_z$ hybridization of carbon. The double-bond selects a partner, that results in two degenerate states. (b) The tunneling process between the two degenerate states leads to a delocalization of the double-bond and thus to an energy reduction. This is a simple example of a resonant valence bond state (RVB = resonating valence bond state).

between the states allows for the delocalization of the π -electrons and results in an energy reduction. Accordingly, the benzene ring can be modelled as a small metallic ring, which results in a large diamagnetic susceptibility.¹⁸

¹⁸The application of a magnetic field induces constant molecular currents, just like in a superconducting ring. In the benzene ring, this is a single-electron effect, whereas the continuous currents in the superconducting ring correspond to a macroscopic quantum effect.

Chapter 16

Second Quantization

Material for this chapter can be found in the book of Fetter and Walecka.

The second quantization is a formalism that gives us an elegant and compact formulation of many-body problems. The second quantization has nothing to do with some sort of 'super quantum mechanics', in which one would 'quantize a second time'; all quantum mechanics is already in the theory we know, which includes states in a Hilbert space (wave functions) and operators (observables), the Schrödinger equation (for the dynamics) and symmetrization conditions for bosons and fermions. We already know an example of a second quantization procedure: we have found the spectrum of the harmonic oscillator in two alternative ways,

conventional → first quantization

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

**elegant → operator technique
→ second quantization**

$$\begin{aligned} H &= \hbar\omega(a^\dagger a + 1/2), \\ [a, a^\dagger] &= 1; \\ \text{vacuum } |0\rangle, & \quad a|0\rangle = 0, \\ \text{states,} & \quad |n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle, \\ E_n &= \hbar\omega(n + 1/2), \\ \Psi_n(q) &= \langle q|n\rangle. \end{aligned}$$

Instead of the wavefunctions Ψ_n , we have introduced the raising and lowering operators a^\dagger and a and then expressed the entire theory using these

operators. The ‘quanta’, which are raised and lowered the states and their energies that we denoted by a^\dagger and a play the same role as ‘particles’¹ that can be created or destroyed. The general scheme we are going to introduce below follows the scheme

classical mechanics		quantum- mechanics		quantum- mechanics
particles orbits	1. Quant. →	fields (waves) wavefunctions	2. quant. →	quanta, particles, operators a^\dagger, a .

In a more general context, one often already starts with a field (field theory), for example in the description of electromagnetism using the potentials $\varphi(\vec{r}, t)$ and $\vec{A}(\vec{r}, t)$, or in the description of the elasticity of bodies that is expressed by the displacement $\vec{u}(\vec{r}, t)$, or in the case of magnetism where the magnetization density $\vec{m}(\vec{r}, t)$ is the central object of interest. In such a case, there is no ‘second’ quantization, rather the original quantization performed in the second-quantization scheme is the only one,

fields	quantization →	‘particles’,
field amplitudes	→	operators.

The resulting ‘particles’ or quanta are mostly bosons and are called photons (electromagnetism), phonons (elasticity theory), and magnons (magnetism).

The transition to the second quantization is a rather tedious procedure. Only the end result becomes a simple and compact formalism. We first formulate the result and afterwards work out a derivation for bosons. Also, we consider different approaches, for example, via standard wave functions and subsequently via a field theory approach.

16.1 Problem in first-quantized form

We start with the many-particle Hamilton operator

$$H = \sum_i \left(\frac{p_i^2}{2m} + U(\vec{r}_i) \right) + \frac{1}{2} \sum_{i \neq j} V(\vec{r}_i, \vec{r}_j), \quad (16.1)$$

¹In general, the particle number is preserved. Here, the quantum number n of the energy can be changed by a^\dagger and a ; the ‘particles’ corresponding to the excitations are not preserved in number.

that involves of the one-particle problem $H_0 = \sum_i [p_i^2/2m + U(\vec{r}_i)]$ and the interaction $V(\vec{r}_i, \vec{r}_j)$. Let $\{\varphi_\alpha(x), \alpha = \text{set of quantum numbers}, x = \vec{r}, s_z\}$ be a (complete) one-particle basis. Then the product wavefunctions

$$\Psi_{\pi N}^{\vec{\alpha}}(x_1, \dots, x_N) = \varphi_{\alpha_1}(x_1) \cdots \varphi_{\alpha_N}(x_N) \quad (16.2)$$

span the space $\mathcal{H}^{\otimes N}$ of the N -particle wave functions. We define the Fock-space as a direct sum

$$\mathcal{F} = \oplus_N \mathcal{H}^{\otimes N}; \quad (16.3)$$

the Fock-space \mathcal{F} contains spaces to describe zero (the vacuum state), one, two, \dots N particles. If fermionic / bosonic particles are considered, the anti-symmetric / symmetric part of \mathcal{F} is relevant. Wave functions for fermions / bosons are obtained from the product basis states $\Psi_{\pi N}^{\vec{\alpha}}$ through superposition and (anti-) symmetrization. The dynamics of the states are given by the Schrödinger equation

$$i\hbar \partial_t \Psi = H \Psi. \quad (16.4)$$

Within non-relativistic physics, the dynamics for N particles remains within $\mathcal{H}^{\otimes N}$. This is the 'first' quantized form of quantum mechanics.

16.2 Problem in second-quantized form

We formulate the rules of the second-quantization formalism. In section 16.3, we discuss the derivation of the formalism from the first-quantized form.

One-particle basis:

A central element is an orthonormal and complete set of one-particle wave-functions $\varphi_\alpha(x)$, for example the planes waves with spin $\varphi_{\vec{k}, \sigma}(\vec{r}, s) = \exp(i\vec{k} \cdot \vec{r}) \chi_\sigma(s)/\sqrt{V}$, the Bloch functions $\varphi_{n\vec{k}, \sigma}(\vec{r}, s) = \exp(i\vec{k} \cdot \vec{r}) u_n(\vec{r}) \chi_\sigma(s)$, the Coulomb wave functions $\varphi_{Elm\sigma}(\vec{r}, s)$ with the discrete bound states $E = E_n < 0$ and the scattering states in the continuum $E \geq 0$, or the harmonic oscillator functions $\Psi_n(q) = N_n H_n(q) e^{-m\omega q^2/2\hbar}$.

Creation and annihilation operators, vacuum:

Let the wave functions $\varphi_\alpha(x)$ be ordered, $\alpha = 1 < \alpha = 2 < \dots \alpha = k \dots$. We define associated creation and annihilation operators

$$a_\alpha^\dagger \quad \text{and} \quad a_\alpha, \quad (16.5)$$

and the vacuum $|0\rangle$ with the following properties:

- Commutation rules: For the operators A and B , let us define

$$\begin{aligned} [A, B] &= AB - BA \quad \text{the commutator,} \\ \{A, B\} &= AB + BA \quad \text{the anti-commutator.} \end{aligned} \quad (16.6)$$

Then, we require the creation and annihilation operators a_α^\dagger and a_α to satisfy

for bosons:

for fermions:

$$\begin{aligned} [a_\alpha, a_{\alpha'}^\dagger] &= \delta_{\alpha\alpha'}, & \{a_\alpha, a_{\alpha'}^\dagger\} &= \delta_{\alpha\alpha'}, \\ [a_\alpha, a_{\alpha'}] &= 0, & \{a_\alpha, a_{\alpha'}\} &= 0, \\ [a_\alpha^\dagger, a_{\alpha'}^\dagger] &= 0, & \{a_\alpha^\dagger, a_{\alpha'}^\dagger\} &= 0. \end{aligned} \quad (16.7)$$

- $|0\rangle$ is the vacuum, annihilated by a_α 's,

$$a_\alpha|0\rangle = 0. \quad (16.8)$$

- a_α^\dagger and a_α are creation and annihilation operators for particles, quanta, modes, \dots ,

$$\begin{aligned} a_\alpha^\dagger &: \mathcal{H}^{\otimes N} \rightarrow \mathcal{H}^{\otimes(N+1)} \quad \text{creates,} \\ a_\alpha &: \mathcal{H}^{\otimes N} \rightarrow \mathcal{H}^{\otimes(N-1)} \quad \text{annihilates} \end{aligned} \quad (16.9)$$

a particle in the state $\varphi_\alpha(x)$, where

$$\varphi_\alpha(x) = \langle x | a_\alpha^\dagger | 0 \rangle. \quad (16.10)$$

Occupation number basis:

The N -particle wavefunctions for bosons and fermions are given by

$$\begin{aligned} &|n_1, n_2, \dots, n_k, \dots, n_\infty\rangle \\ &\stackrel{\text{bosons}}{=} \frac{(a_\infty^\dagger)^{n_\infty}}{\sqrt{n_\infty!}} \dots \frac{(a_k^\dagger)^{n_k}}{\sqrt{n_k!}} \dots \frac{(a_2^\dagger)^{n_2}}{\sqrt{n_2!}} \frac{(a_1^\dagger)^{n_1}}{\sqrt{n_1!}} |0\rangle, \end{aligned} \quad (16.11)$$

$$\stackrel{\text{fermions}}{=} (a_\infty^\dagger)^{n_\infty} \dots (a_k^\dagger)^{n_k} \dots (a_2^\dagger)^{n_2} (a_1^\dagger)^{n_1} |0\rangle, \quad (16.12)$$

where

$$\sum_{k=1}^{\infty} n_k = N. \quad (16.13)$$

This implies that ‘almost all’ $n_k = 0$. In the state $|n_1, n_2, \dots, n_k, \dots, n_\infty\rangle$, the one-particle state k is occupied n_k times. The order is relevant and is given by the order of the one-particle wave functions $\alpha = 1 < \alpha = 2 < \dots < \alpha = k < \dots < \alpha = \infty$. The states $|n_1, \dots, n_k, \dots\rangle$ are orthonormal. Allowed occupation numbers for bosons are $n_\alpha = 0, 1, 2, \dots \in \mathbb{N}_0$. In contrast, allowed occupation numbers for fermions are only $n_\alpha = 0, 1$, such that the Pauli principle is satisfied, which allows a maximum of one particle per state.

The states $|n_1, n_2, \dots, n_k, \dots\rangle$ correspond to the correctly symmetrized N -particles wavefunctions, i.e., for

bosons:

$$\begin{aligned}\Psi(x_1, \dots, x_N) &= \langle x_1, \dots, x_N | n_1, \dots, n_k, \dots \rangle \\ &= U_s \Psi_{\pi N}^{\vec{\alpha}}(x_1, \dots, x_N),\end{aligned}\quad (16.14)$$

fermions:

$$\begin{aligned}\Psi(x_1, \dots, x_N) &= \langle x_1, \dots, x_N | n_1, \dots, n_k, \dots \rangle \\ &= U_a \Psi_{\pi N}^{\vec{\alpha}}(x_1, \dots, x_N),\end{aligned}\quad (16.15)$$

with the symmetrizing operator U_s and the anti-symmetrizing operator U_a , see (13.36), and the product wave functions (16.2). The vector $\vec{\alpha}$ is given by $\vec{\alpha} = n_1$ times 1, n_2 times 2, \dots . A typical state would look like this:

$$\text{for bosons: } |n_1, \dots, n_k, \dots\rangle \stackrel{N=49}{=} |25, 13, 7, 3, 0, 1, 0 \dots\rangle \quad (16.16)$$

the symmetrized product state

$$\begin{aligned}\Psi &= U_s \Psi_{\pi N}^{\vec{\alpha}} \text{ with} \\ \vec{\alpha} &= (\underbrace{1, 1, 1, \dots, 1}_{25}, \underbrace{2, 2, \dots, 2}_{13}, \dots, \underbrace{6}_1)\end{aligned}$$

$$\text{for fermions: } |n_1, \dots, n_k, \dots\rangle \stackrel{N=7}{=} |1, 1, 1, 1, 0, 1, 1, 0, 0 \dots\rangle$$

the anti-symmetrized product state

$$\begin{aligned}\Psi &= U_a \Psi_{\pi N}^{\vec{\alpha}} \text{ with} \\ \vec{\alpha} &= (1, 2, 3, 4, 6, 7, 8).\end{aligned}\quad (16.17)$$

The tuple $|n_1, n_2, \dots, n_k, \dots, n_\infty\rangle$ is called occupation number basis in the (symmetrica / antisymmetric) Fock space sector $\mathcal{F}_{s/a}$. The basis is orthonormal,

$$\langle n'_1, n'_2, \dots, n'_\infty | n_1, n_2, \dots, n_\infty \rangle = \delta_{n'_1 n_1} \delta_{n'_2 n_2} \dots \delta_{n'_\infty n_\infty}, \quad (16.18)$$

and complete,

$$\sum_{n_1, n_2, \dots, n_\infty} |n_1, n_2, \dots, n_\infty\rangle \langle n_1, n_2, \dots, n_\infty| = \mathbb{1} \in \mathcal{F}. \quad (16.19)$$

From the commutation rules, we can find the action of the operators a_α^\dagger and a_α on the occupation number state $|n_1, n_2, \dots, n_\infty\rangle$,

for bosons:

$$\begin{aligned} a_k |\dots, n_k, \dots\rangle &= \sqrt{n_k} |\dots, n_k - 1, \dots\rangle, \\ a_k^\dagger |\dots, n_k, \dots\rangle &= \sqrt{n_k + 1} |\dots, n_k + 1, \dots\rangle, \\ a_k^\dagger a_k |\dots, n_k, \dots\rangle &= n_k |\dots, n_k, \dots\rangle, \\ n_k &= 0, 1, 2, \dots, \infty, \end{aligned} \quad (16.20)$$

for fermions:

$$\begin{aligned} a_k |\dots, n_k, \dots\rangle &= \delta_{n_k, 1} (-1)^{S_k} |\dots, n_k - 1, \dots\rangle, \\ a_k^\dagger |\dots, n_k, \dots\rangle &= \delta_{n_k, 0} (-1)^{S_k} |\dots, n_k + 1, \dots\rangle, \\ a_k^\dagger a_k |\dots, n_k, \dots\rangle &= n_k |\dots, n_k, \dots\rangle, \\ n_k &= 0, 1; \end{aligned} \quad (16.21)$$

$$\begin{aligned} S_k &= n_\infty + \dots + n_{k+1} \\ &= N - (n_1 + n_2 + \dots + n_k). \end{aligned} \quad (16.22)$$

Operators:

The operators in second-quantized form can be written using the creation and annihilation operators. Let $A_1(x)$ be a one-particle operator, then

$$A_1 = \sum_{ij} \langle i|A_1|j\rangle a_i^\dagger a_j \quad (16.23)$$

with²

$$\langle i|A_1|j\rangle = \int dx \varphi_i^*(x) A_1(x) \varphi_j(x). \quad (16.24)$$

Let $A_2(x, x')$ be a two-particle operator, then

$$A_2 = \frac{1}{2} \sum_{ijkl} \langle ij|A_2|kl\rangle a_i^\dagger a_j^\dagger a_l a_k \quad (16.25)$$

with

$$\langle ij|A_2|kl\rangle = \int dx dx' \varphi_i^*(x) \varphi_j^*(x') A_2(x, x') \varphi_k(x) \varphi_l(x'). \quad (16.26)$$

²We define $\int dx = \sum_s \int d^3r$.

The order of the operators is relevant! This manifests itself in the signs for fermions; furthermore, the correct sequence prevents the appearance of terms that describe a self-interaction, e.g., when considering the expression $a_i^\dagger a_j a_l^\dagger a_k$.

Finally, let A_ν be a ν -particle operator, then

$$A_\nu = \frac{1}{\nu!} \sum_{\substack{i_1, \dots, i_\nu \\ k_1, \dots, k_\nu}} \langle i_1, \dots, i_\nu | A_\nu | k_1, \dots, k_\nu \rangle a_{i_1}^\dagger \cdots a_{i_\nu}^\dagger a_{k_\nu} \cdots a_{k_1} \quad (16.27)$$

where

$$\begin{aligned} \langle i_1, \dots, i_\nu | A_\nu | k_1, \dots, k_\nu \rangle &= \int dx_1 \cdots dx_\nu \\ &\varphi_{i_1}^*(x_1) \cdots \varphi_{i_\nu}^*(x_\nu) A_\nu(x_1, \dots, x_\nu) \varphi_{k_1}(x_1) \varphi_{k_\nu}(x_\nu). \end{aligned} \quad (16.28)$$

A typical Hamiltonian with particle-particle interaction then has the form

$$H = \sum_{ij} \langle i | p^2/2m + U(\vec{r}) | j \rangle a_i^\dagger a_j + \frac{1}{2} \sum_{ijkl} \langle ij | V(\vec{r}, \vec{r}') | kl \rangle a_i^\dagger a_j^\dagger a_l a_k. \quad (16.29)$$

As usual, it generates the dynamics of the states. The particle number operator

$$N = \sum_i a_i^\dagger a_i \quad (16.30)$$

counts the particles in the system.

The above elements provide a **complete formalism** that we can use to do calculations in a very compact way. Specifically, in order to calculate expectation values, we need

- States: defined by (16.10), (16.11), and (16.12),
- Operators: defined by (16.23)–(16.28),
- Rules for calculation: defined by the commutation rules (16.7) and the definition of the vacuum (16.8).

16.3 Second quantization: derivation*

We derive the second-quantization formalism from the first-quantized form of quantum mechanics. This derivation serves the continuity of the reasoning, but will no longer play a role later on, e.g., when investigating or

solving specific problems. A short sketch of the proof goes like this: Starting from the Schrödinger equation in first-quantized form, we express the many-body wave function Ψ in the product base. The Schrödinger equation is then written as a system of equations for the coefficients $c(\vec{\alpha}, t)$ in this product representation. These coefficients have to satisfy the symmetrization conditions for Ψ and are therefore partly redundant. This redundancy is eliminated by defining new coefficients $\bar{c}(\vec{n}, t)$ which only depend on the occupation numbers \vec{n} ; the subsequent normalization leads to the final coefficients $f(\vec{n}, t) \propto \bar{c}(\vec{n}, t)$. At this stage, the wave function Ψ is represented by the symmetrized occupation number basis with coefficients $f(\vec{n}, t)$. The Schrödinger equation is again equivalent to a system of equations for the coefficients $f(\vec{n}, t)$. On the right-hand side of these coupled equations there appear coefficients $f(\vec{n})$ with occupation numbers which differ from those in $\partial_t f(\vec{n}, t)$ on the left-hand side by $0, \pm 1, \pm 2$ (± 2 from interaction terms). By defining raising and lower operators which produce these changes in the occupation numbers \vec{n} of $f(\vec{n})$, the complicated system of equations can be written in a simple form, from which the second-quantized form of the Schrödinger equation follows.

We start by formulating a typical problem in the first-quantized form.

- Hamiltonian:

$$H = \sum_{k=1}^N H_0(x_k) + \frac{1}{2} \sum_{k \neq l} V(x_k, x_l), \quad H_0 = \frac{p^2}{2m} + U. \quad (16.31)$$

- One-particle functions: $\varphi_\alpha(x)$, a complete orthonormal system in \mathcal{H} , ordered.
- N -particle solution:

$$i\hbar \partial_t \Psi(x_1, \dots, x_N, t) = H \Psi(x_1, \dots, x_N, t). \quad (16.32)$$

For bosons, $\Psi \in \mathcal{H}_S$ is symmetric, while for fermions, $\Psi \in \mathcal{H}_A$ anti-symmetric.

We express the solution in the product basis of $\mathcal{H}^{\otimes N}$,

$$\Psi(x_1, \dots, x_N, t) = \sum_{\alpha'_1, \dots, \alpha'_N} c(\alpha'_1, \dots, \alpha'_N, t) \varphi_{\alpha'_1}(x_1) \cdots \varphi_{\alpha'_N}(x_N). \quad (16.33)$$

We insert this Ansatz into the Schrödinger equation (16.32), multiply with $\varphi_{\alpha_1}^*(x_1) \cdots \varphi_{\alpha_N}^*(x_N)$ and integrate; the k -th term in the one-particle Hamiltonian H_0 then reduces to the following expression

$$\int dx_1 \cdots dx_N \underbrace{\varphi_{\alpha_1}^*(x_1) \cdots \varphi_{\alpha_k}^*(x_k)}_{\delta_{\alpha_1, \alpha'_1}} \cdots \underbrace{\varphi_{\alpha_N}^*(x_N)}_{\delta_{\alpha_N, \alpha'_N}} H_0(x_k) \varphi_{\alpha'_1}(x_1) \cdots \varphi_{\alpha'_N}(x_N). \quad (16.34)$$

The operator $H_0(x_k)$ only affects the k -th factor $\varphi_{\alpha_k}(x_k)$ and gives the matrix element

$$\langle \alpha_k | H_0 | \alpha'_k \rangle = \int dx_k \varphi_{\alpha_k}^*(x_k) H_0(x_k) \varphi_{\alpha'_k}(x_k). \quad (16.35)$$

Next, we look at the interaction term in the Hamiltonian, specifically the term $V(x_k, x_l)$ in the sum. Here, as well, the integration produces only Kronecker-deltas $\delta_{\alpha_i, \alpha'_i}$ for all $i \neq k, l$,

$$\int dx_1 \cdots dx_N \underbrace{\varphi_{\alpha_1}^*(x_1) \cdots \varphi_{\alpha_k}^*(x_k)}_{\delta_{\alpha_1, \alpha'_1}} \cdots \underbrace{\varphi_{\alpha_l}^*(x_l) \cdots \varphi_{\alpha_N}^*(x_N)}_{\delta_{\alpha_N, \alpha'_N}} V(x_k, x_l) \varphi_{\alpha'_1}(x_1) \cdots \varphi_{\alpha'_N}(x_N) \quad (16.36)$$

and the projection provides the matrix element

$$\langle \alpha_k \alpha_l | V | \alpha'_k \alpha'_l \rangle = \int dx dx' \varphi_{\alpha_k}^*(x) \varphi_{\alpha_l}^*(x') V(x, x') \varphi_{\alpha'_k}(x) \varphi_{\alpha'_l}(x'). \quad (16.37)$$

The above steps show, how the matrix elements for the operators appear in the formalism.

Next, we use the above results and express the Schrödinger equation through the coefficients $c(\alpha_1, \dots, \alpha_N, t)$ in the expansion of Ψ ,

$$\begin{aligned} i\hbar \partial_t c(\alpha_1, \dots, \alpha_N, t) &= \sum_{k=1}^N \sum_{\alpha'_k} \langle \alpha_k | H_0 | \alpha'_k \rangle c(\alpha_1, \dots, \alpha'_k, \dots, \alpha_N, t) \\ &+ \frac{1}{2} \sum_{k \neq l}^N \sum_{\alpha'_k, \alpha'_l} \langle \alpha_k \alpha_l | V | \alpha'_k \alpha'_l \rangle c(\alpha_1, \dots, \alpha'_k, \dots, \alpha'_l, \dots, \alpha_N, t), \end{aligned} \quad (16.38)$$

to obtain a coupled system of equations for the amplitudes $c(\alpha_1, \dots, \alpha_N, t)$. The symmetry of the wave function requires the following properties under

exchange $x_k \leftrightarrow x_l$

$$\Psi(x_1, \dots, x_k, \dots, x_l, \dots) = \begin{cases} +\Psi(x_1, \dots, x_l, \dots, x_k, \dots), & \text{for bosons,} \\ -\Psi(x_1, \dots, x_l, \dots, x_k, \dots), & \text{for fermions.} \end{cases}$$

Applied to the coefficients c , this gives the following symmetry requirement under the exchange of two quantum numbers α_k and α_l ,

$$\begin{aligned} c(\alpha_1, \dots, \alpha_k, \dots, \alpha_l, \dots, \alpha_N, t) \\ = \pm c(\alpha_1, \dots, \alpha_l, \dots, \alpha_k, \dots, \alpha_N, t), \end{aligned} \quad (16.39)$$

where the signs are $+$ for bosons and $-$ for fermions. The dynamics and symmetry of the wave function are now encoded in the coefficients $c(\alpha_1, \dots, \alpha_N, t)$.

Next, we introduce a correctly symmetrized orthonormal basis; for bosons these functions are symmetric linear combinations of the product functions $\Psi_{\pi N}^{\alpha} = \varphi_{\alpha_1}(x_1) \cdots \varphi_{\alpha_N}(x_N)$

$$\Phi_{n_1, \dots, n_{\infty}}(x_1, \dots, x_N) \equiv \left(\frac{n_1! n_2! \cdots n_{\infty}!}{N!} \right)^{\frac{1}{2}} \sum_{\alpha_1, \dots, \alpha_N} \varphi_{\alpha_1}(x_1) \cdots \varphi_{\alpha_N}(x_N), \quad (16.40)$$

where the sum $\sum_{\alpha_1, \dots, \alpha_N}$ is restricted to terms with fixed occupation numbers n_1, \dots, n_{∞} . The new basis functions $\Phi_{n_1, \dots, n_{\infty}}$ are orthonormal,

$$\begin{aligned} \int dx_1 \cdots dx_N \Phi_{n'_1, \dots, n'_{\infty}}^*(x_1, \dots, x_N) \Phi_{n_1, \dots, n_{\infty}}(x_1, \dots, x_N) \\ = \delta_{n_1, n'_1} \cdots \delta_{n_{\infty}, n'_{\infty}}, \end{aligned} \quad (16.41)$$

and complete by construction; the bosonic symmetry (statistics) is taken into account,

$$\begin{aligned} \Phi_{n_1, \dots, n_{\infty}}(x_1, \dots, x_k, \dots, x_l, \dots, x_N) \\ = \Phi_{n_1, \dots, n_{\infty}}(x_1, \dots, x_l, \dots, x_k, \dots, x_N). \end{aligned} \quad (16.42)$$

The states $\Phi_{n_1, \dots, n_{\infty}}$ correspond to the occupation number states $|n_1, \dots, n_{\infty}\rangle$ introduced on page 113,

$$\Phi_{n_1, \dots, n_{\infty}}(x_1, \dots, x_N) = \langle x_1, \dots, x_N | n_1, \dots, n_{\infty} \rangle. \quad (16.43)$$

In order to implement the basis change to the Φ 's, we regroup the terms in the wave function Ψ , (16.33): Consider a configuration in Ψ with amplitude

$c(\alpha_1, \dots, \alpha_N, t)$. The quantum numbers $\alpha_1, \dots, \alpha_N$ come from an ordered set of states $\{1, 2, \dots, \infty\}$. Consequently, to each α_k there corresponds a number from the set $\{1, 2, \dots, \infty\}$; e.g., $\alpha_{23} = 5$ means that particle 23, located at x_{23} , is in the state $\varphi_5(x_{23})$. Several bosonic particles may be in the same state. A bosonic amplitude $c(\alpha_1, \dots, \alpha_N, t)$ can therefore contain several identical quantum numbers $\alpha_i = \alpha_j = \dots$; a typical example is the coefficient

$$c(1, 2, 25, 12, 3, 1, 3, 2, 25, 12, 12, 2, 1, \dots; t) \quad (16.44)$$

which can be written as

$$\begin{aligned} & c(1, 2, 25, 12, 3, 1, 3, 2, 25, 12, 12, 2, 1, \dots; t) \\ &= c(\underbrace{1, 1, 1, \dots}_{n_1 \text{ times}}, \underbrace{2, 2, 2, \dots}_{n_2 \text{ times}}, \dots, \underbrace{12, 12, 12, \dots}_{n_{12} \text{ times}}; t) \\ &\equiv \bar{c}(n_1, n_2, \dots, n_{12}, \dots, n_\infty; t), \end{aligned} \quad (16.45)$$

using the bosonic symmetry. All amplitudes $c(\alpha_1, \dots, \alpha_N)$ with identical quantum numbers $\alpha_1, \dots, \alpha_N$ have the same value \bar{c} for symmetry reasons; it is therefore sufficient to know the new amplitudes $\bar{c}(n_1, \dots, n_\infty, t)$ with the occupation numbers n_1, \dots, n_∞ . With the transition to the new coefficients \bar{c} , we need a new normalisation:

$$\begin{aligned} 1 &= \int \prod_k dx_k |\Psi|^2 = \sum_{\alpha_1, \dots, \alpha_N} |c(\alpha_1, \dots, \alpha_N; t)|^2 \\ &= \sum_{n_1, \dots, n_\infty} |\bar{c}(n_1, \dots, n_\infty, t)|^2 \sum_{\alpha_1, \dots, \alpha_N} 1, \end{aligned} \quad (16.46)$$

where the sum $\sum_{\alpha_1, \dots, \alpha_N}$ is again limited to terms with fixed occupation numbers n_1, \dots, n_∞ . This sum counts the number of possibilities A_{n_1, \dots, n_∞} to create sequences of length $N = \sum_i n_i$ with n_i numbers i , therefore

$$\sum_{\alpha_1, \dots, \alpha_N} 1 = A_{n_1, \dots, n_\infty} = \frac{N!}{n_1! n_2! \dots n_\infty!}. \quad (16.47)$$

We introduce new, normalised coefficients f instead of \bar{c} ,

$$\begin{aligned} f(n_1, n_2, \dots, n_\infty; t) &= \left(\frac{N!}{n_1! \dots n_\infty!} \right)^{\frac{1}{2}} \bar{c}(n_1, \dots, n_\infty; t), \\ \sum_{n_1, \dots, n_\infty} f(n_1, \dots, n_\infty; t) &= 1. \end{aligned} \quad (16.48)$$

We can now express Ψ in the symmetric basis $\Phi_{n_1, \dots, n_\infty}$, see (16.40),

$$\begin{aligned} \Psi(x_1, \dots, x_N; t) &= \sum_{\alpha_1, \dots, \alpha_N} c(\alpha_1, \dots, \alpha_N; t) \varphi_{\alpha_1}(x_1) \cdots \varphi_{\alpha_N}(x_N) \\ &= \sum_{n_1, \dots, n_\infty} f(n_1, \dots, n_\infty; t) \\ &\quad \times \underbrace{\left(\frac{n_1! \cdots n_\infty!}{N!} \right)^{\frac{1}{2}} \sum_{\alpha_1, \dots, \alpha_N} \varphi_{\alpha_1}(x_1) \cdots \varphi_{\alpha_N}(x_N)}_{\Phi_{n_1, \dots, n_\infty}(x_1, \dots, x_N)}. \end{aligned} \quad (16.49)$$

Next, we write the Hamiltonian in the new basis $\Phi_{n_1, \dots, n_\infty}(x_1, \dots, x_N)$ and start with the one-particle term H_0 in (16.38),

$$\begin{aligned} &\sum_{k=1}^N \sum_{\alpha'_k} \langle \alpha_k | H_0 | \alpha'_k \rangle c(\alpha_1, \dots, \alpha'_k, \dots; t) \\ &= \sum_{k=1}^N \sum_{\alpha'_k} \langle \alpha_k | H_0 | \alpha'_k \rangle \bar{c}(n_1, \dots, n_{\alpha_k} - 1, \dots, n_{\alpha'_k} + 1, \dots, n_\infty; t) \end{aligned}$$

Obviously, the α_k state on the left-hand side of (16.38) is replaced by the α'_k state, which we take into account in the \bar{c} coefficients by decreasing n_{α_k} by one and increasing $n_{\alpha'_k}$ by one. In the sum \sum_k , α_k occurs just n_k times and contributes the same amount to the sum each time, so the sum over particles can be replaced by a sum over states, $\sum_k = \sum_\alpha n_\alpha$, and we obtain

$$\begin{aligned} &\sum_{k=1}^N \sum_{\alpha'_k} \langle \alpha_k | H_0 | \alpha'_k \rangle c(\alpha_1, \dots, \alpha'_k, \dots; t) \\ &= \sum_{\alpha} \sum_{\alpha'} \langle \alpha | H_0 | \alpha' \rangle n_\alpha \bar{c}(n_1, \dots, n_\alpha - 1, \dots, n_{\alpha'} + 1, \dots, n_\infty; t) \\ &\downarrow \quad \alpha \rightarrow i, \quad \alpha' \rightarrow j \text{ (ordered quantum numbers)} \\ &= \sum_i \langle i | H_0 | i \rangle n_i \left(\frac{n_1! \cdots n_i! \cdots}{N!} \right)^{\frac{1}{2}} f(n_1, n_2, \dots, n_i, \dots, n_\infty; t) \\ &\quad + \sum_{i \neq j} \langle i | H_0 | j \rangle n_i \left(\frac{\cdots (n_i - 1)! \cdots (n_j + 1)! \cdots}{N!} \right)^{\frac{1}{2}} \\ &\quad \times f(n_1, n_2, \dots, (n_i - 1), \dots, (n_j + 1), \dots, n_\infty; t). \end{aligned} \quad (16.50)$$

The factors n_i , together with the modified normalization factors associated with the f -coefficients, will finally produce the multiplicative factors $\sqrt{n+1}$ and \sqrt{n} of a^\dagger and a in (16.21).

We treat the interaction term V in a similar way,

$$\begin{aligned} & \frac{1}{2} \sum_{k \neq l} \sum_{\alpha'_k \alpha'_l} \langle \alpha_k \alpha_l | V | \alpha_{k'} \alpha_{l'} \rangle c(\alpha_1, \dots, \alpha'_k, \dots, \alpha'_l, \dots; t) \\ &= \frac{1}{2} \sum_{k \neq l} \sum_{\alpha'_k \alpha'_l} \langle \alpha_k \alpha_l | V | \alpha_{k'} \alpha_{l'} \rangle \\ & \quad \bar{c}(n_1, \dots, n_{\alpha_k} - 1, \dots, n_{\alpha'_k} + 1, \dots, n_{\alpha_l} - 1, \dots, n_{\alpha'_l} + 1, \dots, n_\infty; t). \end{aligned} \quad (16.51)$$

Again, we replace the $\sum_{\text{particles}}$ with a \sum_{states} : In the sum $\sum_{k \neq l}$, the quantum number α_k (α_l) occurs just n_{α_k} (n_{α_l}) times and we obtain $n_{\alpha_k} n_{\alpha_l}$ identical terms if $\alpha_k \neq \alpha_l$; if $\alpha_k = \alpha_l$, one term less occurs in the second sum (over k) (since $k \neq l$) and we arrive at $n_{\alpha_l}(n_{\alpha_l} - 1)$ terms. This results in (16.51)

$$\begin{aligned} & \frac{1}{2} \sum_{k \neq l} \sum_{\alpha'_k \alpha'_l} \langle \alpha_k \alpha_l | V | \alpha_{k'} \alpha_{l'} \rangle c(\alpha_1, \dots, \alpha'_k, \dots, \alpha'_l, \dots; t) \\ &= \frac{1}{2} \sum_{\alpha_k \alpha_l} \sum_{\alpha'_k \alpha'_l} n_{\alpha_k} (n_{\alpha_l} - \delta_{\alpha_k \alpha_l}) \langle \alpha_k \alpha_l | V | \alpha_{k'} \alpha_{l'} \rangle \bar{c}(n_1, \dots, n_\infty; t) \\ & \quad \text{mit: } \begin{array}{ccc} \alpha_k & \rightarrow & i, \quad \alpha_{k'} \rightarrow k \\ \alpha_l & \rightarrow & j, \quad \alpha_{l'} \rightarrow l \end{array} \\ &= \frac{1}{2} \sum_{ijkl} n_i (n_j - \delta_{ij}) \langle ij | V | kl \rangle \\ & \quad \times \bar{c}(n_1, \dots, n_i - 1, \dots, n_j - 1, \dots, n_k + 1, \dots, n_l + 1, \dots; t) \\ &= \sum_{i \neq j \neq k \neq l} \frac{1}{2} \langle ij | V | kl \rangle n_i n_j \\ & \quad \times f(\dots, n_i - 1, \dots, n_j - 1, \dots, n_k + 1, \dots, n_l + 1, \dots; t) \\ & \quad \times \left(\frac{\dots (n_i - 1)! \dots (n_j - 1)! \dots (n_k + 1)! \dots (n_l + 1)! \dots}{N!} \right)^{\frac{1}{2}} \\ &+ \sum_{i=j \neq k \neq l} \frac{1}{2} \langle ii | V | kl \rangle n_i (n_i - 1) \\ & \quad \times f(\dots, n_i - 2, \dots, n_k + 1, \dots, n_l + 1, \dots; t) \\ & \quad \times \left(\frac{\dots (n_i - 2)! \dots (n_k + 1)! \dots (n_l + 1)! \dots}{N!} \right)^{\frac{1}{2}} \end{aligned}$$

$$+ \sum_{i=j=k \neq l} \dots + \dots, \quad (16.52)$$

where all sums with combinations of '=' and '≠' relations between the indices i, j, k, l have to be taken into account. After multiplication with $[N!/(n_1! \dots n_\infty!)]^{1/2}$, the combination of the Schrödinger equation (16.32) with the 1-particle-term (16.50) and the interaction term (16.52) results in the system of equations for the coefficients f ,

$$\begin{aligned} i\hbar \partial_t f(n_1, \dots, n_\infty; t) = & \sum_i \langle i | H_0 | i \rangle n_i f(n_1, \dots, n_i, \dots, n_\infty; t) \\ & + \sum_{i \neq j} \langle i | H_0 | j \rangle \sqrt{n_i} \sqrt{n_j + 1} f(\dots, n_i - 1, \dots, n_j + 1, \dots; t) \\ & + \frac{1}{2} \sum_{i \neq j \neq k \neq l} \langle ij | V | kl \rangle \sqrt{n_i} \sqrt{n_j} \sqrt{n_k + 1} \sqrt{n_l + 1} \\ & \quad \times f(\dots, n_i - 1, \dots, n_j - 1, \dots, n_k + 1, \dots, n_l + 1, \dots; t) \\ & + \frac{1}{2} \sum_{i \neq k \neq l} \langle ii | V | kl \rangle \sqrt{n_i} \sqrt{n_i - 1} \sqrt{n_k + 1} \sqrt{n_l + 1} \\ & \quad \times f(\dots, n_i - 2, \dots, n_k + 1, \dots, n_l + 1, \dots; t) \\ & + \sum_{\text{other combinations } \neq \& = \text{ in } i, j, k, l} \end{aligned} \quad (16.53)$$

where the further combinations include a term with three ≠, a term with three =, six terms with two ≠ and one = and six terms with one ≠ and two =. According to the construction, this system of equations describes the Schrödinger equation in the occupation number basis. The introduction of the operators a_α^\dagger and a_α as defined in (16.5), with the commutation relations (16.7) and the definition of the occupation number basis (16.11) provide us with a simplified form of (16.53): We start with the wave function in the occupation number basis

$$|\Psi(t)\rangle = \sum_{n_1, \dots, n_\infty} f(n_1, \dots, n_\infty; t) \underbrace{|n_1, \dots, n_\infty\rangle}_{\Phi_{n_1, \dots, n_\infty}}. \quad (16.54)$$

With the dynamics (16.53) of $f(n_1, \dots, n_\infty; t)$, $|\Psi(t)\rangle$ satisfies the Schrödinger equation. We can now write the result (16.53) term for term in second-quantized form. We demonstrate the procedure on the $i \neq j$ one-particle term and replace

$$n_i \rightarrow n'_i = n_i - 1,$$

$$\begin{aligned} n_j &\rightarrow n'_j = n_j + 1, \\ n_k &\rightarrow n'_k = n_k \quad \text{else.} \end{aligned} \quad (16.55)$$

This re-definition of the occupation numbers leaves the particle number unchanged, $\sum_i n_i = \sum_i n'_i = N$. Furthermore, $n'_i \in \{-1, 0, 1, \dots, \infty\}$ and $n'_j \in \{1, 2, \dots, \infty\}$, but for $n'_i = -1$, $n_i = 0$ and for $n'_j = 0$, $n_j + 1 = 0$. If we now choose all $n'_i \in (0, \dots, \infty)$, the factors $\sqrt{n_i} \sqrt{n_j + 1} = \sqrt{n'_i + 1} \sqrt{n'_j}$ disappear in the additional or missing terms, i.e. for $n'_i = -1$, $n'_j = 0$. The $i \neq j$ one-particle term in (16.53) can thus be rewritten in the form

$$\begin{aligned} &\sum_{n'_1, \dots, n'_\infty} \sum_{i \neq j} \langle i | H_0 | j \rangle f(n'_1, n'_2, \dots, n'_i, \dots, n'_j, \dots, n'_\infty; t) \\ &\quad \times \underbrace{\sqrt{n'_i + 1} \sqrt{n'_j} |n'_1, \dots, n'_i + 1, \dots, n'_j - 1, \dots, n'_\infty\rangle}_{\text{see (16.21)} \quad a_i^\dagger a_j |n'_1, \dots, n'_i, \dots, n'_j, \dots, n'_\infty\rangle} \\ &= \sum_{i \neq j} \langle i | H_0 | j \rangle a_i^\dagger a_j |\Psi(t)\rangle. \end{aligned} \quad (16.56)$$

The same analysis can be carried out for the other terms and we arrive at the Schrödinger equation in second-quantized form

$$i\hbar \partial_t |\Psi(t)\rangle = H |\Psi(t)\rangle \quad (16.57)$$

with

$$H = \sum_{ij} \langle i | H_0 | j \rangle a_i^\dagger a_j + \frac{1}{2} \sum_{ijkl} \langle ij | V | kl \rangle a_i^\dagger a_j^\dagger a_l a_k. \quad (16.58)$$

We can see that the second quantized form (16.54) is exactly equivalent to the first quantized form (16.49). Once we have made the somewhat tedious transition to the second-quantized form, we have obtained a compact and clear formalism.

16.3.1 Fermions

When describing fermions, analogous steps have to be followed, with the complication arising from the additional signs appearing under an exchange of two particles. On the other hand, the procedure is simplified by the fact that the occupation numbers n_i can only assume values 0 or 1. The orthonormalized anti-symmetrized basis states $\Phi_{n_1, \dots, n_\infty}(x_1, \dots, x_N) = \langle x_1, \dots, x_N | n_1, \dots, n_\infty \rangle$ correspond precisely to the Slater determinants.

16.4 Change of basis

Let the one-particle states $\{\varphi_\alpha\}$ and $\{\chi_\nu\}$ define two complete orthonormal systems (CONS). The change of basis

$$\begin{aligned} |\varphi_\alpha\rangle &= \sum_\nu \langle \chi_\nu | \varphi_\alpha \rangle |\chi_\nu\rangle, \\ |\chi_\nu\rangle &= \sum_\alpha \langle \varphi_\alpha | \chi_\nu \rangle |\varphi_\alpha\rangle \end{aligned} \quad (16.59)$$

corresponds to a transformation of the one-particle operators a_α^\dagger and a_α and b_ν^\dagger and b_ν : We start from the relationship (16.10) between the creation operators and the one-particle basis functions,

$$\varphi_\alpha(x) = \langle x | \varphi_\alpha \rangle = \langle x | a_\alpha^\dagger | 0 \rangle = \sum_\nu \langle \chi_\nu | \varphi_\alpha \rangle \underbrace{\langle x | b_\nu^\dagger | 0 \rangle}_{\chi_\nu(x)}, \quad (16.60)$$

and arrive at the transformation between operators in the form,

$$\begin{aligned} a_\alpha^\dagger &= \sum_\nu \langle \chi_\nu | \varphi_\alpha \rangle b_\nu^\dagger, & b_\nu^\dagger &= \sum_\alpha \langle \varphi_\alpha | \chi_\nu \rangle a_\alpha^\dagger, \\ a_\alpha &= \sum_\nu \langle \varphi_\alpha | \chi_\nu \rangle b_\nu, & b_\nu &= \sum_\alpha \langle \chi_\nu | \varphi_\alpha \rangle a_\alpha. \end{aligned} \quad (16.61)$$

As an example, consider fermions on a lattice. The operators c_i^\dagger and c_i create and annihilate a fermion on site i at the location \vec{r}_i , $i = 1, \dots, I$, respectively. On the other hand, the operators $c_{\vec{k}}^\dagger$ and $c_{\vec{k}}$ create / annihilate plane waves with wave vector \vec{k} . Then

$$\begin{aligned} c_{\vec{k}}^\dagger &= \sum_i \langle i | \vec{k} \rangle c_i^\dagger = \sum_i \frac{e^{i\vec{k} \cdot \vec{r}_i}}{\sqrt{I}} c_i^\dagger, \\ c_i^\dagger &= \sum_{\vec{k}} \frac{e^{-i\vec{k} \cdot \vec{r}_i}}{\sqrt{I}} c_{\vec{k}}^\dagger. \end{aligned} \quad (16.62)$$

Next, we are interested in the two-particle state $|\vec{k} \vec{q}\rangle$, especially, its representation in the position basis on the lattice, $\langle i j | \vec{k} \vec{q} \rangle$. We define

$$\begin{aligned} |i j\rangle &= c_i^\dagger c_j^\dagger |0\rangle & |\vec{k} \vec{q}\rangle &= c_{\vec{k}}^\dagger c_{\vec{q}}^\dagger |0\rangle \\ \Rightarrow \langle i j | \vec{k} \vec{q} \rangle &= \langle 0 | c_j c_i c_{\vec{k}}^\dagger c_{\vec{q}}^\dagger | 0 \rangle \end{aligned} \quad (16.63)$$

$$\begin{aligned}
&= \frac{1}{I} \sum_{lm} \langle 0 | c_j c_i e^{i\vec{k} \cdot \vec{r}_l} c_l^\dagger e^{i\vec{q} \cdot \vec{r}_m} c_m^\dagger | 0 \rangle \\
&= \frac{1}{I} \sum_{lm} e^{i\vec{k} \cdot \vec{r}_l} e^{i\vec{q} \cdot \vec{r}_m} \langle 0 | c_j c_i c_l^\dagger c_m^\dagger | 0 \rangle. \tag{16.64}
\end{aligned}$$

The calculation of the matrix elements $\langle 0 | c_j c_i c_l^\dagger c_m^\dagger | 0 \rangle$ illustrates how calculations are performed in the second-quantized formalism by using the commutation relations (16.7) and the definition of the vacuum (16.8),

$$\begin{aligned}
\langle 0 | c_j c_i c_l^\dagger c_m^\dagger | 0 \rangle &= \langle 0 | c_j (\{c_i, c_l^\dagger\} - c_l^\dagger c_i) c_m^\dagger | 0 \rangle \\
&= \langle 0 | c_j (\delta_{il} - c_l^\dagger c_i) c_m^\dagger | 0 \rangle \\
&= \delta_{il} \underbrace{\langle 0 | c_j c_m^\dagger | 0 \rangle}_{\delta_{jm}} - \underbrace{\langle 0 | c_j c_l^\dagger}_{\delta_{jl}} \underbrace{c_i c_m^\dagger}_{\delta_{im}} | 0 \rangle \\
&= \delta_{il} \delta_{jm} - \langle 0 | c_j c_l^\dagger (\delta_{im} - c_m^\dagger c_i) | 0 \rangle \\
&= \delta_{il} \delta_{jm} - \delta_{im} \underbrace{\langle 0 | \{c_j c_l^\dagger\}}_{\delta_{jl}} - c_l^\dagger \underbrace{c_j}_{0} | 0 \rangle \\
&= \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}. \tag{16.65}
\end{aligned}$$

We insert this result into (16.64) and obtain the two-particle matrix element

$$\begin{aligned}
\langle i j | \vec{k} \vec{q} \rangle &= \sum_{lm} (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) \frac{e^{i\vec{k} \cdot \vec{r}_l}}{\sqrt{I}} \frac{e^{i\vec{q} \cdot \vec{r}_m}}{\sqrt{I}} \\
&= \begin{vmatrix} \frac{e^{i\vec{k} \cdot \vec{r}_i}}{\sqrt{I}} & \frac{e^{i\vec{k} \cdot \vec{r}_j}}{\sqrt{I}} \\ \frac{e^{i\vec{q} \cdot \vec{r}_i}}{\sqrt{I}} & \frac{e^{i\vec{q} \cdot \vec{r}_j}}{\sqrt{I}} \end{vmatrix} = \text{Slater det} \{ \varphi_{\vec{k}}, \varphi_{\vec{q}} \}, \tag{16.66}
\end{aligned}$$

which is nothing but the Slater determinant (up to normalization) of the plane-wave product function $\varphi_{\vec{k}}(\vec{r}_i) \varphi_{\vec{q}}(\vec{r}_j)$. Accordingly, $\langle i j | \vec{k} \vec{q} \rangle$ is already properly antisymmetrized.

The result (16.65) is reminiscent of the Wick (or Hartree-Fock) decomposition of the 4-point function (2-particle correlator) into a sum of products of 2-point functions (1-particle correlators)

$$\langle 0 | c_j c_i c_l^\dagger c_m^\dagger | 0 \rangle = \sum_{\text{pairings}} \prod (\pm) \langle 0 | c c^\dagger | 0 \rangle \langle 0 | c c^\dagger | 0 \rangle. \tag{16.67}$$

In order to list the sum over pairings, we mark the pairs to be contracted with a line,

$$\langle 0|c_j c_i c_l^\dagger c_m^\dagger|0\rangle \rightarrow \overline{\langle 0|c_j c_i c_l^\dagger c_m^\dagger|0\rangle} + \overline{\langle 0|c_j c_i c_l^\dagger c_m^\dagger|0\rangle}. \quad (16.68)$$

In a next step, we bring the operators to their proper position via exchanges; every exchange of two fermions contributes a factor (-1) (we shift j by two positions in the first term ($\rightarrow 2$ exchanges) and i by one position in the second ($\rightarrow 1$ exchange)),

$$\langle 0|c_j c_i c_l^\dagger c_m^\dagger|0\rangle \rightarrow \overline{\langle 0|c_i c_l^\dagger c_j c_m^\dagger|0\rangle} - \overline{\langle 0|c_j c_l^\dagger c_i c_m^\dagger|0\rangle}. \quad (16.69)$$

In a last step, we factor the expressions and arrive at the sum over pairings,

$$\langle 0|c_j c_i c_l^\dagger c_m^\dagger|0\rangle = \langle 0|c_i c_l^\dagger|0\rangle \langle 0|c_j c_m^\dagger|0\rangle - \langle 0|c_j c_l^\dagger|0\rangle \langle 0|c_i c_m^\dagger|0\rangle. \quad (16.70)$$

The evaluation of the 2-point function gives $\langle 0|c_m c_n^\dagger|0\rangle = \delta_{mn}$ and we obtain the result (16.65). We will come back to the Wick-decomposition in the discussion of the pair-correlator $\langle \Phi_0|\Psi_s^\dagger(\vec{r})\Psi_{s'}^\dagger(\vec{r}')\Psi_{s'}(\vec{r}')\Psi_s(\vec{r})|\Phi_0\rangle$ in (17.18).

16.5 Field operators

The basis transformation to the position base $\{\psi_{\vec{r}}\}$ generates the field operator $\Psi(\vec{r})$. We start from the definition of position basis functions and use the completeness of the basis $\{\varphi_\alpha\}$,

$$\delta(\vec{r} - \vec{r}') = \langle \vec{r}'|\psi_{\vec{r}}\rangle = \sum_\alpha \underbrace{\langle \varphi_\alpha|\psi_{\vec{r}}\rangle}_{\varphi_\alpha^*(\vec{r})} \underbrace{\langle \vec{r}'|\varphi_\alpha\rangle}_{\varphi_\alpha(\vec{r}')}. \quad (16.71)$$

In analogy to $|\varphi_\alpha\rangle = a_\alpha^\dagger|0\rangle$, we define $|\psi_{\vec{r}}\rangle \equiv \Psi^\dagger(\vec{r})|0\rangle$, i.e., the field operator $\Psi^\dagger(\vec{r})$ creates one particle at location \vec{r} . The relation (16.71) can then be written as

$$\Psi^\dagger(\vec{r})|0\rangle = \sum_\alpha \varphi_\alpha^*(\vec{r}) a_\alpha^\dagger|0\rangle \quad (16.72)$$

and we obtain the transformation between conventional mode and field operators (see (16.61))

$$\begin{aligned} \Psi^\dagger(\vec{r}) &= \sum_\alpha \varphi_\alpha^*(\vec{r}) a_\alpha^\dagger, \\ \Psi(\vec{r}) &= \sum_\alpha \varphi_\alpha(\vec{r}) a_\alpha. \end{aligned} \quad (16.73)$$

As an example, we generate the field operators from the plane waves created via $a_{\vec{k}}^\dagger$,

$$\langle \vec{r} | a_{\vec{k}}^\dagger | 0 \rangle = \frac{e^{i\vec{k} \cdot \vec{r}}}{\sqrt{V}}. \quad (16.74)$$

Then the field operator

$$\Psi^\dagger(\vec{r}) = \sum_{\vec{k}} \frac{e^{-i\vec{k} \cdot \vec{r}}}{\sqrt{V}} a_{\vec{k}}^\dagger \quad (16.75)$$

creates a particle at \vec{r} , because

$$\langle \vec{r}' | \Psi^\dagger(\vec{r}) | 0 \rangle = \langle \vec{r}' | \sum_{\vec{k}} \frac{e^{-i\vec{k} \cdot \vec{r}}}{\sqrt{V}} a_{\vec{k}}^\dagger | 0 \rangle = \sum_{\vec{k}} \frac{e^{-i\vec{k} \cdot (\vec{r} - \vec{r}')}}{V} = \delta(\vec{r} - \vec{r}'). \quad (16.76)$$

The generalization to particles with spin is trivial: let $a_{\alpha\sigma}^\dagger$ generate a state with spin σ in mode α described by the wave function

$$\langle \vec{r}, s | a_{\alpha\sigma}^\dagger | 0 \rangle = \varphi_\alpha(\vec{r}) \chi_\sigma(s). \quad (16.77)$$

Then the field operator

$$\Psi_s^\dagger(\vec{r}) = \sum_{\alpha, \sigma} \varphi_\alpha^*(\vec{r}) \chi_\sigma(s) a_{\alpha\sigma}^\dagger \quad (16.78)$$

creates a particle with spin s at \vec{r} . Alternatively, one can also write

$$\Psi_s^\dagger(\vec{r}) = \sum_{\alpha} \varphi_{\alpha s}^*(\vec{r}) a_{\alpha s}^\dagger \quad (16.79)$$

by defining $\varphi_{\alpha s}(\vec{r})$ for fixed spin s (instead of the spinor $\varphi_\alpha^*(\vec{r}) \chi_\sigma(s)$).

16.5.1 Commutation Rules

The commutation rules for the field operators follow from those of the operators $a_{\alpha s}^\dagger$ and $a_{\alpha s}$ (with s the spin quantum number),

$$\begin{aligned} \Psi_s(\vec{r}) \Psi_{s'}(\vec{r}') \pm \Psi_{s'}(\vec{r}') \Psi_s(\vec{r}) &= \sum_{\alpha\alpha'} \varphi_{\alpha s}(\vec{r}) \varphi_{\alpha' s'}(\vec{r}') (a_{\alpha s} a_{\alpha' s'} \pm a_{\alpha' s'} a_{\alpha s}) \\ &= 0, \end{aligned}$$

$$\begin{aligned}
\Psi_s^\dagger(\vec{r})\Psi_{s'}^\dagger(\vec{r}') \pm \Psi_{s'}^\dagger(\vec{r}')\Psi_s^\dagger(\vec{r}) &= 0, \\
\Psi_s(\vec{r})\Psi_{s'}^\dagger(\vec{r}') \pm \Psi_{s'}^\dagger(\vec{r}')\Psi_s(\vec{r}) &= \sum_{\alpha\alpha'} \varphi_{\alpha s}(\vec{r})\varphi_{\alpha' s'}^*(\vec{r}') \underbrace{(a_{\alpha s}a_{\alpha' s'}^\dagger \pm a_{\alpha' s'}^\dagger a_{\alpha s})}_{\delta_{\alpha\alpha'}\delta_{ss'}} \\
&= \sum_{\alpha} \underbrace{\varphi_{\alpha s}(\vec{r})\varphi_{\alpha s}^*(\vec{r}')}_{\delta(\vec{r}-\vec{r}')} \delta_{ss'} \\
&= \delta(\vec{r}-\vec{r}')\delta_{ss'}.
\end{aligned} \tag{16.80}$$

For bosons

For fermions

$$\begin{aligned}
[\Psi_s(\vec{r}), \Psi_{s'}(\vec{r}')] &= 0, & \{\Psi_s(\vec{r}), \Psi_{s'}(\vec{r}')\} &= 0, \\
[\Psi_s^\dagger(\vec{r}), \Psi_{s'}^\dagger(\vec{r}')] &= 0, & \{\Psi_s^\dagger(\vec{r}), \Psi_{s'}^\dagger(\vec{r}')\} &= 0, \\
[\Psi_s(\vec{r}), \Psi_{s'}^\dagger(\vec{r}')] &= \delta(\vec{r}-\vec{r}')\delta_{ss'}. & \{\Psi_s(\vec{r}), \Psi_{s'}^\dagger(\vec{r}')\} &= \delta(\vec{r}-\vec{r}')\delta_{ss'}.
\end{aligned} \tag{16.81}$$

16.5.2 Many-particle position basis

We define the states³

$$|\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N\rangle = \frac{1}{\sqrt{N!}} \Psi^\dagger(\vec{r}_N) \Psi^\dagger(\vec{r}_{N-1}) \cdots \Psi^\dagger(\vec{r}_1) |0\rangle \tag{16.82}$$

with the following properties:

1. $|\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N\rangle$ has the correct symmetry under exchange,

$$|\vec{r}_1, \dots, \vec{r}_l, \dots, \vec{r}_k, \dots\rangle = \pm |\vec{r}_1, \dots, \vec{r}_k, \dots, \vec{r}_l, \dots\rangle, \tag{16.83}$$

where ‘+’ applies to bosons and ‘−’ applies to fermions.

2. $\Psi^\dagger(\vec{r})$ creates a particle with correct symmetry

$$\Psi^\dagger(\vec{r}) |\vec{r}_1, \dots, \vec{r}_N\rangle = \sqrt{N+1} |\vec{r}_1, \dots, \vec{r}_N, \vec{r}\rangle. \tag{16.84}$$

3. $\Psi(\vec{r})$ annihilates a particle \vec{r} ; each of the N particles is destroyed (with correct symmetry), such that the resulting $N-1$ particle state has the correct⁴ symmetry,

$$\Psi(\vec{r}) |\vec{r}_1, \dots, \vec{r}_N\rangle$$

³We suppress the spin quantum number.

⁴Symmetric for bosons and antisymmetric for fermions.

$$\begin{aligned}
&= \frac{1}{\sqrt{N!}} \underbrace{\Psi(\vec{r})\Psi^\dagger(\vec{r}_N)}_{\pm\Psi^\dagger(\vec{r}_N)\Psi(\vec{r})+\delta(\vec{r}-\vec{r}_N)} \cdots \Psi^\dagger(\vec{r}_1) |0\rangle \\
&= \frac{1}{\sqrt{N!}} [\delta(\vec{r}-\vec{r}_N) \pm \Psi^\dagger(\vec{r}_N)\Psi(\vec{r})] \Psi^\dagger(\vec{r}_{N-1}) \cdots \Psi^\dagger(\vec{r}_1) |0\rangle \\
&= \frac{1}{\sqrt{N!}} [\delta(\vec{r}-\vec{r}_N) |\vec{r}_1, \dots, \vec{r}_{N-1}\rangle \pm \delta(\vec{r}-\vec{r}_{N-1}) |\vec{r}_1, \dots, \vec{r}_{N-2}, \vec{r}_N\rangle \\
&\quad + \cdots + (\pm 1)^{N-1} \delta(\vec{r}-\vec{r}_1) |\vec{r}_2, \dots, \vec{r}_N\rangle], \tag{16.85}
\end{aligned}$$

where we commuted the annihilation operator $\Psi(\vec{r})$ through every term up until $\Psi(\vec{r})|0\rangle = 0$. The particle at \vec{r} can only be removed if one of the N coordinates $\vec{r}_1, \dots, \vec{r}_N$ is equal to \vec{r} .

4. $\Psi^\dagger(\vec{r}) |\vec{r}_1, \dots, \vec{r}_N\rangle$ creates, $\Psi(\vec{r}) |\vec{r}_1, \dots, \vec{r}_N\rangle$ annihilates a particle.

$\langle \vec{r}_1, \dots, \vec{r}_N | \Psi^\dagger(\vec{r})$ annihilates, $\langle \vec{r}_1, \dots, \vec{r}_N | \Psi(\vec{r})$ creates a particle.

Furthermore

$$\begin{aligned}
\langle \vec{r}_1, \dots, \vec{r}_N | &= \langle 0 | \Psi(\vec{r}_1) \cdots \Psi(\vec{r}_N) \frac{1}{\sqrt{N!}}, \\
|\vec{r}_1, \dots, \vec{r}_N\rangle &= \frac{1}{\sqrt{N!}} \Psi^\dagger(\vec{r}_N) \cdots \Psi^\dagger(\vec{r}_1) |0\rangle. \tag{16.86}
\end{aligned}$$

5. The states $|\vec{r}_1, \dots, \vec{r}_N\rangle$ are orthonormal,

$$\langle \vec{r}'_1, \dots, \vec{r}'_{N'} | \vec{r}_1, \dots, \vec{r}_N \rangle = \frac{\langle 0 | \Psi(\vec{r}'_1) \cdots \Psi(\vec{r}'_{N'}) \Psi^\dagger(\vec{r}_N) \cdots \Psi^\dagger(\vec{r}_1) |0\rangle}{\sqrt{N'}! \sqrt{N!}},$$

by commuting all destruction operators Ψ up until $\Psi(\vec{r}_l)|0\rangle = 0$ the vacuum, one finds

$$\begin{aligned}
\langle \vec{r}'_1, \dots, \vec{r}'_{N'} | \vec{r}_1, \dots, \vec{r}_N \rangle &= \frac{\delta_{NN'}}{N!} \sum_{\pi \in S_n} (\pm 1)^\pi \delta(\vec{r}_1 - \vec{r}'_{\pi(1)}) \\
&\quad \times \delta(\vec{r}_2 - \vec{r}'_{\pi(2)}) \cdots \delta(\vec{r}_N - \vec{r}'_{\pi(N)}). \tag{16.87}
\end{aligned}$$

6. $|\vec{r}_1, \dots, \vec{r}_N\rangle$ is a basis,

$$|\Phi\rangle = \int d^3r_1 \cdots d^3r_N \varphi(\vec{r}_1, \dots, \vec{r}_N) |\vec{r}_1, \dots, \vec{r}_N\rangle; \tag{16.88}$$

$|\Phi\rangle$ has the correct symmetry for arbitrary amplitude functions $\varphi(\vec{r}_1, \dots, \vec{r}_N)$, even those which are not (anti-) symmetrized,

$$\begin{aligned} \langle \vec{r}'_1, \dots, \vec{r}'_N | \Phi \rangle &= \int d^3 r_1 \cdots d^3 r_N \varphi(\vec{r}_1, \dots, \vec{r}_N) \langle \vec{r}'_1, \dots, \vec{r}'_N | \vec{r}_1, \dots, \vec{r}_N \rangle \\ &= \frac{1}{N!} \sum_{\pi} (\pm)^{\pi} \varphi(\vec{r}'_{\pi(1)}, \dots, \vec{r}'_{\pi(N)}) = \varphi(\vec{r}'_1, \dots, \vec{r}'_N), \end{aligned} \quad (16.89)$$

where the last equation applies if φ is already (anti-) symmetrized. If φ is also normalized, such that $\int d^3 r_1 \cdots d^3 r_N |\varphi(\vec{r}_1, \dots, \vec{r}_N)|^2 = 1$, then $|\Phi\rangle$ is also normalized, $\langle \Phi | \Phi \rangle = 1$.

It holds that

$$\mathbb{1}_{\mathcal{F}} = \sum_N \mathbb{1}_N = \sum_N \int d^3 r_1 \cdots d^3 r_N |\vec{r}_1, \dots, \vec{r}_N\rangle \langle \vec{r}_1, \dots, \vec{r}_N| \quad (16.90)$$

the identity in Fock space \mathcal{F} . Note that $\mathbb{1}_0 = |0\rangle\langle 0|$ must be taken into account in the sum.

16.5.3 Operators expressed by field operators

1-particle operators

$$\begin{aligned} A_1 &= \sum_{ij} \langle i | A_1 | j \rangle a_i^{\dagger} a_j \\ &= \sum_{ij} \int d^3 r \varphi_i^*(\vec{r}) A_1(\vec{r}) \varphi_j(\vec{r}) a_i^{\dagger} a_j \\ &= \int d^3 r \Psi^{\dagger}(\vec{r}) A_1(\vec{r}) \Psi(\vec{r}). \end{aligned} \quad (16.91)$$

It is interesting to compare this second-quantized form with the first-quantized form of expectation values in the state $\varphi(\vec{r})$,

$$\langle A_1 \rangle_{\varphi} = \int d^3 r \varphi^*(\vec{r}) A_1 \varphi(\vec{r}). \quad (16.92)$$

A simple recipe for obtaining the operator expressed through field operators then is to take the expectation value $\langle A_1 \rangle_{\varphi}$ and replace the wave function $\varphi(\vec{r})$ by the field operator $\Psi(\vec{r})$. This idea also works for 2-particle operators, etc. Here are a few recurrent examples:

density operator: The expectation value of the density

$$\begin{aligned}\langle \rho(\vec{r}) \rangle_\varphi &= \int d^3r' \varphi^*(\vec{r}') \delta(\vec{r} - \vec{r}') \varphi(\vec{r}') \\ &= \varphi^*(\vec{r}) \varphi(\vec{r}),\end{aligned}\quad (16.93)$$

provides us with the density operator in second-quantized form

$$\rho(\vec{r}) = \Psi^\dagger(\vec{r}) \Psi(\vec{r}). \quad (16.94)$$

The particle number operator has the form

$$N = \int d^3r \Psi^\dagger(\vec{r}) \Psi(\vec{r}). \quad (16.95)$$

current operator: The expression for the current density

$$\langle \vec{j}(\vec{r}) \rangle = \frac{\hbar}{2mi} \left[\varphi^*(\vec{r}) \vec{\nabla} \varphi(\vec{r}) - \varphi(\vec{r}) \nabla \varphi^*(\vec{r}) \right] \quad (16.96)$$

becomes the current-density operator

$$\vec{j}(\vec{r}) = \frac{\hbar}{2mi} \left[\Psi^\dagger(\vec{r}) \vec{\nabla} \Psi(\vec{r}) - (\vec{\nabla} \Psi^\dagger(\vec{r})) \Psi(\vec{r}) \right]. \quad (16.97)$$

Note the order of the operators: First $\Psi(\vec{r})$ has to remove a particle from the state $|\Phi\rangle$, then the operator acts at this location, and finally the particle is added back to the system by $\Psi^\dagger(\vec{r})$. Of course, the current operator (with its derivatives) can also be derived from the usual second-quantized form with operators a_i and a_i^\dagger ,

$$\begin{aligned}\vec{j}(\vec{r}) &= \sum_{ij} \langle i | \vec{j}(\vec{r}) | j \rangle a_i^\dagger a_j \\ &= \sum_{ij} \int d^3r' \varphi_i^*(\vec{r}') \frac{1}{2m} [\vec{p}' \delta(\vec{r}' - \vec{r}) + \delta(\vec{r}' - \vec{r}) \vec{p}'] \varphi_j(\vec{r}') a_i^\dagger a_j \\ &= \frac{\hbar}{2mi} \sum_{ij} \int d^3r' [-\vec{\nabla} \varphi_i^*(\vec{r}') a_i^\dagger \delta(\vec{r}' - \vec{r}) \varphi_j(\vec{r}') a_j \\ &\quad + \varphi_i^*(\vec{r}') a_i^\dagger \delta(\vec{r}' - \vec{r}) \vec{\nabla} \varphi_j(\vec{r}') a_j] \\ &= \frac{\hbar}{2mi} \int d^3r' [-\vec{\nabla} \Psi^\dagger(\vec{r}') \Psi(\vec{r}') + \Psi^\dagger(\vec{r}') \vec{\nabla} \Psi(\vec{r}')] \delta(\vec{r}' - \vec{r}) \\ &= \frac{\hbar}{2mi} [\Psi^\dagger(\vec{r}) \vec{\nabla} \Psi(\vec{r}) - (\vec{\nabla} \Psi^\dagger(\vec{r})) \Psi(\vec{r})].\end{aligned}\quad (16.98)$$

kinetic energy: The expectation value

$$\langle T \rangle_\varphi = \int d^3r \varphi^*(\vec{r}) \frac{-\hbar^2 \nabla^2}{2m} \varphi(\vec{r}) = \frac{\hbar^2}{2m} \int d^3r \vec{\nabla} \varphi^*(\vec{r}) \vec{\nabla} \varphi(\vec{r})$$

leads to the operator of the kinetic energy

$$T = \frac{\hbar^2}{2m} \int d^3r \vec{\nabla} \Psi^\dagger(\vec{r}) \vec{\nabla} \Psi(\vec{r}). \quad (16.99)$$

spin density: The spin density for spin 1/2 particles

$$\langle \vec{s}(\vec{r}) \rangle_{\varphi\chi} = \frac{\hbar}{2} \sum_{s s'} \varphi^*(\vec{r}) \chi^*(s) \vec{\sigma}_{s s'} \varphi(\vec{r}) \chi(s') \quad (16.100)$$

defines the spin-density operator

$$\vec{s}(\vec{r}) = \frac{\hbar}{2} \sum_{s s'} \Psi_s^\dagger(\vec{r}) \vec{\sigma}_{s s'} \Psi_{s'}(\vec{r}). \quad (16.101)$$

Note that the kinetic energy and the density and current density operators expressed in Fourier space have the following form,

$$\begin{aligned} T &= \sum_{\vec{k}} \frac{\hbar^2 k^2}{2m} a_{\vec{k}}^\dagger a_{\vec{k}}, \\ \rho(\vec{q}) &= \sum_{\vec{k}} a_{\vec{k}+\vec{q}}^\dagger a_{\vec{k}}, \\ \vec{j}(\vec{q}) &= \frac{\hbar}{m} \sum_{\vec{k}} (\vec{k} + \vec{q}/2) a_{\vec{k}+\vec{q}}^\dagger a_{\vec{k}}. \end{aligned} \quad (16.102)$$

2-particle operators

$$\begin{aligned} A_2 &= \frac{1}{2} \sum_{ijkl} \langle ij | A_2 | kl \rangle a_i^\dagger a_j^\dagger a_l a_k \\ &= \frac{1}{2} \int d^3r d^3r' \Psi^\dagger(\vec{r}) \Psi^\dagger(\vec{r}') A_2(\vec{r}, \vec{r}') \Psi(\vec{r}') \Psi(\vec{r}). \end{aligned} \quad (16.103)$$

Note the order of the operators,

$$\Psi^\dagger(\vec{r}) \Psi^\dagger(\vec{r}') \Psi(\vec{r}') \Psi(\vec{r}). \quad (16.104)$$

Hamiltonian: The typical Hamiltonian of interacting particles with spin has the form (the interaction is assumed to be spin independent)

$$H = \frac{\hbar^2}{2m} \sum_s \int d^3r \vec{\nabla} \Psi_s^\dagger(\vec{r}) \cdot \vec{\nabla} \Psi_s(\vec{r}) + \frac{1}{2} \sum_{s,s'} \int d^3r d^3r' \Psi_s^\dagger(\vec{r}) \Psi_{s'}^\dagger(\vec{r}') V(\vec{r} - \vec{r}') \Psi_{s'}(\vec{r}') \Psi_s(\vec{r}). \quad (16.105)$$

For an electron gas, $s = \pm 1/2$, $m = m_e$, $V = e^2/r$, and $\Psi_s^\dagger(\vec{r})$, $\Psi_s(\vec{r})$ are fermionic field operators. In Fourier space, this operator has the form ($\Psi_s(\vec{r}) \rightarrow a_{\vec{k}s}$ and $V_{\vec{q}} = 4\pi e^2/q^2$)

$$H = \sum_{\vec{k},s} \frac{\hbar^2 k^2}{2m} a_{\vec{k}s}^\dagger a_{\vec{k}s} + \frac{1}{2} \sum_{\vec{k},\vec{k}',\vec{q},s,s'} V_{\vec{q}} a_{\vec{k}+\vec{q}s}^\dagger a_{\vec{k}'-\vec{q}s'}^\dagger a_{\vec{k}'s'} a_{\vec{k}s}. \quad (16.106)$$

16.6 Equation of motion

As is the case in the first-quantized form of quantum mechanics, we can also choose between Schrödinger, Heisenberg, and interaction pictures in the second-quantized formulation. In the Heisenberg picture, the operators (and thus also the field operators $\Psi_s^\dagger(\vec{r})$ and $\Psi_s(\vec{r})$) depend on time,

$$A_H(t) = U^\dagger(t) A(t) U(t), \quad (16.107)$$

with (T the time ordering operator)

$$U(t) = T \exp \left(-\frac{i}{\hbar} \int_0^t H(t') dt' \right). \quad (16.108)$$

$$i\hbar \partial_t U = H(t) U. \quad (16.109)$$

For a time independent Hamiltonian H ,

$$U(t) = e^{-iHt/\hbar}, \quad H_H(t) = H. \quad (16.110)$$

The equation of motion for the operator $A_H(t)$ is

$$i\hbar \frac{dA_H(t)}{dt} = [A_H(t), H_H(t)] + i\hbar \partial_t A_H(t)$$

$$\begin{aligned} \partial_t H=0 \quad [A_H(t), H] + i\hbar \partial_t A_H(t) \\ \partial_t A=0 \quad [A_H(t), H], \end{aligned} \quad (16.111)$$

$$\frac{dH_H(t)}{dt} = \partial_t H_H(t). \quad (16.112)$$

The time derivative $\partial_t A_H(t)$ is defined as $\partial_t A_H(t) = U^\dagger \partial_t A(t) U$. Note that $[H_H(t), H_H(t)] = 0$ but in general $[H_H(t), H_H(t')]\big|_{t \neq t'} \neq 0$.

We calculate the time derivative of the field operator $d\Psi_H(\vec{r}, t)/dt$ for a Hamiltonian operator H of the form

$$\begin{aligned} H = \int d^3r \Psi^\dagger(\vec{r}) \frac{-\hbar^2 \nabla^2}{2m} \Psi(\vec{r}) \\ + \frac{1}{2} \int d^3r d^3r' \Psi^\dagger(\vec{r}) \Psi^\dagger(\vec{r}') V(\vec{r} - \vec{r}') \Psi(\vec{r}') \Psi(\vec{r}) \end{aligned} \quad (16.113)$$

with $\partial_t H = 0$, $\Psi = \text{F, B-Fields}$, and we suppress the spin quantum numbers. We have to calculate, using (16.111),

$$\begin{aligned} i\hbar \frac{d\Psi_H(\vec{r}, t)}{dt} &= [\Psi_H(\vec{r}, t), H] \\ &\stackrel{(16.110)}{=} e^{iHt/\hbar} [\Psi(\vec{r}), H] e^{-iHt/\hbar}, \end{aligned} \quad (16.114)$$

$$\begin{aligned} [\Psi(\vec{r}), H] &= \int d^3r' \left[\Psi(\vec{r}), \Psi^\dagger(\vec{r}') \frac{-\hbar^2 \nabla'^2}{2m} \Psi(\vec{r}') \right] \\ &\quad + \frac{1}{2} \int d^3r' d^3r'' V(\vec{r}' - \vec{r}'') \\ &\quad \times [\Psi(\vec{r}), \Psi^\dagger(\vec{r}') \Psi^\dagger(\vec{r}'') \Psi(\vec{r}'') \Psi(\vec{r}')] . \end{aligned} \quad (16.115)$$

To calculate the commutators $[A, BC]$, we use the relationships⁵

$$[A, BC] = \begin{cases} [A, B]C - B[C, A], \\ \{A, B\}C - B\{C, A\}. \end{cases} \quad (16.116)$$

For fermions, one finds

$$\begin{aligned} [\Psi(\vec{r}), \Psi^\dagger(\vec{r}') \Psi(\vec{r}')] &= \underbrace{\{\Psi(\vec{r}), \Psi^\dagger(\vec{r}')\}}_{\delta(\vec{r}-\vec{r}')} \Psi(\vec{r}') - \Psi^\dagger(\vec{r}') \underbrace{\{\Psi(\vec{r}), \Psi(\vec{r}')\}}_0 \\ [\Psi(\vec{r}), \Psi^\dagger(\vec{r}') \Psi^\dagger(\vec{r}'') \Psi(\vec{r}'') \Psi(\vec{r}')] &= \underbrace{\{\Psi(\vec{r}), \Psi^\dagger(\vec{r}')\}}_{\delta(\vec{r}-\vec{r}')} \Psi^\dagger(\vec{r}'') \Psi(\vec{r}'') \Psi(\vec{r}') \end{aligned}$$

⁵Use $[A, BC] = ABC - BCA = ABC - BAC + BAC - BCA$.

$$\begin{aligned}
& -\Psi^\dagger(\vec{r}') \underbrace{\{\Psi(\vec{r}), \Psi^\dagger(\vec{r}'')\}}_{\delta(\vec{r}-\vec{r}'')} \Psi(\vec{r}'') \Psi(\vec{r}') \\
& \underbrace{\{\Psi(\vec{r}), \Psi^\dagger(\vec{r}'')\}}_{\delta(\vec{r}-\vec{r}'')} \Psi(\vec{r}'') \Psi(\vec{r}') \\
& = \delta(\vec{r}-\vec{r}') \Psi^\dagger(\vec{r}'') \Psi(\vec{r}'') \Psi(\vec{r}') - \delta(\vec{r}-\vec{r}'') \Psi^\dagger(\vec{r}') \Psi(\vec{r}'') \Psi(\vec{r}'). \quad (16.117)
\end{aligned}$$

The combination of (16.115) – (16.117) gives

$$\begin{aligned}
[\Psi(\vec{r}), H] &= \frac{-\hbar^2}{2m} \vec{\nabla}^2 \Psi(\vec{r}) + \frac{1}{2} \int d^3 r' \int d^3 r'' V(\vec{r}' - \vec{r}'') \\
&\times [\delta(\vec{r} - \vec{r}') \Psi^\dagger(\vec{r}'') \Psi(\vec{r}'') \Psi(\vec{r}') - \delta(\vec{r} - \vec{r}'') \Psi^\dagger(\vec{r}') \Psi(\vec{r}'') \Psi(\vec{r}')]. \quad (16.118)
\end{aligned}$$

Exchanging $\Psi(\vec{r}') \Psi(\vec{r}'') = -\Psi(\vec{r}'') \Psi(\vec{r}')$ in the last term and using (16.114) provides us with the result

$$\begin{aligned}
i\hbar \frac{d\Psi_H(\vec{r}, t)}{dt} &= -\frac{\hbar^2}{2m} \nabla^2 \Psi_H(\vec{r}, t) \\
&+ \left(\int d^3 r' \Psi_H^\dagger(\vec{r}', t) V(\vec{r} - \vec{r}') \Psi_H(\vec{r}', t) \right) \Psi_H(\vec{r}, t). \quad (16.119)
\end{aligned}$$

The result for bosons is identical. Note that⁶ the calculation of the commutators at the equal times $t = t'$ gives the results

$$\begin{aligned}
[\Psi_H(\vec{r}, t), \Psi_H^\dagger(\vec{r}', t)]_\pm &= \delta(\vec{r} - \vec{r}'), \\
[\Psi_H(\vec{r}, t), \Psi_H(\vec{r}', t)]_\pm &= 0, \\
[\Psi_H^\dagger(\vec{r}, t), \Psi_H^\dagger(\vec{r}', t)]_\pm &= 0, \quad (16.120)
\end{aligned}$$

but that the same problem for different times t and t' is a difficult many-particle problem,

$$\begin{aligned}
[\Psi_H(\vec{r}, t), \Psi_H^\dagger(\vec{r}', t')]_\pm &= \text{many-particle problem}, \\
[\Psi_H(\vec{r}, t), \Psi_H(\vec{r}', t')]_\pm &= \text{many-particle problem}, \\
[\Psi_H^\dagger(\vec{r}, t), \Psi_H^\dagger(\vec{r}', t')]_\pm &= \text{many-particle problem}. \quad (16.121)
\end{aligned}$$

For the general case with a spin quantum number, no new phenomena arise, see Fetter-Walecka, pages 67ff.

Of course we can use the a operators instead of the Ψ operators. With

$$\langle x | a_\alpha | 0 \rangle = e^{i\vec{k} \cdot \vec{r}} \chi_\sigma(s), \quad (16.122)$$

⁶We denote $[\cdot, \cdot]_-$ for the bosonic commutator $[\cdot, \cdot]$ and $[\cdot, \cdot]_+$ for the anticommutator $\{\cdot, \cdot\}$.

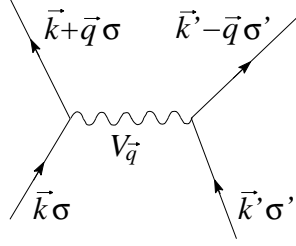


Fig. 16.1: Interaction vertex for the Hamiltonian (16.123). The incident vectors represent the annihilation operators $a_{\vec{k}\sigma}$, $a_{\vec{k}'\sigma'}$, the outgoing vectors represent the creation operators $a_{\vec{k}+\vec{q}\sigma}^\dagger$, $a_{\vec{k}'-\vec{q}\sigma'}^\dagger$. The exchange particle / potential $V_{\vec{q}}$ is described by the wavy line.

and $x = \vec{r}, s$, $\alpha = \vec{k}, \sigma$, and the interaction $V(\vec{r}, \vec{r}') = V(\vec{r} - \vec{r}')$, the above Hamiltonian has the form [see (??) and Fig. 16.1]

$$H = \sum_{\vec{k}\sigma} \frac{\hbar^2 k^2}{2m} a_{\vec{k}\sigma}^\dagger a_{\vec{k}\sigma} + \frac{1}{2} \sum_{\vec{k}\vec{k}'\vec{q}\sigma'\sigma} V_{\vec{q}} a_{\vec{k}+\vec{q}\sigma}^\dagger a_{\vec{k}'-\vec{q}\sigma'}^\dagger a_{\vec{k}'\sigma'} a_{\vec{k}\sigma} \quad (16.123)$$

and the equation of motion for the operators $a_{H\vec{k},\sigma}(t)$ is

$$i\hbar \frac{da_{H\vec{k}\sigma}(t)}{dt} = \frac{\hbar^2 k^2}{2m} a_{H\vec{k}\sigma}(t) + \sum_{\vec{k}'\vec{q}\sigma} V_{\vec{q}} a_{H\vec{k}'+\vec{q}\sigma}^\dagger(t) a_{H\vec{k}'\sigma}(t) a_{H\vec{k}\sigma}(t). \quad (16.124)$$

$\underbrace{\hspace{10em}}_{\rightarrow V_{\vec{q}} \rho_{-\vec{q}}}$

Chapter 17

Fermi sea: electrons

Material for this chapter can be found in the book by Fetter and Walecka and any text on solid state physics, see the topics Fermi sea, metals.

In this chapter, we apply the second-quantization formalism to the description of the electronic Fermi sea. This is one of the most prominent systems in condensed matter and, in more recent times, in quantum optics. It appears in two variants, in the continuum, where the electrons are described by plane waves, and on a lattice, where the electrons hop between sites, having in mind that each site represents one localized atomic orbital state. While free continuum-electrons with mass m are described by a parabolic dispersion $\varepsilon_k = \hbar^2 k^2 / 2m$, the dispersion of the lattice- (or so-called tight-binding) electrons is given by a cosine dispersion, $\varepsilon_k = -2t \cos ka$, where t is the hopping amplitude between sites and a is the lattice constant (see exercises, all exemplified in one dimension). The cosine dispersion is an energy-band characteristic of periodic systems, see also the Kronig-Penney model discussed in the quantum mechanics I lecture. Here, we focus on the case of free continuum electrons that describe well the physics of simple metals. We first describe the Fermi sea in the second quantization formalism and then study correlations arising due to the (fermionic) particle statistics. We then discuss the interacting free Fermi gas within the Hartree Fock approximation. Effects of finite temperatures will be discussed in chapter 23.

17.1 Fermi sea

We briefly summarize the description of the Fermi gas, see paragraph 13.6.7, and cast it into the second-quantized formalism. Starting from the plane wave states with spin $|\vec{k}, \sigma\rangle$ describing free fermions,

$$\langle \vec{r}, s | \vec{k}, \sigma \rangle = \varphi_{\vec{k}}(\vec{r}) \chi_{\sigma}(s) = \frac{1}{\sqrt{V}} e^{i\vec{k} \cdot \vec{r}} \chi_{\sigma}(s), \quad (17.1)$$

we define the associated creation operators ($x = \vec{r}, s$)

$$\langle x | c_{\vec{k}\sigma}^{\dagger} | 0 \rangle = \varphi_{\vec{k}}(\vec{r}) \chi_{\sigma}(s). \quad (17.2)$$

In second quantization, the free particle Hamiltonian $H = \sum_i p_i^2/2m$ takes the form

$$H = \sum_{\vec{k}, \sigma} \frac{\hbar^2 k^2}{2m} c_{\vec{k}\sigma}^{\dagger} c_{\vec{k}\sigma}. \quad (17.3)$$

The ground state accounts for the Fermi statistics (with the Pauli exclusion principle forbidding double-occupation of states, see Fig. 17.1) and is given by the filled Fermi sea

$$|\phi_0\rangle = \prod_{k \leq k_F, \sigma} c_{\vec{k}\sigma}^{\dagger} | 0 \rangle. \quad (17.4)$$

The associated occupation numbers $n_{\vec{k}\sigma}$

$$n_{\vec{k}\sigma} = \langle \Phi_0 | c_{\vec{k}\sigma}^{\dagger} c_{\vec{k}\sigma} | \phi_0 \rangle = \begin{cases} 1, & k \leq k_F, \\ 0, & \text{else} \end{cases}. \quad (17.5)$$

define the Fermi function $n_F(\varepsilon)$ at $T = 0$,

$$n_F(\varepsilon) = \Theta(\varepsilon - \varepsilon_F), \quad (17.6)$$

where we have substituted $k \rightarrow \varepsilon$ and $k_F \rightarrow \varepsilon_F$. The parameters k_F and ε_F (Fermi wavenumber and energy) are determined by the density n of the Fermi gas

$$k_F = (3\pi^2 n)^{1/3}, \quad \varepsilon_F = \frac{\hbar^2 k_F^2}{2m}. \quad (17.7)$$

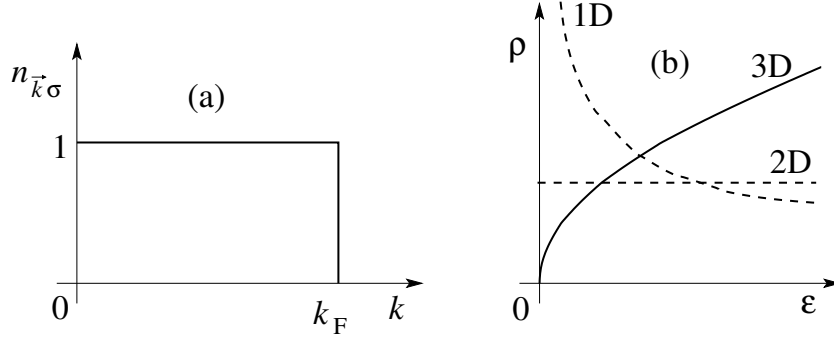


Fig. 17.1: (a) Occupation number $n_{\vec{k}\sigma}$ of the electron states with wave vector \vec{k} and spin σ for the Fermi sea. (b) Densities of state $\rho(\epsilon)$ for free fermions in 3D (solid line) and in 2D and 1D.

Typical values for k_F (for $n = N/V \sim 10^{23} \text{cm}^{-3}$ in metals) are $k_F \sim 1 \text{ \AA}^{-1}$ and correspondingly $\epsilon_F \sim \text{eV}$. The density of states (DOS) is given by (see Fig. 17.1(b))

$$\rho(\epsilon) = \frac{m}{2\pi^2 \hbar^3} \sqrt{2m\epsilon} = \frac{3}{4} \frac{n}{\epsilon_F} \left(\frac{\epsilon}{\epsilon_F} \right)^{1/2} \quad (17.8)$$

and their value at the Fermi edge is

$$\rho(\epsilon_F) = \frac{3}{4} \frac{n}{\epsilon_F}, \quad \rho = \text{DOS/Spin}, \quad 2\rho = \text{DOS}. \quad (17.9)$$

As an exercise, calculate the density of states for the Fermi gases in one dimension ($\rho_{1D} \propto 1/\sqrt{\epsilon}$) and in two dimensions ($\rho_{2D} \propto \text{const.}$). In the following, we calculate the density, the density matrix $\langle \Phi_0 | \Psi_s^\dagger(\vec{r}) \Psi_s(\vec{r}') | \Phi_0 \rangle$, and the pair correlations $\langle \Phi_0 | \Psi_s^\dagger(\vec{r}) \Psi_{s'}^\dagger(\vec{r}') \Psi_{s'}(\vec{r}') \Psi_s(\vec{r}) | \Phi_0 \rangle$ in the Fermi gas.

17.1.1 Homogeneous density

Let us first show that the density $n = N/V$ is homogeneous in the Fermi sea. With the definition (16.23) of the one-particle operators, the density $\rho(\vec{r}_0) = \delta(\vec{r} - \vec{r}_0)$ can be expressed as¹

$$\rho = \sum_{ij} \langle i | \rho | j \rangle c_i^\dagger c_j, \quad (17.10)$$

¹We define $i = \vec{k}'\sigma'$ and $j = \vec{k}\sigma$.

$$= \sum_{\vec{k}'\sigma', \vec{k}\sigma} \langle \vec{k}'\sigma' | \rho | \vec{k}\sigma \rangle c_{\vec{k}'\sigma'}^\dagger c_{\vec{k}\sigma}, \quad (17.11)$$

$$\begin{aligned} \langle \vec{k}'\sigma' | \rho | \vec{k}\sigma \rangle &= \delta_{\sigma,\sigma'} \int \frac{d^3r}{V} e^{i(\vec{k}-\vec{k}')\cdot\vec{r}} \delta(\vec{r}-\vec{r}_0) \\ &= \delta_{\sigma,\sigma'} e^{i(\vec{k}-\vec{k}')\cdot\vec{r}_0} / V, \end{aligned} \quad (17.12)$$

$$\begin{aligned} \langle \rho(\vec{r}_0) \rangle &= \langle \Phi_F | \rho(\vec{r}_0) | \Phi_F \rangle \\ &= \frac{1}{V} \sum_{\vec{k}\vec{k}'\sigma} e^{i(\vec{k}-\vec{k}')\cdot\vec{r}_0} \underbrace{\langle \Phi_F | c_{\vec{k}'\sigma}^\dagger c_{\vec{k}\sigma} | \Phi_F \rangle}_{(16.21): \delta_{\vec{k}\vec{k}'} n_{\vec{k}\sigma}} \end{aligned} \quad (17.13)$$

$$= \frac{1}{V} \sum_{\vec{k}\sigma} n_{\vec{k}\sigma} = \frac{2}{V} \sum_{|\vec{k}| \leq k_F} n_{\vec{k}} = n. \quad (17.14)$$

Therefore, we find a homogeneous density $n = N/V$ in the Fermi gas. Alternatively, we can formulate the density through field operators, $\rho_s(\vec{r}) = \Psi_s^\dagger(\vec{r})\Psi_s(\vec{r})$, and go directly to (17.13) using a basis transformation. We will do this below for the more general density-matrix.

17.1.2 Density Matrix

The local density operator for electrons with spin s is given by $\rho_s(\vec{r}) = \Psi_s^\dagger(\vec{r})\Psi_s(\vec{r})$. The density matrix of the Fermi sea involves its generalization to different positions \vec{r} and \vec{r}' (non-diagonal density operator),

$$\begin{aligned} G_s(\vec{r}-\vec{r}') &\equiv \langle \Phi_0 | \Psi_s^\dagger(\vec{r})\Psi_s(\vec{r}') | \Phi_0 \rangle \\ &= \langle \Phi_0 | \sum_{\vec{k}\vec{k}'} \frac{\exp(-i\vec{k}\cdot\vec{r})}{\sqrt{V}} \frac{\exp(i\vec{k}'\cdot\vec{r}')}{\sqrt{V}} \underbrace{c_{\vec{k}s}^\dagger c_{\vec{k}'s}}_{\delta_{\vec{k}\vec{k}'} n_{\vec{k}s}} | \Phi_0 \rangle. \end{aligned} \quad (17.15)$$

$G_s(\vec{r}-\vec{r}')$ gives us the amplitude for the state with an electron shifted from \vec{r}' to \vec{r} . The evaluation of the Kronecker $\delta_{\vec{k}\vec{k}'}$ symbol and a transition to the continuum leads to

$$\begin{aligned} G_s(\vec{r}-\vec{r}') &= \frac{1}{V} \sum_{\vec{k}} e^{-i\vec{k}\cdot(\vec{r}-\vec{r}')} n_{\vec{k}s} = \frac{n}{2} g_s(\vec{r}-\vec{r}') \\ &= \int \frac{d^3k}{(2\pi)^3} e^{-i\vec{k}\cdot(\vec{r}-\vec{r}')} \theta(k_F - k) \\ &= \frac{1}{4\pi^2} \int_0^{k_F} dk k^2 \int_{-1}^1 dz e^{-ik|\vec{r}-\vec{r}'|z} \end{aligned} \quad (17.16)$$

$$= \frac{3n}{2} \frac{\sin x - x \cos x}{x^3} \Big|_{x=k_F|\vec{r}-\vec{r}'|} = \frac{3n}{2} \frac{j_1(k_F|\vec{r}-\vec{r}'|)}{k_F|\vec{r}-\vec{r}'|},$$

with the spherical Bessel function j_1 . The result is shown in Fig. 17.2; the limits $x = k_F r \rightarrow 0, \infty$ are given by

$$\begin{aligned} x \rightarrow 0 : \quad & (x - x^3/6 - x + x^3/2)/x^3|_{x \rightarrow 0} = 1/3 \Rightarrow G_s(0) = n/2 \\ x \rightarrow \infty : \quad & G_s(x) \sim 3n(\cos x)/2x^2. \end{aligned} \quad (17.17)$$

With $G_s(0) = n/2$, we find a homogeneous density $\rho = G_\uparrow(0) + G_\downarrow(0) = n$.

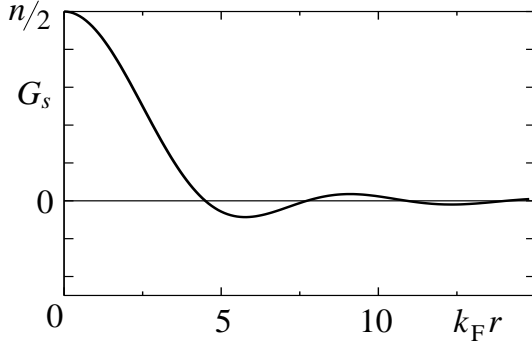


Fig. 17.2: Density matrix $G_s(k_F r)$.

17.2 Pair correlations

The pair correlation gives us the relative probability to find an electron at \vec{r}' , s' if there is an electron at \vec{r} , s ,

$$\left(\frac{n}{2}\right)^2 g_{ss'}(\vec{r} - \vec{r}') = \langle \Phi_0 | \Psi_s^\dagger(\vec{r}) \Psi_{s'}^\dagger(\vec{r}') \Psi_{s'}(\vec{r}') \Psi_s(\vec{r}) | \Phi_0 \rangle. \quad (17.18)$$

The non-trivial correlations we discuss here arise due to the particle statistics or Pauli's exclusion principle; they affect only equal spins. Additional correlations appear when the Coulomb interaction is taken into account.

The calculation of the pair correlator is a nice exercise in calculating with the second-quantization formalism. We first express the field operators $\Psi_s(\vec{r})$ by the annihilators $c_{\vec{k}s}$,

$$\left(\frac{n}{2}\right)^2 g_{ss'}(\vec{r} - \vec{r}') = \frac{1}{V^2} \sum_{\vec{p}\vec{p}'\vec{q}\vec{q}'} e^{-i(\vec{p}-\vec{p}')\cdot\vec{r}} e^{-i(\vec{q}-\vec{q}')\cdot\vec{r}'} \langle \Phi_0 | c_{\vec{p}s}^\dagger c_{\vec{q}s}^\dagger c_{\vec{q}'s'} c_{\vec{p}'s} | \Phi_0 \rangle.$$

First, we examine the case $s \neq s'$, then $\vec{q}' = \vec{q}$ and $\vec{p}' = \vec{p}$, otherwise the expectation value $\langle \dots \rangle = 0$ vanishes² and we find

$$\begin{aligned} \langle \Phi_0 | c_{\vec{p}s}^\dagger c_{\vec{q}s'}^\dagger c_{\vec{q}'s'} c_{\vec{p}'s} | \Phi_0 \rangle &= \delta_{\vec{p}\vec{p}'} \delta_{\vec{q}\vec{q}'} \langle \Phi_0 | c_{\vec{p}s}^\dagger c_{\vec{q}s'}^\dagger c_{\vec{q}s'} c_{\vec{p}s} | \Phi_0 \rangle \\ &= \delta_{\vec{p}\vec{p}'} \delta_{\vec{q}\vec{q}'} \langle \phi_0 | c_{\vec{p}s}^\dagger n_{\vec{q}s'} c_{\vec{p}s} | \phi_0 \rangle \\ &= \delta_{\vec{p}\vec{p}'} \delta_{\vec{q}\vec{q}'} n_{\vec{p}s} n_{\vec{q}s'}, \end{aligned} \quad (17.19)$$

$$\begin{aligned} \left(\frac{n}{2}\right)^2 g_{ss'}(\vec{r}) &= \frac{1}{V^2} \sum_{\vec{p}\vec{q}} n_{\vec{p}s} n_{\vec{q}s'} = n_s n_{s'} \stackrel{n_s=n/2}{=} \left(\frac{n}{2}\right)^2 \\ \Rightarrow g_{s \neq s'}(\vec{r}) &= 1. \end{aligned} \quad (17.20)$$

For $s \neq s'$, we find a constant value $g_{s \neq s'}(k_F |\vec{r} - \vec{r}'|) = 1$ and hence electrons with opposite spin are uncorrelated, see Fig. 17.3.

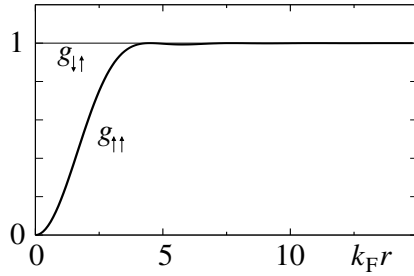


Fig. 17.3: Pair correlators $g_{\uparrow\uparrow}(k_F r)$ (with exchange hole due to the Pauli principle) and $g_{\uparrow\downarrow}(k_F r)$ (uncorrelated, as spin \uparrow and spin \downarrow states are distinguishable).

Second, we focus on the case $s = s'$ and obtain contributions for $\vec{p}' = \vec{p}$, $\vec{q}' = \vec{q}$ and for $\vec{p}' = \vec{q}$, $\vec{q}' = \vec{p}$,

$$\begin{aligned} \langle \Phi_0 | c_{\vec{p}s}^\dagger c_{\vec{q}s}^\dagger c_{\vec{q}'s} c_{\vec{p}'s} | \Phi_0 \rangle &= \delta_{\vec{p}\vec{p}'} \delta_{\vec{q}\vec{q}'} \langle \Phi_0 | c_{\vec{p}s}^\dagger c_{\vec{q}s}^\dagger c_{\vec{q}s} c_{\vec{p}s} | \Phi_0 \rangle \\ &\quad + \delta_{\vec{p}'\vec{q}} \delta_{\vec{q}'\vec{p}} \langle \Phi_0 | c_{\vec{p}s}^\dagger c_{\vec{q}s}^\dagger \underbrace{c_{\vec{p}s} c_{\vec{q}s}}_{-c_{\vec{q}s} c_{\vec{p}s}} | \Phi_0 \rangle \\ &= (\delta_{\vec{p}\vec{p}'} \delta_{\vec{q}\vec{q}'} - \delta_{\vec{p}'\vec{q}} \delta_{\vec{q}'\vec{p}}) n_{\vec{p}s} n_{\vec{q}s}, \end{aligned} \quad (17.21)$$

$$\begin{aligned} \left(\frac{n}{2}\right)^2 g_{ss}(\vec{r}) &= \frac{1}{V^2} \sum_{\vec{p}\vec{q}} \left(1 - e^{-i(\vec{p}-\vec{q}) \cdot \vec{r}}\right) n_{\vec{p}s} n_{\vec{q}s} \\ &= \left(\frac{n}{2}\right)^2 - \left(\frac{1}{V} \sum_{\vec{p}} e^{-i\vec{p} \cdot \vec{r}} n_{\vec{p}s}\right)^2 \end{aligned}$$

²The expectation value $\langle \dots \rangle$ vanishes for $\vec{q}' \neq \vec{q}$ or $\vec{p}' \neq \vec{p}$ because different states in $|\Phi_0\rangle$ and in $c_{\vec{p}s}^\dagger c_{\vec{q}s}^\dagger c_{\vec{q}'s'} c_{\vec{p}'s} |\Phi_0\rangle$ are occupied.

$$\begin{aligned}
&= \left(\frac{n}{2}\right)^2 - G_s^2(\vec{r}) \equiv \left(\frac{n}{2}\right)^2 [1 - g_s^2(\vec{r})], \quad (17.22) \\
&= \left(\frac{n}{2}\right)^2 \left[1 - (9/x^6) (\sin x - x \cos x)^2\right]_{x=k_F r}.
\end{aligned}$$

As a result, we find that the correlator $g_{ss}(k_F|\vec{r}-\vec{r}'|)$ for equal spin shows 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 a spin s electron, see Fig. 17.3. The pair correlator $g(\vec{r}) = (g_{\uparrow\uparrow}(\vec{r}) + g_{\uparrow\downarrow}(\vec{r}))/2$ of the overall system is shown in Fig. 17.4.

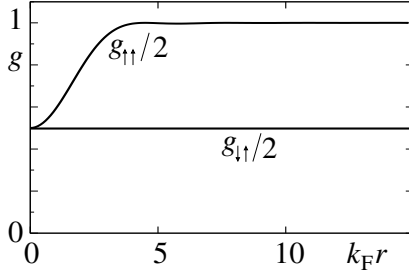


Fig. 17.4: Pair correlator $g = [g_{\uparrow\uparrow} + g_{\uparrow\downarrow}]/2$ of the entire electronic system.

These above results have been calculated at the Hartree-Fock level: the optimal Slater determinants for the homogeneous (and hence translation invariant) Fermi gas are made from plane-wave one-particle states. The state $|\Phi_0\rangle$ is thus an optimized Slater determinant. If one additionally considers correlations due to Coulomb interactions, one finds corrections as shown in Fig. 17.5. Again one finds that exactly one electron is missing around an

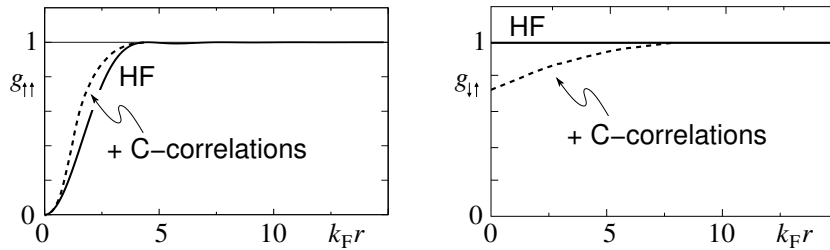


Fig. 17.5: Corrections (schematic) to the pair correlators $g_{\uparrow\uparrow}$ and $g_{\uparrow\downarrow}$ taking correlation effects due to the Coulomb interaction into account.

electron,

$$n \int d^3r [g(\vec{r}) - 1] = -1; \quad (17.23)$$

the exchange hole therefore has the charge $e > 0$.

Finally, a few comments on notation and jargon. The correlator $g_{ss'}(\vec{r} - \vec{r}')$ is calculated by ‘contraction’ of the field operators (the contraction is often indicated by corresponding lines, for the first term in (17.24) $\langle \Phi_0 | \overline{\Psi_s^\dagger(\vec{r}) \Psi_{s'}^\dagger(\vec{r}') \Psi_{s'}(\vec{r}') \Psi_s(\vec{r})} | \Phi_0 \rangle$ and for the second term $\langle \Phi_0 | \overline{\Psi_s^\dagger(\vec{r}) \Psi_{s'}^\dagger(\vec{r}') \Psi_s(\vec{r}) \Psi_{s'}(\vec{r}')} | \Phi_0 \rangle$),

$$\begin{aligned} \left(\frac{n}{2}\right)^2 g_{ss'}(\vec{r} - \vec{r}') &= \langle \Phi_0 | \Psi_s^\dagger(\vec{r}) \Psi_{s'}^\dagger(\vec{r}') \Psi_{s'}(\vec{r}') \Psi_s(\vec{r}) | \Phi_0 \rangle \\ &= \langle \Phi_0 | \Psi_s^\dagger(\vec{r}) \Psi_s(\vec{r}) | \Phi_0 \rangle \langle \Phi_0 | \Psi_{s'}^\dagger(\vec{r}') \Psi_{s'}(\vec{r}') | \Phi_0 \rangle \\ &\quad - \langle \Phi_0 | \Psi_s^\dagger(\vec{r}) \Psi_{s'}(\vec{r}') | \Phi_0 \rangle \langle \Phi_0 | \Psi_{s'}^\dagger(\vec{r}') \Psi_s(\vec{r}) | \Phi_0 \rangle. \end{aligned} \quad (17.24)$$

For a proof, calculate $\langle \Phi_0 | \Psi_s^\dagger(\vec{r}) \Psi_s(\vec{r}) | \Phi_0 \rangle = n/2$ and $\langle \Phi_0 | \Psi_s^\dagger(\vec{r}) \Psi_{s'}(\vec{r}') | \Phi_0 \rangle = \delta_{ss'}(n/2)g_s$. For $s \neq s'$, one finds $(n/2)^2 g_{s \neq s'} = (n/2)^2 - 0$ and for $s = s'$, $(n/2)^2 g_{ss} = (n/2)^2 - (n/2)^2 g_s^2$. 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)

While the factorization for the non-interacting system is exact, the Hartree-Fock ground state is only an approximation when interactions between the particles are taken into account. In this sense, the above factorization is only approximately valid if the correlations $\langle \Phi | \Psi_s^\dagger \Psi_{s'}^\dagger \Psi_{s'} \Psi_s | \Phi \rangle$ are considered in the real ground state $|\Phi\rangle$ of the interacting system,

$$\begin{aligned} \langle \Phi | \Psi_s^\dagger \Psi_{s'}^\dagger \Psi_{s'} \Psi_s | \Phi \rangle &\stackrel{\text{HFA}}{\approx} \langle \Phi | \Psi_s^\dagger \Psi_s | \Phi \rangle \langle \Phi | \Psi_{s'}^\dagger \Psi_{s'} | \Phi \rangle \\ &\quad - \langle \Phi | \Psi_s^\dagger \Psi_{s'} | \Phi \rangle \langle \Phi | \Psi_{s'}^\dagger \Psi_s | \Phi \rangle. \end{aligned} \quad (17.25)$$

17.3 Ground state energy for interacting electrons

The Hamiltonian of the interacting electron system is given in (16.105). We calculate $E_{\text{HF}} = \langle H \rangle$ in Hartree-Fock-Approximation: For the (homogeneous) electron gas, the optimal Hartree-Fock single-particle waves are plane waves $e^{i\vec{k}\cdot\vec{r}} \chi_\sigma(s_z)/\sqrt{V}$ and $|\Phi_0\rangle$ defined in (17.4) is an optimized Hartree-Fock ansatz. If we combine (16.105) and (16.106), we obtain

$$E_{\text{HF}} = \langle \Phi_0 | H | \Phi_0 \rangle = \langle \phi_0 | T | \Phi_0 \rangle + \langle \Phi_0 | V | \Phi_0 \rangle, \quad (17.26)$$

with the kinetic energy

$$\begin{aligned}
\langle T \rangle &= \langle \Phi_0 | \sum_{\vec{k}\sigma} \frac{\hbar^2 k^2}{2m} c_{\vec{k}\sigma}^\dagger c_{\vec{k}\sigma} | \Phi_0 \rangle \\
&= \underbrace{2}_{\text{Spin}} V \int_0^{k_F} \frac{4\pi k^2 dk}{(2\pi)^3} \frac{\hbar^2 k^2}{2m} \\
&= \frac{V}{2\pi^2} \frac{\hbar^2}{m} \frac{1}{5} k_F^5 \\
&= \frac{V}{\pi^2} \varepsilon_F \frac{k_F^3}{5} \\
&= \frac{3}{5} \varepsilon_F N,
\end{aligned} \tag{17.27}$$

with $N = nV$. For the interaction energy $\langle V \rangle$, we obtain

$$\begin{aligned}
\langle V \rangle &= \frac{1}{2} \sum_{ss'} \int d^3r d^3r' V(\vec{r} - \vec{r}') \underbrace{\langle \Phi_0 | \Psi_s^\dagger(\vec{r}) \Psi_{s'}^\dagger(\vec{r}') \Psi_{s'}(\vec{r}') \Psi_s(\vec{r}) | \Phi_0 \rangle}_{(n/2)^2 g_{ss'}(\vec{r} - \vec{r}')} \\
&= \frac{1}{2} \int d^3r d^3r' V(\vec{r} - \vec{r}') \left[\left(\frac{n}{2} \right)^2 \cdot \underbrace{2}_{s \neq s'} + \left(\frac{n}{2} \right)^2 (1 - g_s^2) \underbrace{2}_{s=s'} \right] \\
&= \frac{n^2}{2} V \int d^3r V(\vec{r}) - \frac{n^2}{4} V \int d^3r V(\vec{r}) g_s^2(\vec{r}),
\end{aligned} \tag{17.28}$$

where

$$g_s^2(\vec{r}) = \frac{9}{x^6} (\sin x - x \cos x)^2 \Big|_{x=k_F r}. \tag{17.29}$$

The first term in (17.28) is the Hartree energy; it takes the repulsion of the electrons into account. The second term has its origin in the symmetry properties of $|\Phi_0\rangle$ and is called exchange energy. The two terms can be visualized using Feynman diagrams³, see Fig. 17.6 where the Hartree and

³Feynman 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 $\langle c^\dagger c^\dagger \dots c \rangle$ are decomposed into products of free propagators $\langle c^\dagger c \rangle$. 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) come with additional signs, one integrates over \vec{k} values of the inner lines, etc. Once the rules for the construction of diagrams have been established (each set of rules results from the problem-specific expansion in a small parameter), an analytical expression can be written down for each diagram.

exchange diagrams for the Green's function $G(\vec{k}, t - t') = -i\langle T[c_{\vec{k}}(t) c_{\vec{k}}^\dagger(t')] \rangle$ (T the time ordering operator and $\langle \dots \rangle$ the ground state expectation value) are shown.

For the Coulomb interaction $V(r) = e^2/r$, the integral $\int d^3r V(r) = 4\pi \int dr r^2 e^2/r$ is divergent.⁴ However, we should take into account the (approximately homogeneous) positively charged background, which is generated by the ions and has the same charge density as the electron gas. We then obtain several compensating terms,

$$E_H = \frac{Nn}{2} \int d^3r \left(\frac{e^2}{r} + \frac{e^2}{r} - 2\frac{e^2}{r} \right) = 0; \quad (17.30)$$

the first term describes the repulsive electron-electron interaction, the second term results from the (also repulsive) ion-ion interaction and the third term takes into account the interaction between the electrons and the ions; the factor 2 is a consequence of the distinction between electrons and ions.

That leaves only the (non-divergent) exchange term that contributes

$$E_{\text{exc}} = -\frac{9}{4}nN \frac{e^2}{k_F^2} 4\pi \underbrace{\int_0^\infty dx \frac{(\sin x - x \cos x)^2}{x^5}}_{1/4} \quad (17.31)$$

$$= -\frac{3}{4\pi} N e^2 k_F \quad (17.32)$$

to the energy.

To describe an electron gas / an electron liquid, the Seitz radius r_s is usually introduced, which gives the free volume available for an electron. If we measure r_s in Bohr radii a_B , we obtain

$$\frac{4\pi}{3}(r_s a_B)^3 = \frac{1}{n}, \quad r_s = \frac{5.44}{\sqrt[3]{n_{22}}}, \quad (17.33)$$

where $n = n_{22} \cdot 10^{22} \text{cm}^{-3}$ is the electron density. For typical metals

$$n \approx \underbrace{0.9 \cdot 10^{22}}_{\text{Cesium}} - \underbrace{2.5 \cdot 10^{23}}_{\text{Beryllium}} \text{cm}^{-3}, \quad (17.34)$$

⁴The divergence is a consequence of the long range of the Coulomb interaction and creates an incompressible system. The overall system can only be compressed if the compensating (here ionogenic) background takes part in the deformation. The finite energy of the plasmons in the $\vec{k} \rightarrow 0$ limit is a consequence of the incompressibility of the electrons. If the ions resonate (in a phononic excitation), the energy of the mode goes to 0 for $\vec{k} \rightarrow 0$.

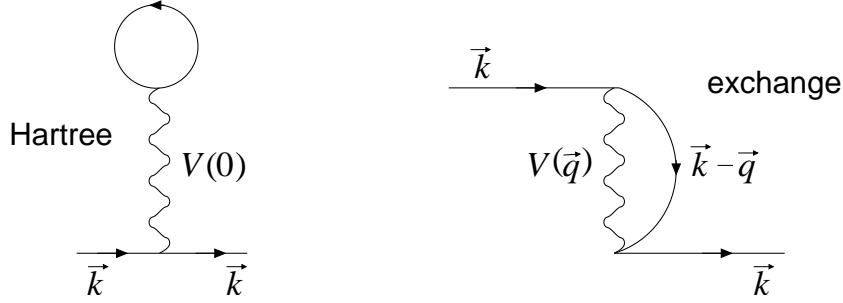


Fig. 17.6: Hartree (left) and exchange (right) self-energy graph for the Green's function. The free Green's function $G_0(\vec{k}, \omega) = 1/(\omega - \epsilon_{0\vec{k}} + i\delta)$, $\epsilon_{0\vec{k}} = \hbar^2 k^2/2m$, $\delta = 0^+$ (for the retarded component), picks up the interaction-induced self-energy $\Sigma(\vec{k}, \omega)$, $G_0 \rightarrow G(\vec{k}, \omega) = 1/[\omega - (\epsilon_{0\vec{k}} + \Sigma(\vec{k}, \omega)) + i\delta]$. Poles in the complex ω plane at a fixed \vec{k} mark the energies $\epsilon_{\vec{k}}$ of the quasiparticles. In the lowest order of e^2 the self-energy has two terms $\Sigma(\vec{k}, \omega) \approx \Sigma_H + \Sigma_{\text{exc}}(\vec{k})$, Hartree (disappears with the positive background) and exchange (see (17.31)). The k -sum of these terms gives the corrections $E_H + E_{\text{exc}}$ to the kinetic term in the Hartree-Fock energy E_{HF} .

and thus r_s assumes typical values $r_s \approx 2 - 6$. Using this notation, we arrive at the standard Hartree-Fock result for the energy of the electron gas in the form

$$\frac{E_{\text{HF}}}{N} = \frac{3}{5} \epsilon_F - \frac{3}{4\pi} e^2 k_F \approx \left(\frac{2.21}{r_s^2} - \frac{0.916}{r_s} \right) \text{Ry}. \quad (17.35)$$

Note that $E_0/N < E_{\text{HF}}/N$, because $|\Phi_0\rangle$ is just a (Rayleigh-Ritz type) approximation to the real ground state $|\Phi\rangle$.

The result (17.35) is correct for high densities and correspondingly small radii r_s . The correlation energies due to the Coulomb interaction are not taken into account in (17.35). They are calculated using diagrammatic techniques and one can find the following expansion for $r_s < 2.5$ (note that $(r_s)^0$ is either a constant or a logarithm $\ln r_s$)

$$\frac{E_0}{N \text{ Ry}} \approx \underbrace{\frac{2.21}{r_s^2} - \frac{0.916}{r_s}}_{E_{\text{HF}}/N} - \underbrace{0.094 + 0.0622 \ln r_s + 0.018 r_s \ln r_s + \dots}_{E_{\text{Corr}}/N < 0}. \quad (17.36)$$

The result (17.36) is not analytical for large densities $r_s \rightarrow 0$.

For small densities $r_s \rightarrow \infty$, one expects the transition to a crystal named after Wigner. The expression for the correlation energy of the Wigner crystal can be cast into the form

$$\frac{E_{\text{kor}}}{N} \approx -\frac{0.88}{r_s + 7.8} \text{Ry} < 0. \quad (17.37)$$

Figure 17.7 finally shows the phases to be expected in a system with an interaction of the type $e^{-r/\lambda}/r$ that interpolates between long ($\lambda \rightarrow \infty$) and short ($\lambda \rightarrow 0$) ranges.

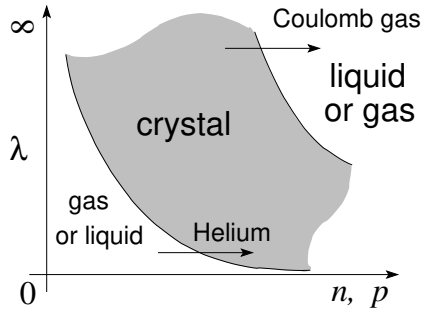


Fig. 17.7: Phases in a system with (shielded) Coulomb interaction; λ = interaction-range, n = density, p = pressure, dimensionless parameter $= n\lambda^3$. Such a phase diagram with re-entrance in the liquid / gaseous phase occurs (e.g.) in the vortex matter of type II superconductors and in Helium.

Chapter 18

Quantization of a string: phonons

Further material on this chapter can be found in the book by Franz Gross (including chapter 14 for the path integrals).

We apply the second quantized formalism of the previous chapter to the linear chain and string problems. The elastic deformation modes in such a one-dimensional medium are called phonons; these are massless Bosons with/without dispersion in the two cases of the chain/string. Three-dimensional extensions are considered in lectures on solid state physics (for a bulk elastic medium) and in the following chapter, where we will quantize the electromagnetic field—the resulting modes are the photons, massless Bosons with linear dispersion and spin.¹ In the second part of the chapter, we will develop the path-integral formalism of bosonic fields, a generalization of the path-integral introduced for particle in QM I; the role of the trajectories $x(t)$ in Chapter 1 is now assumed by the bosonic field $\phi(x, t)$; fermionic path integrals can be defined using Grassman variables, which is beyond the scope of this lecture, however. Both chapters, on phonons and on photons, have a nice contextual significance for the particle–wave duality in quantum mechanics: so far, we have assigned modes and wave functions to particles. In this and the next chapter, we will do the reverse by assigning particles to elastic or radiation modes/waves (phonons or photons). The particles are associated with creation and ‘annihilation’ operators, which add or remove a quantum to/from the mode, just as in the discussion of the

¹A system with linear dispersion is usually said to be ‘non-dispersive’; in such a system, a wave packet does not disperse but keeps its form.

harmonic oscillator using the operators a^\dagger and a .

18.1 One-dimensional chain/string

We consider a discrete system of point masses (mass m , position $x_i = il$, l the separation between masses) connected by springs with spring constant k , see Fig. 18.1(a). Going over to the continuum limit defines a massive, elastic string, see Fig. 18.1(b), which is characterized by its mass density $\mu = m/l$ and elasticity $\varepsilon = kl$.

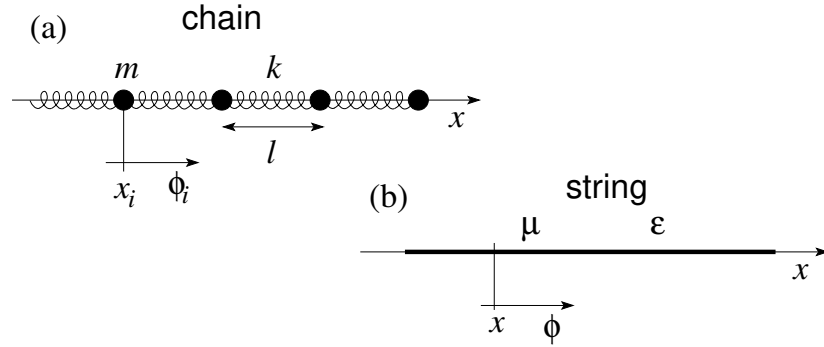


Fig. 18.1: (a) Discrete linear chain with point masses m , at positions $x_i = il$ separated by the distance l and connected by springs with spring constant k . The (longitudinal) displacement of the point masses are described by the amplitudes ϕ_i . (b) Continuous limit characterized by the mass density $\mu = m/l$ and the elasticity $\varepsilon = kl$.

In the following, we develop both the discrete and continuous versions, the linear chain and the string. We denote the system length $L = Nl$ and define the (longitudinal) displacements ϕ_i (chain) and $\phi(x)$ (string). The kinetic and potential energies are

$$\begin{array}{ll}
 \text{chain} & \text{string} \\
 E_{\text{kin}} = \frac{m}{2} \sum_{i=1}^N (\partial_t \phi_i)^2, & E_{\text{kin}} = \frac{\mu}{2} \int_0^L dx (\partial_t \phi)^2, \\
 E_{\text{pot}} = \frac{k}{2} \sum_{i=1}^{N-1} (\phi_{i+1} - \phi_i)^2, & E_{\text{pot}} = \frac{\varepsilon}{2} \int_0^L dx (\partial_x \phi)^2.
 \end{array}$$

Combining these two, we obtain the Lagrangians

$$\mathcal{L}_c = \sum_i \left[\frac{m}{2} \dot{\phi}_i^2 - \frac{k}{2} (\phi_{i+1} - \phi_i)^2 \right], \quad \mathcal{L}_s = \int_0^L dx \left[\frac{\mu}{2} \dot{\phi}^2 - \frac{\varepsilon}{2} \phi'^2 \right]. \quad (18.1)$$

The equations of motion can be derived from the principle of least action $\delta S = 0$, where $S = \int dt \mathcal{L}(t)$, and we find the results

$$\begin{aligned} \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\phi}_i} - \frac{\partial \mathcal{L}}{\partial \phi_i} &= 0, & \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\phi}} + \frac{d}{dx} \frac{\partial \mathcal{L}}{\partial \phi'} - \frac{\partial \mathcal{L}}{\partial \phi} &= 0, \\ m\ddot{\phi}_i - k(\phi_{i+1} - 2\phi_i + \phi_{i-1}) &= 0, & \mu\ddot{\phi} - \varepsilon\phi'' &= 0. \end{aligned} \quad (18.2)$$

We choose periodic boundary conditions,

$$\phi_0 = \phi_N, \quad \phi(0) = \phi(L), \quad (18.3)$$

$$\partial_t \phi_0 = \partial_t \phi_N, \quad \partial_t \phi(0) = \partial_t \phi(L), \quad (18.4)$$

and find the normal modes of the chain and the string in the form

$$\begin{aligned} \phi_{n,j}(t) &= \frac{1}{\sqrt{N}} e^{i(k_n j l - \omega_n t)}, & \phi_n(x, t) &= \frac{1}{\sqrt{L}} e^{i(k_n x - \omega_n t)}, \\ k_n &= \frac{2\pi}{L} n, & k_n &= \frac{2\pi}{L} n, \\ \omega_n &= 2\sqrt{\frac{k}{m}} \left| \sin \frac{k_n l}{2} \right|, & \omega_n &= \sqrt{\frac{\varepsilon}{\mu}} |k_n| > 0. \end{aligned} \quad (18.5)$$

The dispersions $\omega(k)$ are shown schematically in Fig. 18.2; while the spectrum for the chain is dispersive, the one of the string is linear, i.e., non-dispersive. As a result, a wave-packet in the chain disperses, while the one in a string keeps its form (as it has a constant group-velocity).

The modes satisfy orthonormality and completeness relations,

$$\sum_{j=0}^{N-1} \phi_{n,j}^*(t) \phi_{m,j}(t) = \delta_{nm}, \quad \int_0^L dx \phi_n^*(x, t) \phi_m(x, t) = \delta_{nm}, \quad (18.6)$$

$$\sum_{n=-N/2}^{N/2} \phi_{n,j}^*(t) \phi_{n,i}(t) = \delta_{ji}, \quad \sum_{n=-\infty}^{\infty} \phi_n^*(x, t) \phi_n(y, t) = \delta(x - y). \quad (18.7)$$

The transition from the discrete chain to the continuous string can be done in each step by requiring $L = Nl = \text{const.}$, $\mu = m/l$, and $\varepsilon = kl$ by letting $l \rightarrow 0$.

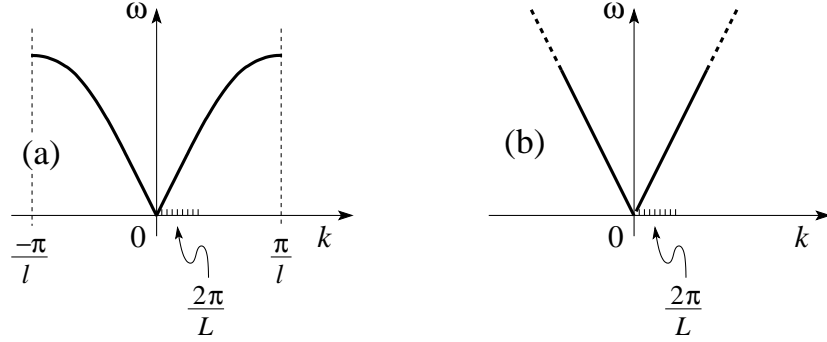


Fig. 18.2: Spectra of the discrete chain (a) and continuous string (b). Note the curvature of the dispersion at the Brillouin Zone boundaries $\pm\pi/l$ for the discrete chain: modes with crystal momentum π/l (and those with $-\pi/l$) are perfectly reflected on the periodic structure. The spectrum of the continuous string has no limitation on short-wave modes corresponding to large k -values: the entire continuous space can be sampled by modes with arbitrarily large k -values. Note that both spectra define the same speed of sound, $\omega_n \sim 2\sqrt{k/m}(|k_n|l/2) \sim \sqrt{(\varepsilon/l)/(\mu l)}l|k_n| \sim \sqrt{\varepsilon/\mu}|k_n|$ at small values of k_n .

We continue with the string. With the help of the complete set of modes (18.5), we can expand an arbitrary solution $\Phi(x, t)$

$$\begin{aligned}\Phi(x, t) &= \sum_{n=-\infty}^{\infty} \sqrt{\frac{h}{2\mu\omega_n}} [a_n \phi_n(x, t) + a_n^* \phi_n^*(x, t)], \\ &= \sum_{n=-\infty}^{\infty} \sqrt{\frac{h}{2\mu\omega_n L}} \left[a_n e^{i(k_n x - \omega_n t)} + a_n^* e^{-i(k_n x - \omega_n t)} \right], \quad (18.8)\end{aligned}$$

where the factor $\sqrt{h/2\mu\omega_n}$ with arbitrary action h ensures that the coefficients a_n are dimensionless: with $(h/\mu\omega)^{1/2} = (\text{Length})^{3/2}$ and $[\phi_n] = 1/\sqrt{\text{Length}}$, we have a displacement $[\Phi]$ measured in units of length. This choice makes full sense once the theory is quantized by going over to dimensionless operators (in place of amplitudes) and replacing $h \rightarrow \hbar$, see later.

Taking a closer look at the amplitudes in the expansion (18.8), one often uses the following different forms,

$$a_n \phi_n(x, t) = a_n(t) \phi_n(x),$$

$$\begin{aligned} a_n(t) &= a_n(0)e^{-i\omega_n t}, \quad \omega_n > 0, \\ \ddot{a}_n(t) + \omega_n^2 a_n(t) &= 0; \end{aligned} \quad (18.9)$$

the amplitudes $a_n(t)$ therefore satisfy the differential equation of a harmonic oscillator.

The energy of the string includes the sum of kinetic and potential energies in (18.1) and can be written as

$$\begin{aligned} & \frac{\mu}{2} \int_0^L dx [\dot{\Phi}(x, t)]^2 + \frac{\varepsilon}{2} \int_0^L dx [\Phi'(x, t)]^2 \\ &= \frac{1}{2} \sum_{n=-\infty}^{\infty} \left\{ \mu \omega_n^2 \frac{h}{2\mu\omega_n} [(2|a_n|^2 - a_n a_{-n} - a_n^* a_{-n}^*)(t)] \right. \\ & \quad \left. + \varepsilon k_n^2 \frac{h}{2\mu\omega_n} [(2|a_n|^2 + a_n a_{-n} + a_n^* a_{-n}^*)(t)] \right\} \\ &= \sum_{n=-\infty}^{\infty} h\omega_n a_n^*(t) a_n(t) \equiv H \end{aligned} \quad (18.10)$$

and assumes the form of a sum of independent harmonic oscillators as expressed in amplitudes a_n of normal modes. Here we have used the relation

$$\int_0^L dx \phi_n(x, t) \phi_m(x, t) = \delta_{n, -m} e^{-2i\omega_n t}, \quad (18.11)$$

and the fact that $\dot{a}_n(t) = -i\omega_n a_n(t)$, as well as $\omega_n = \omega_{-n}$ and $k_n = -k_{-n}$. Note that all ϕ_n have frequency $\omega_n > 0$. The wave numbers $k_n > 0$ correspond to right propagating waves (so-called right-movers), while $k_n < 0$ correspond to left propagating waves (left-movers). The frequencies $-\omega_n < 0$ belong to $\phi_n^*(x, t)$.

We also introduce the (real) coordinates

$$\begin{aligned} q_n(t) &= \sqrt{\frac{h}{2M\omega_n}} [a_n(t) + a_n^*(t)], \\ p_n(t) &= -i\sqrt{\frac{Mh\omega_n}{2}} [a_n(t) - a_n^*(t)], \\ \rightarrow a_n &= \frac{ip_n + M\omega_n q_n}{\sqrt{2Mh\omega_n}}, \\ \rightarrow a_n^* &= \frac{-ip_n + M\omega_n q_n}{\sqrt{2Mh\omega_n}}, \end{aligned} \quad (18.12)$$

where $M = \mu L$ is the total mass of the system. With these coordinates, the Hamiltonian assumes (again) the form of a sum over independent harmonic oscillators,

$$H = \sum_{n=-\infty}^{\infty} \left(\frac{p_n^2}{2M} + \frac{M\omega_n^2}{2} q_n^2 \right). \quad (18.13)$$

Correspondingly, the equations of motion are those of a harmonic oscillator,

$$\begin{aligned} \dot{q}_n &= \frac{\partial H}{\partial p_n} = \frac{p_n}{M}, \\ \dot{p}_n &= -\frac{\partial H}{\partial q_n} = -M\omega_n^2 q_n, \\ \ddot{q}_n + \omega_n^2 q_n &= 0. \end{aligned} \quad (18.14)$$

Note the three types of coordinates:

$$\begin{aligned} \Phi &= \text{real-valued displacement,} \\ a_n &= \text{complex amplitude of the } n\text{-th normal mode,} \\ q_n &= \text{displacement of the } n\text{-th mode} = q_n(0) \cos(k_n x - \omega_n t). \end{aligned}$$

18.2 Quantization

Starting from the harmonic oscillator in canonical coordinates, as described in (18.13), we can quantize the theory by introducing operators p_n and q_n with commutation relations

$$[q_n, p_m] = i\hbar\delta_{nm}, \quad [q_n, q_m] = 0, \quad [p_n, p_m] = 0. \quad (18.15)$$

The second method of quantizing the theory involves replacing the amplitudes a_n^* and a_n in (18.12) with creation and annihilation operators a_n^\dagger and a_n alongside the substitution $\hbar \rightarrow \hbar$. Their commutation relations can be derived from Eq. (18.15) and the definition (18.12),

$$[a_n, a_m^\dagger] = \delta_{nm}, \quad [a_n, a_m] = 0, \quad [a_n^\dagger, a_m^\dagger] = 0. \quad (18.16)$$

Furthermore, we define the vacuum $|0\rangle$, which is annihilated by the operators a_n , $a_n|0\rangle = 0$. Finally, we can go over to field operators by using the expansion (18.8) in normal modes and finding the corresponding commutators [our time-dependent normal modes define the field operators in the

Heisenberg picture, see Eq (16.120)],

$$\begin{aligned} [\Phi(x, t), \Pi(x', t)] &= i\hbar\delta(x - x'), \\ [\Phi(x, t), \Phi(x', t)] &= 0, \\ [\Pi(x, t), \Pi(x', t)] &= 0, \end{aligned} \tag{18.17}$$

where the conjugated field $\Pi(x, t)$ is, as usual, defined by

$$\begin{aligned} \Pi(x, t) &= \partial\mathcal{L}/\partial\dot{\Phi} = \mu\dot{\Phi} \\ &= -i \sum_{n=-\infty}^{\infty} \left(\frac{\hbar\mu\omega_n}{2} \right)^{1/2} [a_n\phi_n(x, t) - a_n^\dagger\phi_n^*(x, t)]. \end{aligned} \tag{18.18}$$

We have now quantized the string in all three representations. We note, that we actually do not have to trace back the system to the harmonic oscillator in order to be able to quantize it. Alternatively, we can take the following direct route:

Alternative Quantization: Let $\mathcal{L}(\phi, \dot{\phi})$ be the classical Lagrangian density of a system. We consider the field ϕ as a canonical coordinate and define the canonically conjugate momentum variable

$$\pi = \frac{\partial\mathcal{L}}{\partial\dot{\phi}}. \tag{18.19}$$

The theory is quantized by going over to field operators Φ and Π and requiring them to satisfy the commutation relations

$$\begin{aligned} [\Phi(x, t), \Pi(x', t)] &= i\hbar\delta(x - x'), \\ [\Phi(x, t), \Phi(x', t)] &= 0, \\ [\Pi(x, t), \Pi(x', t)] &= 0. \end{aligned} \tag{18.20}$$

The quantization rules via normal modes (18.16) and directly via canonical commutation relations (18.20) are in principle equivalent. Assuming the route via normal modes, however, we are sure to take only independent degrees of freedom into account, which is not always the case for the direct quantization (18.20), as becomes clear when comparing the two approaches in the quantization of the electromagnetic field in the next chapter. The expansion in normal nodes thus is the safest way for a consistent quantization of the theory, albeid not the fastest.

Example: Starting from the single particle-theory of a massive (m) particle in a potential (U), we derive the many-body field-theory of corresponding non-interacting bosons. The Lagrangian density

$$\mathcal{L} = i\hbar\psi^*\dot{\psi} - \frac{\hbar^2}{2m}\vec{\nabla}\psi^* \cdot \vec{\nabla}\psi - U(\vec{r})\psi^*\psi \quad (18.21)$$

for the wave function $\psi(\vec{r})$ generates the Schrödinger equation

$$i\hbar\partial_t\psi = \left(-\frac{\hbar^2}{2m}\nabla^2 + U(\vec{r})\right)\psi. \quad (18.22)$$

We second-quantize this problem by passing from the wave function ψ to the field operator Ψ ; its conjugate variable

$$\pi = \frac{\partial\mathcal{L}}{\partial\dot{\psi}} = i\hbar\psi^* \quad (18.23)$$

defines the creation operator Ψ^\dagger and the commutation relations have the form [see (16.120)],

$$\begin{aligned} [\Psi(\vec{r}, t), \Psi^\dagger(\vec{r}', t)] &= \delta(\vec{r} - \vec{r}'), \\ [\Psi(\vec{r}, t), \Psi(\vec{r}', t)] &= 0, \\ [\Psi^\dagger(\vec{r}, t), \Psi^\dagger(\vec{r}', t)] &= 0. \end{aligned} \quad (18.24)$$

For fermions, one substitutes the commutator $[\cdot, \cdot]$ with the anti-commutator $\{\cdot, \cdot\}$.

Note that we first interpreted the first-quantized one-particle wave mechanics as a classical field theory for the wave field ψ , and then second-quantized this theory. With this procedure, we obtain a many-particle theory with field operators from a one-particle theory. The Hamiltonian can be obtained through a Legendre transform,

$$H = \int d^3r [\Pi\dot{\Psi} - \mathcal{L}] = \int d^3r [\Psi^\dagger \left(-\frac{\hbar^2\nabla^2}{2m} + U\right) \Psi]. \quad (18.25)$$

Occupation number basis

With (18.16), we can introduce the occupation number basis

$$|n_1, \dots, n_k, \dots, n_\infty\rangle = \frac{(a_\infty^\dagger)^{n_\infty}}{\sqrt{n_\infty!}} \dots \frac{(a_k^\dagger)^{n_k}}{\sqrt{n_k!}} \dots \frac{(a_1^\dagger)^{n_1}}{\sqrt{n_1!}} |0\rangle, \quad (18.26)$$

with the vacuum

$$a_k|0\rangle = 0, \quad (18.27)$$

and their interrelations via creation and annihilation operators

$$\begin{aligned} a_k^\dagger |n_1, \dots, n_k, \dots\rangle &= \sqrt{n_k + 1} |n_1, \dots, n_k + 1, \dots\rangle, \\ a_k |n_1, \dots, n_k, \dots\rangle &= \sqrt{n_k} |n_1, \dots, n_k - 1, \dots\rangle, \\ a_k^\dagger a_k |n_1, \dots, n_k, \dots\rangle &= n_k |n_1, \dots, n_k, \dots\rangle. \end{aligned} \quad (18.28)$$

The vibrational state $|n_1, \dots, n_k, \dots\rangle$ of the string contains n_k quanta of the vibration mode $\phi_k(x, t)$. In state $|0\rangle$, we have vacuum vibrations: a string at complete rest would violate Heisenberg's uncertainty principle, as position and momentum would be simultaneously known. In the state $|0, \dots, n_k, \dots, 0\rangle$, the average displacement $\langle n_k | \Phi(x, t) | n_k \rangle$ vanishes at every point, hence displacements are fluctuational,

$$\begin{aligned} \langle n_k | a_n | n_k \rangle &= 0 = \langle n_k | a_n^\dagger | n_k \rangle, \\ \Rightarrow \langle n_k | \Phi(x, t) | n_k \rangle &= 0. \end{aligned} \quad (18.29)$$

However, the expectation value of its square Φ^2 does not vanish; using $\langle n_k | a_n a_m | n_k \rangle = 0 = \langle n_k | a_n^\dagger a_m^\dagger | n_k \rangle$, we find that

$$\begin{aligned} \langle n_k | \Phi^2(x, t) | n_k \rangle &= \langle n_k | \sum_{n, m} \frac{\hbar}{2M\sqrt{\omega_n\omega_m}} \left(\underbrace{a_n a_m^\dagger e^{i(k_n - k_m)x} e^{-i(\omega_n - \omega_m)t}}_{\delta_{nm} + n_k \delta_{nk} \delta_{mk}} \right. \\ &\quad \left. + \underbrace{a_m^\dagger a_n e^{-i(k_n - k_m)x} e^{i(\omega_n - \omega_m)t}}_{n_k \delta_{nk} \delta_{mk}} \right) | n_k \rangle \end{aligned}$$

$$= \frac{\hbar}{M\omega_k} n_k + \underbrace{\sum_n \frac{\hbar}{2M\omega_n}}_{\rightarrow \infty \text{ like } \sum 1/n} \quad (18.30)$$

$$= \sum_n \frac{\hbar}{M\omega_n} \left(\delta_{nk} n_n + \frac{1}{2} \right). \quad (18.31)$$

The result (18.31) includes two terms, a finite mean-squared displacement due to the n_k -fold excitation of the k -mode and a divergent sum due to zero-point fluctuations. This divergence appears both at small and large frequencies—we call them infrared (IR) and ultra-violet (UV) divergencies. We can easily deal with them if we regularize the problem by going back from

the string to a chain. Indeed, the chain has only N degrees of freedom, which leads to a cutoff for large values of n , thus getting rid of the UV divergence in the sum. Furthermore, by fixing one mass in space, the translation mode (or so-called zero-mode) at $n = 0$ is eliminated, which produces a cutoff at small values of n , removing the IR divergence and thus leading to a finite result. For the string, the IR divergence is removed the same way by fixing the string at some point in space; the UV divergence is removed by a physical argument, e.g., every string has a finite thickness d and cannot elastically deform at smaller scale, thus $|k| < \pi/d$. Alternatively, every string is made from something (at least in condensed matter) and thus comes with a microscopic scale l defined by the string's constituents.

For the second term arising from zero-point fluctuations, we can construct an analogy with the thermal excitation of a string at finite temperatures T : following the equipartition theorem, the displacement amplitude of the k -mode is given by the relation $M\omega_n^2|\Phi_k|^2 = k_B T$. Once the thermal energy $k_B T$ drops below the energy $\hbar\omega_n/2$ of zero-point fluctuations, we have to replace the thermal drive $k_B T$ by the quantum energy $\hbar\omega_n/2$. This substitution provides us with the second term in Eq. (18.31)—going to the limit $T \rightarrow 0$, the fluctuating motion of the string does not stop but goes from thermal to quantum. Within the fluctuation-dissipation theorem (FDT), this transition from thermal (at T) to quantum (with frequency ω) fluctuations manifests through a weight factor $\coth(\hbar\omega/k_B T)$.

Discussion²: Calculating the mean-squared displacement $\langle \Phi(x, t) \rangle$, we can learn something about the diffusion (or Brownian motion) of a string. For a point-like particle coupled to a thermal bath one finds (from the FDT) the diffusion law $\langle [\Phi(t) - \Phi(0)]^2 \rangle_T \sim (k_B T / \eta) t$, where η is the dissipative coefficient in the (overdamped) equation of motion. Here, we have two differences with respect to this well known system: i) rather than a particle, here we discuss a string, which is an extended object including many but constrained degrees of freedom. These constraints (e.g., masses bound together by springs) reduce the effectiveness of fluctuations and the string diffuses more slowly, $\langle [\Phi(x, t) - \Phi(x, 0)]^2 \rangle_T \propto \sqrt{t}$; a fluctuating plane would diffuse as $\ln(t)$, while a bulk body does not diffuse at all. ii) Rather than thermal fluctuations, here we discuss quantum fluctuations. It turns out, that quantum fluctuations are less efficient, e.g., the quantum Brownian motion of a particle is only

²You might need some knowledge in (quantum) statistical mechanics to appreciate these statements. In particular, note that we discuss systems (particle, string, membrane, bulk body) in contact with a reservoir that produces a dissipative dynamics, i.e., a term $\eta\dot{\Phi}$ in the (classical) equation of motion

logarithmic, $\langle [\Phi(t) - \Phi(0)]^2 \rangle_{\hbar} \sim (\hbar/\eta) \ln t/t_0$, with t_0 a cutoff. Note, this incoherent diffusion (overdamped motion with η) is different from the coherent ‘diffusion’ of a Gaussian wavepacket $\langle [\Phi(t) - \Phi(0)]^2 \rangle_{\hbar} \sim (\hbar^2/m^2\sigma) t^2$ as described by the Schrödinger equation.

Given the (strong) quantum fluctuations of the string, we can ask the question what state(s) minimize the uncertainty in the position of the string. Obviously, we cannot fix the string at $\Phi(x, t) = f(x)$, because

$$\begin{aligned} \Phi(x, t) = \sum_n \sqrt{\frac{\hbar}{2M\omega_n}} [(a_n + a_n^\dagger) \cos(k_n x - \omega_n t) \\ + i(a_n - a_n^\dagger) \sin(k_n x - \omega_n t)], \end{aligned}$$

and the operators $A_n = (a_n + a_n^\dagger)$ and $B_n = i(a_n - a_n^\dagger)$ cannot be made precise simultaneously, that would require $[A_m, B_n] = 0$ for all m, n . But $[A_n, B_n] = -2i \neq 0$, and thus

$$\begin{aligned} \Delta A_n \Delta B_n &> 1 \quad \text{oder} \\ \Delta q_n \Delta p_n &> \hbar/2. \end{aligned} \tag{18.32}$$

Since we cannot define the field $\Phi(x, t)$ exactly, we would like to have at least optimal states with $\Delta A \simeq \Delta B \simeq 1$, i.e., states with minimal uncertainty. Such states exist and are given by the coherent states.

18.2.1 Coherent States

With $\alpha_i \in \mathbb{C}$, we define the coherent state

$$\begin{aligned} |\vec{\alpha}\rangle &= |\alpha_1, \dots, \alpha_k, \dots\rangle \\ &= \sum_{n_1, \dots, n_k, \dots=0}^{\infty} \frac{(\alpha_1)^{n_1} \dots (\alpha_k)^{n_k} \dots}{\sqrt{n_1! \dots n_k! \dots}} |n_1, \dots, n_k, \dots\rangle \\ &= \sum_{n_1, \dots, n_k, \dots=0}^{\infty} \frac{(\alpha_1 a_1^\dagger)^{n_1} \dots (\alpha_k a_k^\dagger)^{n_k} \dots}{n_1! \dots n_k! \dots} |0\rangle. \end{aligned} \tag{18.33}$$

These states are eigenstates of the annihilation operators a_k with eigenvalues α_k ,

$$a_k |\dots, \alpha_k, \dots\rangle = \alpha_k |\dots, \alpha_k, \dots\rangle. \tag{18.34}$$

We show this for a single mode,

$$a|\alpha\rangle = \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} \sqrt{n} |n-1\rangle = \alpha \underbrace{\sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle}_{|\alpha\rangle}. \quad (18.35)$$

The operator a^\dagger has no eigenstate but satisfies the relation

$$a^\dagger|\alpha\rangle = \frac{d}{d\alpha}|\alpha\rangle, \quad (18.36)$$

instead, since

$$a^\dagger|\alpha\rangle = \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} \sqrt{n+1} |n+1\rangle = \sum_{n=0}^{\infty} n \frac{\alpha^{n-1}}{\sqrt{n!}} |n\rangle = \frac{d}{d\alpha}|\alpha\rangle.$$

The coherent states $|\vec{\alpha}\rangle$ are not normalized, though; with

$$\langle\alpha|\alpha\rangle = \sum_n \frac{(|\alpha|^2)^n}{n!} = e^{\alpha\alpha^*} \quad (18.37)$$

we obtain

$$\langle\vec{\alpha}|\vec{\alpha}\rangle = e^{\sum_k \alpha_k^* \alpha_k}. \quad (18.38)$$

However, they do have the desired property of minimizing $\langle\Delta A\rangle = \langle\Delta B\rangle = 1$, i.e., they carry the smallest possible uncertainty compatible with the Heisenberg uncertainty principle. In this respect, note the following duality:

$$\begin{aligned} n_k \text{ sharp} &\rightarrow \Phi \text{ strongly fluctuates in the state } |0, \dots, n_k, \dots, 0\rangle, \\ n_k \text{ not sharp} &\rightarrow \Phi \text{ has minimal fluctuations in the state } |\alpha\rangle. \end{aligned} \quad (18.39)$$

To prove the above statement, we make use of the relations

$$\begin{aligned} \langle\alpha|a|\alpha\rangle &= \alpha N, \\ \langle\alpha|a^2|\alpha\rangle &= \alpha^2 N, \\ \langle\alpha|a a^\dagger|\alpha\rangle &= (1 + |\alpha|^2)N, \end{aligned} \quad (18.40)$$

as well as

$$\begin{aligned} \langle\alpha|a^\dagger|\alpha\rangle &= \alpha^* N, \\ \langle\alpha|(a^\dagger)^2|\alpha\rangle &= (\alpha^*)^2 N, \\ \langle\alpha|a^\dagger a|\alpha\rangle &= |\alpha|^2 N, \end{aligned} \quad (18.41)$$

where $N = \langle \alpha | \alpha \rangle = \exp(|\alpha|^2)$, and find that

$$\langle A \rangle = \frac{\langle \alpha | A | \alpha \rangle}{\langle \alpha | \alpha \rangle} = (\alpha + \alpha^*), \quad \langle B \rangle = i(\alpha - \alpha^*), \quad (18.42)$$

$$\begin{aligned} \langle A^2 \rangle &= (\alpha^2 + (\alpha^*)^2 + 1 + 2\alpha\alpha^*) = (1 + (\alpha + \alpha^*)^2), \\ \langle B^2 \rangle &= -(\alpha^2 + (\alpha^*)^2 - 1 - 2\alpha\alpha^*) = (1 - (\alpha - \alpha^*)^2), \end{aligned} \quad (18.43)$$

$$\begin{aligned} (\Delta A)^2 &= (1 + (\alpha + \alpha^*)^2 - (\alpha + \alpha^*)^2) = 1, \\ (\Delta B)^2 &= (1 - (\alpha - \alpha^*)^2 + (\alpha - \alpha^*)^2) = 1. \end{aligned} \quad (18.44)$$

Note that $\langle a^\dagger a \rangle = |\alpha|^2$, therefore the square modulus $|\alpha|^2$ of the normalized state α provides us with the number of excitation quanta in this state. The relative uncertainties are

$$\begin{aligned} \frac{\langle \Delta A \rangle}{\langle A \rangle} &= \frac{1}{\alpha + \alpha^*} = \frac{1}{2|\alpha| \cos \varphi}, \\ \frac{\langle \Delta B \rangle}{\langle B \rangle} &= \frac{-1}{2|\alpha| \sin \varphi}, \end{aligned} \quad (18.45)$$

where $\alpha = |\alpha| \exp(i\varphi)$. Hence, the (relative) uncertainties in position and momentum of the string disappear for $|\alpha| = \sqrt{\# \text{ quanta}} \rightarrow \infty$. This statement tells us that the classical limit of a bosonic field theory corresponds to the limit of large boson number, e.g., the classical elastic theory emerges for a large number of phonons and classical electromagnetism is the result of a large photon number.

18.3 Path integrals

In Chapter 1 of Quantum Mechanics I, we introduced the quantum mechanics of a single particle with the help of path integrals. The basic idea in the transition from classical to quantum mechanics was that all paths $x(t)$ (rather than a single one) contribute with their weight $\exp(iS[x(t)]/\hbar)$ to the propagator, where $S[x(t)]$ is the action accumulated on the path $x(t)$. The classical path with $\delta S = 0$ adds the largest contribution to the sum over all paths. In the following, we show that this formalism also applies to many-body systems. We will see that, instead of working with operators, within the path integral formalism we deal with integrals over complex-valued functions. We give a derivation for Bose-systems on a concrete example, the 1D string; the formalism is more complicated for fermions because we need Grassmann-variables. For comparison, we perform the transition to path

integrals for a bosonic many-body system alongside with the discussion of a one-particle problem.

18.3.1 Preparation

We work in the Heisenberg picture, where operators are time-dependent while states are not,

Schrödinger picture: Ops time indep. \leftrightarrow states time dep.
 Heisenberg picture: Ops time dep. \leftrightarrow states time indep.

Let q be a generalized coordinate with $Q_H(t)$ the corresponding (time-dependent) operator in the Heisenberg picture. The *eigenstates* of $Q_H(t)$ then are also time-dependent and we denote them by $|q, t\rangle_H$. Note that this is different from the standard statement that states are time-independent in the Heisenberg picture; this difference is owed to that fact that $|q, t\rangle_H$ is an eigenstate rather than a state.

With $U(t, t_0) (= \exp[-iH(t - t_0)/\hbar])$ the (simplest) time-evolution operator (for a situation where $\partial_t H = 0$), we obtain the eigenstate $|q\rangle_S$ belonging to Q in the Schrödinger picture

$$|q\rangle_S = U(t, t_0)|q, t\rangle_H \quad (18.46)$$

In order to familiarize us with these operators and eigenstates, we consider an example with $t_0 = 0$ and $\hbar = 1$, then

$$\begin{aligned} \text{Schrödinger:} \quad Q|q\rangle_S &= q|q\rangle_S, \\ |q, t\rangle_H &= \exp(iHt)|q\rangle_S \\ \text{Heisenberg:} \quad Q_H(t) &= \exp(iHt)Q\exp(-iHt), \\ Q_H(t)|q, t\rangle_H &= \exp(iHt)Q \underbrace{\exp(-iHt)\exp(iHt)}_{\mathbb{I}}|q\rangle_S \\ &= q\exp(iHt)|q\rangle_S \\ &= q|q, t\rangle_H. \end{aligned}$$

A state $|m\rangle_H$ in the Heisenberg picture is time independent. The associated time dependent wave function in coordinate representation is

$$\begin{aligned} \Psi_m(q, t) &= {}_S\langle q | e^{-iHt} | m \rangle_H \\ &= {}_H\langle q, t | m \rangle_H \\ &= {}_S\langle q | m, t \rangle_S, \end{aligned} \quad (18.47)$$

with $|m, t\rangle_S = \exp(-iHt)|m\rangle_H$, defined as usual. In the following, we only write $|q, t\rangle$ without the index H .

Coherent states Moving now to the path integral formalism for the bosonic many-body problem, we base its derivation on the coherent states,

$$|\vec{\alpha}\rangle = \prod_j \sum_{n_j} \frac{(\alpha_j a_j^\dagger)^{n_j}}{n_j!} |0\rangle. \quad (18.48)$$

These are right (left) eigenstates of the operators a_i (a_i^\dagger),

$$\begin{aligned} a_i |\vec{\alpha}\rangle &= \alpha_i |\vec{\alpha}\rangle, & a_i^\dagger |\vec{\alpha}\rangle &= \partial_{\alpha_i} |\vec{\alpha}\rangle, \\ \langle \vec{\alpha} | a_i^\dagger &= \langle \vec{\alpha} | \alpha_i^*, & \langle \vec{\alpha} | a_i &= \partial_{\alpha_i^*} \langle \vec{\alpha} |, \\ \langle \vec{\alpha} | \vec{\alpha} \rangle &= \exp(\vec{\alpha}^* \cdot \vec{\alpha}). \end{aligned} \quad (18.49)$$

We will use the states $|\vec{\alpha}\rangle$ in analogy to the position states $|q\rangle$ in the one-particle problem, in particular, we will find the correspondence (see (18.59) and (18.64) below)

$$\alpha_j^* \leftrightarrow q_j. \quad (18.50)$$

Note that the field operators $\Phi(x)$ do not have eigenstates, but the annihilation operators a_j do. This is why we build the path integral formalism on the operators a_j instead of the field operators $\Phi(x)$. In the end, we can use the complex numbers $\vec{\alpha}$, $\vec{\alpha}^*$ to convert the result back to complex-valued fields Φ and Π .

Below, we make use of an important property of coherent states when building matrix elements of operators that are ‘normal ordered’:

Normal ordering We call the operator $:O(a_i^\dagger, a_j):$ normal ordered, if all of the annihilation operators a_j are to the right of the creation operators a_i^\dagger . For an arbitrary operator $O(a_i^\dagger, a_j)$, we write the associated normal ordered operator $:O(a_i^\dagger, a_j):$. We then have

$$\langle 0 | :O(a_i^\dagger, a_j): | 0 \rangle = 0. \quad (18.51)$$

Let $O(a_i^\dagger, a_j)$ be normal ordered, then

$$\begin{aligned} \langle \vec{\alpha} | O(a_i^\dagger, a_j) | \vec{\alpha}' \rangle &= O(\alpha_i^*, \alpha'_j) \langle \vec{\alpha} | \vec{\alpha}' \rangle \\ &= O(\alpha_i^*, \alpha'_j) e^{\vec{\alpha}^* \cdot \vec{\alpha}'}. \end{aligned} \quad (18.52)$$

Completeness Furthermore, the coherent states $|\vec{\alpha}\rangle$ are complete in the following sense

$$\mathbb{1} = \int \prod_i \frac{d\alpha_i d\alpha_i^*}{2\pi i} e^{-\vec{\alpha}^* \cdot \vec{\alpha}} |\vec{\alpha}\rangle \langle \vec{\alpha}|. \quad (18.53)$$

Proof: The proof of (18.53) is split into two parts:

a) $\int \prod \dots$ commutes with all a_i, a_i^\dagger and it follows that $\int \prod \dots = c \cdot \mathbb{1}$.

b) $c = 1$.

a) is shown by explicit calculation of the commutators

$$\begin{aligned} [a_i, \int \prod_i \frac{d\alpha_i d\alpha_i^*}{2\pi i} e^{-\vec{\alpha}^* \cdot \vec{\alpha}} |\vec{\alpha}\rangle \langle \vec{\alpha}|] \\ = \int \prod_i \frac{d\alpha_i d\alpha_i^*}{2\pi i} e^{-\vec{\alpha}^* \cdot \vec{\alpha}} \left(\underbrace{a_i |\vec{\alpha}\rangle \langle \vec{\alpha}|}_{\alpha_i |\vec{\alpha}\rangle} + |\vec{\alpha}\rangle \underbrace{\langle \vec{\alpha}| a_i}_{\partial_{\alpha_i^*} \langle \vec{\alpha}|} \right) \\ = \int \prod_i \frac{d\alpha_i d\alpha_i^*}{2\pi i} e^{-\vec{\alpha}^* \cdot \vec{\alpha}} (\alpha_i - \partial_{\alpha_i^*}) |\vec{\alpha}\rangle \langle \vec{\alpha}| = 0, \end{aligned} \quad (18.54)$$

where the derivative $\partial_{\alpha_i^*}$ generates a term α_i via partial integration which compensates for the already existing term. Note, that $(\partial/\partial\alpha^*)f(\alpha) = f'(\alpha)(\partial\alpha/\partial\alpha^*) = 0$, and $\partial_{\alpha^*} = (\partial_{\alpha_x} + i\partial_{\alpha_y})/2$. The same calculation can also be carried out for a_i^\dagger .

For b), we calculate the expectation value

$$\int \prod_i \frac{d\alpha_i d\alpha_i^*}{2\pi i} e^{-\vec{\alpha}^* \cdot \vec{\alpha}} \langle 0|\vec{\alpha}\rangle \langle \vec{\alpha}|0\rangle = \int \prod_i \frac{d\alpha_i d\alpha_i^*}{2\pi i} e^{-\vec{\alpha}^* \cdot \vec{\alpha}}.$$

We interpret the measure $d\alpha_i d\alpha_i^*$ as a skew product $d\alpha \wedge d\alpha^*$. With the substitution $\alpha = r \exp(i\theta)$, $\alpha^* = r \exp(-i\theta)$, we find its areal measure

$$d\alpha d\alpha^* = \begin{vmatrix} \partial_\theta \alpha & \partial_r \alpha \\ \partial_\theta \alpha^* & \partial_r \alpha^* \end{vmatrix} dr d\theta \quad (18.55)$$

$$= \begin{vmatrix} ire^{i\theta} & e^{i\theta} \\ -ire^{-i\theta} & e^{-i\theta} \end{vmatrix} dr d\theta = 2ir dr d\theta. \quad (18.56)$$

The remaining integral then is easily calculated,

$$\begin{aligned} \int \frac{d\alpha d\alpha^*}{2\pi i} e^{-\alpha\alpha^*} &= \int \frac{r dr d\theta}{\pi} e^{-r^2} = 1. \\ \Rightarrow c &= 1. \end{aligned} \quad (18.57)$$

□

Finally, in analogy to the one-particle problem, we need the eigenstates of the annihilation operator $a_i(t)$ in the Heisenberg picture,

$$|\vec{\alpha}, t\rangle_H \quad \text{with} \quad a_i(t)|\vec{\alpha}, t\rangle_H = \alpha_i|\vec{\alpha}, t\rangle_H. \quad (18.58)$$

A state $|v\rangle$ in the coherent-state basis is then described by the wave function

$$\begin{aligned} \Psi_v(\vec{\alpha}^*, t) &= {}_H\langle\vec{\alpha}, t|v\rangle_H \\ &= {}_S\langle\vec{\alpha}|v, t\rangle_S, \end{aligned} \quad (18.59)$$

where

$$|v, t\rangle_S = e^{-iHt}|v\rangle_H. \quad (18.60)$$

Again, we omit the index H in the following.

We now develop the path integral formalism in parallel for a particle (with coordinate q) and for a system of bosons (with coherent-state complex coordinates $\vec{\alpha}$). For convenience, we introduce the abbreviations

$$\begin{aligned} \int [d\alpha^*] &\equiv \int \prod_j d\alpha_j^*, \\ \int [d\alpha] &\equiv \int \prod_j \frac{d\alpha_j}{2\pi i}, \end{aligned} \quad (18.61)$$

and set $\hbar = 1$.

18.3.2 Path integrals: particles versus bosons

Particles, are described by the wave function $\Psi_m(q, t) = \langle q, t|m\rangle$. The propagator K can be derived by inserting a unit operator $\mathbb{1} = \int dq_0 |q_0, t_0\rangle\langle q_0, t_0|$,

$$\begin{aligned} \Psi_m(q_n, t_n) &= \int dq_0 \underbrace{\langle q_n, t_n|q_0, t_0\rangle}_K \underbrace{\langle q_0, t_0|m\rangle}_\Psi \\ &= \int dq_0 K(q_n, t_n; q_0, t_0) \Psi(q_0, t_0) \end{aligned} \quad (18.62)$$

$$K(q_n, t_n; q_0, t_0) = \langle q_n, t_n|q_0, t_0\rangle. \quad (18.63)$$

Bosons, are described by the wave function $\Psi_v(\vec{\alpha}^*, t) = \langle \vec{\alpha}, t | v \rangle$. Note the correspondence between $\Psi_m(q, t) = \langle q, t | m \rangle$ and $\Psi_v(\vec{\alpha}^*, t) = \langle \vec{\alpha}, t | v \rangle$; accordingly, $q \leftrightarrow \alpha^*$ correspond to each other. The propagator U can again be found by inserting $\mathbb{1}$,

$$\begin{aligned} \Psi_v(\vec{\alpha}_n^*, t_n) &= \int \prod_i \frac{d\alpha_{i0} d\alpha_{i0}^*}{2\pi i} e^{-\vec{\alpha}_0^* \cdot \vec{\alpha}_0} \langle \vec{\alpha}_n, t_n | \vec{\alpha}_0, t_0 \rangle \langle \vec{\alpha}_0, t_0 | v \rangle \\ &= \int [d\alpha_0^*] U(\vec{\alpha}_n^*, t_n; \vec{\alpha}_0^*, t_0) \Psi_v(\vec{\alpha}_0^*, t_0), \end{aligned} \quad (18.64)$$

$$U(\vec{\alpha}_n^*, t_n; \vec{\alpha}_0^*, t_0) = \int [d\alpha_0] e^{-\vec{\alpha}_0^* \cdot \vec{\alpha}_0} \langle \vec{\alpha}_n, t_n | \vec{\alpha}_0, t_0 \rangle. \quad (18.65)$$

For particles as well as for Bosons, the time interval $[t_0, t_n]$ can be divided into n intervals $[t_i, t_{i+1}]$, $i = 0, 1, \dots, n-1$ with $t_{i+1} - t_i = \varepsilon \rightarrow 0$. For each time step, a $\mathbb{1}$ is inserted, and one finds for K and U ,

$$K(q_n, t_n; q_0, t_0) = \int \left[\prod_{i=1}^{n-1} dq_i \right] \left[\prod_{i=0}^{n-1} \langle q_{i+1}, t_{i+1} | q_i, t_i \rangle \right], \quad (18.66)$$

$$U(\vec{\alpha}_n^*, t_n; \vec{\alpha}_0^*, t_0) = \int \left[\prod_{i=1}^{n-1} [d\alpha_i^*] \right] \left[\prod_{i=0}^{n-1} [d\alpha_i] e^{-\vec{\alpha}_i^* \cdot \vec{\alpha}_i} \langle \vec{\alpha}_{i+1}, t_{i+1} | \vec{\alpha}_i, t_i \rangle \right].$$

The matrix elements $\langle \cdot | \cdot \rangle$ are calculated under the assumption that the Hamiltonian is constant over the small ε -intervals $[t_i, t_{i+1}]$ for both particles as well as bosonic modes,

$$\begin{aligned} \langle q_{i+1}, t_{i+1} | q_i, t_i \rangle &= \langle q_{i+1} | U(t_{i+1}, t_0) U^\dagger(t_i, t_0) | q_i \rangle \\ &\cong \langle q_{i+1} | e^{-iH\varepsilon} | q_i \rangle \cong \langle q_{i+1} | 1 - iH\varepsilon | q_i \rangle \\ &= \delta(q_{i+1} - q_i) - i\varepsilon \langle q_{i+1} | H | q_i \rangle \\ &= \int \frac{dp}{2\pi} e^{ip(q_{i+1} - q_i)} - i\varepsilon \langle q_{i+1} | P^2/2m + V(q_i) | q_i \rangle \\ &= \int \frac{dp}{2\pi} e^{ip(q_{i+1} - q_i)} [1 - i\varepsilon V(q_i)] \\ &\quad - i\varepsilon \int \frac{dp}{2\pi} \int \frac{dp'}{2\pi} \langle q_{i+1} | p' \rangle \langle p' | P^2/2m | p \rangle \langle p | q_i \rangle \\ &= \int \frac{dp}{2\pi} e^{ip(q_{i+1} - q_i)} \{1 - i\varepsilon [p^2/2m + V(q_i)]\} \\ &\cong \int \frac{dp}{2\pi} e^{i[p(q_{i+1} - q_i) - \varepsilon H(p, q_i)]} \end{aligned}$$

$$= \int \frac{dp_i}{2\pi} e^{i\varepsilon[p_i \dot{q}_i - H(p_i, q_i)]}, \quad (18.67)$$

$$\begin{aligned} \langle \vec{\alpha}_{i+1}, t_{i+1} | \vec{\alpha}_i, t_i \rangle &= \langle \vec{\alpha}_{i+1} | U(t_{i+1}, t_0) U^\dagger(t_i, t_0) | \vec{\alpha}_i \rangle \\ &\cong \langle \vec{\alpha}_{i+1} | \vec{\alpha}_i \rangle - i\varepsilon \langle \vec{\alpha}_{i+1} | H(\vec{a}^\dagger, \vec{a}) | \vec{\alpha}_i \rangle \\ &\cong e^{\vec{\alpha}_{i+1}^* \cdot \vec{\alpha}_i} [1 - i\varepsilon H(\vec{\alpha}_{i+1}^*, \vec{\alpha}_i)] \\ &\cong e^{\vec{\alpha}_{i+1}^* \cdot \vec{\alpha}_i - i\varepsilon H(\vec{\alpha}_{i+1}^*, \vec{\alpha}_i)}. \end{aligned} \quad (18.68)$$

In (18.68), we used the condition that the coherent states are eigenstates of the annihilation operators a_k and that the Hamiltonian is normal ordered. We insert these matrix elements into the expression (18.66) for the propagators K and U and obtain

$$\begin{aligned} K(q_n, t_n; q_0, t_0) &= \int \left[\prod_{i=1}^{n-1} dq_i \right] \left[\prod_{i=0}^{n-1} \frac{dp_i}{2\pi} e^{i\varepsilon[p_i \dot{q}_i - H(p_i, q_i)]} \right] \\ &= \int \frac{\mathcal{D}[q(t)] \mathcal{D}[p(t)]}{2\pi\hbar} \exp\left(\frac{i}{\hbar} \int_{t_0}^{t_n} dt [p\dot{q} - H(p, q)]\right), \end{aligned} \quad (18.69)$$

$$\begin{aligned} U(\vec{\alpha}_n^*, t_n; \vec{\alpha}_0^*, t_0) &= \int \left[\prod_{i=1}^{n-1} [d\alpha_i^*] \right] \left[\prod_{i=0}^{n-1} [d\alpha_i] e^{[(\vec{\alpha}_{i+1}^* - \vec{\alpha}_i^*) \cdot \vec{\alpha}_i - i\varepsilon H(\vec{\alpha}_{i+1}^*, \vec{\alpha}_i)]} \right] \\ &\downarrow \vec{\alpha} \rightarrow i\vec{p}, \quad \vec{\alpha}^* \rightarrow \vec{q}, \quad [dq_i] = \prod_j dq_{j,i}, \quad [dp_i] = \prod_j \frac{dp_{j,i}}{2\pi}, \\ &= \int \left[\prod_{i=1}^{n-1} [dq_i] \right] \left[\prod_0^{n-1} [dp_i] e^{i\varepsilon[\vec{q}_i \cdot \vec{p}_i - H(\vec{q}_{i+1}, \vec{p}_i)]} \right] \\ &= \int \frac{\mathcal{D}[\Phi(x, t)] \mathcal{D}[\Pi(x, t)]}{2\pi\hbar} \exp\left(\frac{i}{\hbar} \int_{t_0}^{t_n} dt \int_0^L dx [\dot{\Phi}\Pi - H(\Phi, \Pi)]\right), \end{aligned} \quad (18.70)$$

where we made replacement $a_j \rightarrow \alpha_j \rightarrow ip_j$ and $a_j^\dagger \rightarrow \alpha_j^* \rightarrow q_j$, see also (18.50).³ We also reintroduced Planck's constant in the final formulae (18.69) and (18.70).

In the last step, we transitioned to the the complex-valued fields Φ and Π

³The variables $q_j(t)$ and $p_j(t)$ here should not be identified with the real coordinates in (18.12).

as follows:⁴

$$\begin{aligned}\Phi(x, t) &= \sum_{j=-\infty}^{\infty} \sqrt{\frac{\hbar}{2\mu\omega_j}} [ip_j(t)\phi_j(x) + q_j(t)\phi_j^*(x)], \\ \Pi(x, t) &= \sum_{j=-\infty}^{\infty} \sqrt{\frac{\hbar\mu\omega_j}{2}} [p_j(t)\phi_j(x) + iq_j(t)\phi_j^*(x)],\end{aligned}\quad (18.71)$$

We introduce a ‘coarse graining’ over the length l : using

$$\phi_{j,x} = \frac{1}{\sqrt{l}} \int_{l_x} dx' \phi_j(x') \quad (18.72)$$

we define

$$\begin{aligned}\Phi(x, t) \rightarrow \Phi_x(t) &= \sum_j \sqrt{\frac{\hbar}{2\mu\omega_j}} [ip_j(t)\phi_{j,x} + q_j(t)\phi_{j,x}^*], \\ \Pi(x, t) \rightarrow \Pi_x(t) &= \sum_j \sqrt{\frac{\hbar\mu\omega_j}{2}} [p_j(t)\phi_{j,x} + iq_j(t)\phi_{j,x}^*],\end{aligned}\quad (18.73)$$

resulting in a countable number of degrees of freedom for the fields. With (18.71) and (18.73), we go from the modes with index j to the positions x , where x counts the intervals of length l ; this corresponds to a canonical transformation $p_j, q_j \leftrightarrow \Pi_x, \Phi_x$. Finally, we transform the measure of integration

$$\prod_{i,j} \frac{dq_{j,i} dp_{j,i}}{2\pi} \longrightarrow \prod_{i,x} \frac{d\Phi_{x,i} d\Pi_{x,i}}{2\pi\hbar} = \frac{\mathcal{D}[\Phi(x, t)] \mathcal{D}[\Pi(x, t)]}{2\pi\hbar}, \quad (18.74)$$

where the index i goes over time, j goes over modes, and x goes over positions. At the same time, the exponent is re-written according to

$$\begin{aligned}\prod_i \exp[i\varepsilon(\dot{\vec{q}}_i \cdot \vec{p}_i - H)] &= \exp\left(i\varepsilon \sum_{ij} [\dot{q}_{ij} p_{ij} - H(q_{ij}, p_{ij})]\right) \\ &= \exp\left(i\varepsilon \sum_{i,x} [\dot{\Phi}_{i,x} \Pi_{i,x} - H(\Phi_{i,x}, \Pi_{i,x})]\right), \\ &= \exp\left(i \int dt dx [\dot{\Phi}(x, t) \Pi(x, t) - H(\Phi(x, t), \Pi(x, t))]\right),\end{aligned}\quad (18.75)$$

⁴The transition from α -coordinates to real-valued fields Φ and Π can also be done directly using Eqs. (18.8) and (18.18).

where we formally replaced the discrete sum over time and position i, x with amplitude $\Phi_{x,i}$ and $\Pi_{x,i}$ by continuous time and position integrals over the fields $\Phi(x, t)$ and $\Pi(x, t)$ (note, that $\Phi_{x,i}$ and $\Pi_{x,i}$ have an additional factor $1/\sqrt{l}$ compared to $\Phi(x, t)$ and $\Pi(x, t)$, see (18.72)). As usual, the integrals $\int \mathcal{D}[\Phi(x, t)] \mathcal{D}[\Pi(x, t)]$ are formal objects—in a real calculation, these have to be replaced by the discrete version $\int \prod_{i,x} d\Phi_{x,i} d\Pi_{x,i}$.

18.3.3 Remarks

1. Of course (18.69) for one particle is identical to our result from the QM I lecture; show this by integrating over p (complete the square) and find Feynman's measure A .
2. For fermions, we need the anti-commutation properties of the Fermi field operators. This means that the classical \mathbb{C} -valued fields in (18.70) should have anti-commutation properties. The mathematics of anti-commuting \mathbb{C} -numbers was developed in the 19th century by the mathematician Grassman, so-called Grassman algebras, but for time reasons, we will not discuss it here. This topic can, e.g., be looked up in Negele-Orland's book, 'Quantum Many Particle Systems'.
3. Why are path integrals useful? (18.70) is probably the nicest way to construct the quantum mechanics of a particle or a field. The final formula for the path integral is simple and physically transparent. Many papers start with (18.70)—one then should be able to read these first few lines. Specifically, the path integral is helpful (18.70)
 - in finding Feynman rules (perturbation theory),
 - in finding exact (numerical) solutions on grids,
 - in the transition from quantum mechanics to statistical mechanics, quantum statistical physics, via Wick rotation (imaginary time formalism),

$$t \rightarrow i\tau, \quad \int_0^t dt' \rightarrow \int_{-\hbar/2T}^{\hbar/2T} d\tau, \quad (18.76)$$

- in performing abstract calculations in field theory.

Chapter 19

Quantization of electromagnetism: photons

More material on this topic can be found in the books of Jackson, Sakurai, and Gross.

The quantization of the electromagnetic field requires a careful discussion of the true dynamical degrees of freedom in the theory. Starting with 6 fields for the electric (\vec{E}) and magnetic (\vec{B}) fields, we have to take into account the Maxwell equations and the gauge invariance. Going over to potentials¹ with the four-vector (ϕ, \vec{A}) , we have 4 fields remaining. Another degree of freedom vanishes by fixing the gauge, which leads to 3 remaining fields². When choosing the Coulomb gauge, one of the fields (ϕ) is non-dynamic, which results in two dynamical (i.e., fluctuating) degrees of freedom. These two fields will give us the photons with spin ± 1 (along the direction of propagation \hat{k}). The spin-0 photon does not exist because the photon has no mass.

19.1 Maxwell equations

$$\vec{\nabla} \cdot \vec{B} = 0, \quad \text{no magnetic monopoles,} \quad (19.1)$$

¹The homogeneous equations $\nabla \cdot \vec{B} = 0$ and $\vec{\nabla} \wedge \vec{E} = -\partial_t \vec{B}/c$ are satisfied by $\vec{B} = \vec{\nabla} \wedge \vec{A}$ and $\vec{E} = -\vec{\nabla} \phi - \partial_t \vec{A}/c$.

²Otherwise, we would keep an artificial gauge degree of freedom that would, e.g., imply overestimating fluctuations.

$$\vec{\nabla} \wedge \vec{E} = -\frac{1}{c} \partial_t \vec{B}, \quad \text{Faraday, induction,} \quad (19.2)$$

$$\vec{\nabla} \cdot \vec{E} = 4\pi\rho, \quad \text{Coulomb, charge sources,} \quad (19.3)$$

$$\vec{\nabla} \wedge \vec{B} = \frac{1}{c} \partial_t \vec{E} + \frac{4\pi}{c} \vec{j}, \quad \text{Ampère, current sources.} \quad (19.4)$$

Equations (19.1) and (19.2) are the homogeneous, (19.3) and (19.4) the inhomogeneous Maxwell equations. The charge density ρ and the current density \vec{j} are sources of the electromagnetic fields; we choose them independent of the fields \vec{E} and \vec{B} .³ We integrate the homogeneous equations by introducing the potentials ϕ and \vec{A} , with ϕ a scalar field (potential) and \vec{A} a vector field (gauge field). By expressing the fields \vec{E} and \vec{B} through the potentials,

$$\vec{B} = \vec{\nabla} \wedge \vec{A}, \quad (19.5)$$

$$\vec{E} = -\frac{1}{c} \partial_t \vec{A} - \vec{\nabla} \phi, \quad (19.6)$$

we automatically satisfy the homogeneous equations (19.1) and (19.2).

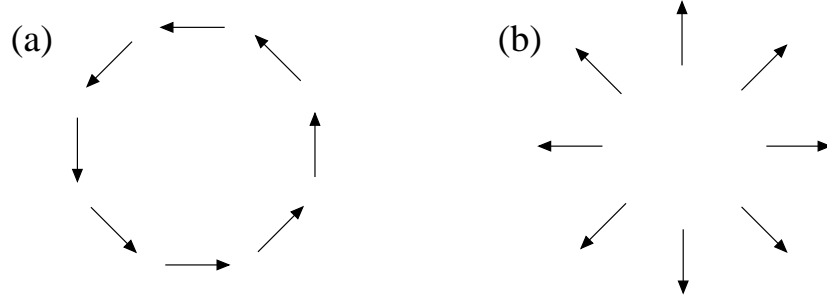


Fig. 19.1: (a) Pure rotation field $\vec{B} = \vec{\nabla} \wedge \vec{A}$ with no divergence $\vec{\nabla} \cdot \vec{B} = 0$. (b) Pure divergence field $\vec{E} = \vec{\nabla} \phi$ with no rotation $\vec{\nabla} \wedge \vec{E} = 0$.

The magnetic field $\vec{B} = \vec{\nabla} \wedge \vec{A}$ is a pure rotation field, $\vec{\nabla} \wedge \vec{B} = (\partial_t \vec{E} + 4\pi \vec{j})/c$, as sketched in Fig. 19.1(a), while the electric field \vec{E} has both divergent, $\vec{\nabla} \cdot \vec{E} = 4\pi\rho$ [as sketched in Fig. 19.1(b)] as well as rotating components, $\vec{\nabla} \wedge \vec{E} = -(1/c) \partial_t \vec{B}$. Magnetic fields and dynamic electric fields thus describe purely transverse physics; longitudinal physics is generated by charges.

³In the presence of matter, $\vec{j} = \vec{j}(\vec{E}, \vec{B})$ and $\rho = \rho(\vec{E}, \vec{B})$ depend on the fields and we need additional constitutive equations, e.g., in the metal, Ohm's law tells that $\vec{j} = \vec{\sigma}(\vec{B}) \vec{E}$, $\rho \cong 0$ with the tensor of conductivities $\vec{\sigma}$. In a superconductor, the London equation reads $\vec{j} = -(c/4\pi\lambda^2) \vec{A}$ with λ denoting the (Ginzburg-Landau) penetration depth.

Inserting the fields (19.5) and (19.6) into the inhomogeneous Maxwell equations provides us with the equations for the potentials,

$$\begin{aligned} \nabla^2 \phi + \frac{1}{c} \partial_t \vec{\nabla} \cdot \vec{A} &= -4\pi\rho, \\ \underbrace{\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{A})}_{-\nabla^2 \vec{A} + \vec{\nabla}(\vec{\nabla} \cdot \vec{A})} + \frac{1}{c^2} \partial_t^2 \vec{A} + \frac{1}{c} \vec{\nabla} \partial_t \phi &= \frac{4\pi}{c} \vec{j}. \end{aligned} \quad (19.7)$$

The consistency of the equations (19.7) enforces the continuity equation for the charge and current fields (we take ∂_t of the first, $\vec{\nabla} \cdot$ of the second equation, subtract, and use the relation $\vec{\nabla} \cdot (\vec{\nabla} \wedge \cdots) = 0$),

$$0 = \vec{\nabla} \cdot (\vec{\nabla} \wedge (\vec{\nabla} \wedge \vec{A})) = \frac{4\pi}{c} (\partial_t \rho + \vec{\nabla} \cdot \vec{j}). \quad (19.8)$$

Furthermore, we have a gauge freedom in the choice of \vec{A} and ϕ : changing the gauge via the scalar gauge field $\chi(\vec{r}, t)$,

$$\begin{aligned} \vec{A}' &= \vec{A} - \vec{\nabla} \chi, \\ \phi' &= \phi + \frac{1}{c} \partial_t \chi, \end{aligned} \quad (19.9)$$

we obtain the same physical fields \vec{E} and \vec{B} ,

$$\begin{aligned} \vec{\nabla} \wedge \vec{A}' &= \vec{\nabla} \wedge \vec{A} = \vec{B}, \\ -\frac{1}{c} \partial_t \vec{A}' - \vec{\nabla} \phi' &= -\frac{1}{c} \partial_t \vec{A} - \vec{\nabla} \phi = \vec{E}. \end{aligned} \quad (19.10)$$

There are no real physical degrees of freedom associated with the field $\chi(\vec{r}, t)$ and we have to fix the gauge (otherwise we would introduce non-physical fluctuations in the gauge field). Typical choices for a gauge are the⁴

- Coulomb gauge: $\vec{\nabla} \cdot \vec{A} = 0$, and the
- Lorenz gauge: $\vec{\nabla} \cdot \vec{A} + (1/c) \partial_t \phi = 0$.

By solving the Laplace problem (\rightarrow Coulomb gauge)

$$\nabla^2 \chi = \vec{\nabla} \cdot \vec{A} \quad (19.11)$$

⁴Ludvig Lorenz, not to be confused with the Hendric Antoon Lorentz; the latter is associated with the theory of the electron, the Lorentz transformation (1904), and the Lorentz-Lorentz formula (Classic-Mosotti). The Lorenz gauge is named after Ludvig Lorenz.

or the wave equation (\rightarrow Lorenz-gauge)

$$\left(\frac{1}{c^2} \partial_t^2 - \nabla^2\right) \chi = \vec{\nabla} \cdot \vec{A} + \frac{1}{c} \partial_t \phi \quad (19.12)$$

we can always switch between the Coulomb and Lorenz gauge.

In the Coulomb-gauge, $\vec{\nabla} \cdot \vec{A} = 0$, and the Coulomb field $\phi(\vec{r}, t)$ is instantaneous,

$$\begin{aligned} \nabla^2 \phi(\vec{r}, t) &= -4\pi \rho(\vec{r}, t), \\ \Rightarrow \phi(\vec{r}, t) &= \int d^3 r' \frac{\rho(\vec{r}', t)}{|\vec{r} - \vec{r}'|}. \end{aligned} \quad (19.13)$$

The (transverse) vector field \vec{A} satisfies the dynamic equation (see (19.7b) and use $\vec{\nabla} \cdot \vec{A} = 0$)

$$\nabla^2 \vec{A} - \frac{1}{c^2} \partial_t^2 \vec{A} = -\frac{4\pi}{c} \vec{j} + \frac{1}{c} \vec{\nabla} \partial_t \phi = -\frac{4\pi}{c} \vec{j}_\perp. \quad (19.14)$$

Indeed, it is easy to check that the source term in (19.14) is purely transverse,

$$\vec{\nabla} \cdot \vec{j}_\perp = \vec{\nabla} \cdot \vec{j} - \nabla^2 \partial_t \phi / 4\pi = (\vec{\nabla} \cdot \vec{j} + \partial_t \rho) = 0. \quad (19.15)$$

By expressing the scalar potential ϕ through the charge density [see (19.13)] and using the relation $\Delta(1/|\vec{r} - \vec{r}'|) = -4\pi\delta(\vec{r} - \vec{r}')$, we can write the source term in the explicitly transverse form⁵

$$\begin{aligned} \vec{j}_\perp(\vec{r}, t) &= \vec{j} - \vec{\nabla} \partial_t \phi / 4\pi \\ &= -\nabla^2 \int d^3 r' \frac{\vec{j}(\vec{r}', t)}{4\pi|\vec{r} - \vec{r}'|} - \vec{\nabla} \int d^3 r' \frac{-\vec{\nabla}' \cdot \vec{j}(\vec{r}', t)}{4\pi|\vec{r} - \vec{r}'|} \\ &= \underbrace{\left(-\nabla^2 + \vec{\nabla} \vec{\nabla} \cdot\right)}_{\vec{\nabla} \wedge \vec{\nabla} \wedge} \int d^3 r' \frac{\vec{j}(\vec{r}', t)}{4\pi|\vec{r} - \vec{r}'|} \\ &= \vec{\nabla} \wedge \left(\vec{\nabla} \wedge \int d^3 r' \frac{\vec{j}(\vec{r}', t)}{4\pi|\vec{r} - \vec{r}'|} \right) \end{aligned} \quad (19.16)$$

The equation for the vector potential \vec{A} then can be written in the form

$$\nabla^2 \vec{A} - \frac{1}{c^2} \partial_t^2 \vec{A} = -\frac{1}{c} \vec{\nabla} \wedge \left(\vec{\nabla} \wedge \int d^3 r' \frac{\vec{j}(\vec{r}', t)}{|\vec{r} - \vec{r}'|} \right). \quad (19.17)$$

⁵Note, that $\vec{n} \wedge \vec{n} \wedge \vec{a} \parallel \vec{a}$; for $\vec{j} = \vec{j}_0 \exp(i\vec{k} \cdot \vec{r})$, we have $\vec{j}_\perp = -\vec{k} \wedge (\vec{k} \wedge \vec{j}) / k^2$.

The interaction between two charges is generated by the sum of the ϕ - and \vec{A} -fields. Which of the fields carries the main part of the interaction depends on the physical situation: in a dynamical setup (with frequency ω) and large distances $L \gg c/\omega$ the \vec{A} -field creates a term that compensates for the instant ϕ -field and only the retarded component of \vec{A} survives. In quasi-stationary situations, on the other hand, the Coulomb field ϕ is dominant and retardation effects are negligible.

The dynamic equations for the potentials in the Lorenz gauge are

$$\begin{aligned}\nabla^2 \phi - \frac{1}{c^2} \partial_t^2 \phi &= -4\pi\rho, \\ \nabla^2 \vec{A} - \frac{1}{c^2} \partial_t^2 \vec{A} &= -\frac{4\pi}{c} \vec{j}.\end{aligned}\quad (19.18)$$

These equations are manifestly covariant and both are retarded. Another gauge choice is $\phi = 0$. Then \vec{A} also includes a longitudinal component, which just takes over the role of ϕ and generates an instantaneous Coulomb interaction between two charges.

19.2 Lagrange function

The Lagrange function or Lagrangian that generates the Maxwell equations has the form

$$L = \int d^3r \left[\frac{E^2(\vec{r}, t) - B^2(\vec{r}, t)}{8\pi} - \rho(\vec{r}, t)\phi(\vec{r}, t) + \frac{1}{c} \vec{j}(\vec{r}, t) \cdot \vec{A}(\vec{r}, t) \right]. \quad (19.19)$$

The variation with respect to the scalar potential ϕ generates the Coulomb law⁶ (we write $\phi_{,i}$ as an abbreviation of $\partial_{x_i}\phi$)

$$\begin{aligned}\frac{\delta L}{\delta \phi} &= -\rho(\vec{r}, t), & \frac{\delta L}{\delta \phi_{,i}} &= -E_i/4\pi, & \frac{\delta L}{\delta \dot{\phi}} &= 0, \\ \Rightarrow \vec{\nabla} \cdot \vec{E} &= 4\pi\rho,\end{aligned}\quad (19.20)$$

⁶The Euler-Lagrange equation to the Lagrangian $L[\eta, \partial_i \eta, \dot{\eta}]$ for the field $\eta(\vec{r}, t)$ has the form

$$\frac{d}{dt} \frac{\delta L}{\delta \dot{\eta}} + \sum_j \partial_j \frac{\delta L}{\delta (\partial_j \eta)} - \frac{\delta L}{\delta \eta} = 0.$$

whereas the variation with respect to the vector potential \vec{A} generates Ampère's law.

$$\begin{aligned} \frac{\delta L}{\delta A_i} &= j_i/c, \quad \frac{\delta L}{\delta A_{i,j}} = \varepsilon_{ijk} B_k/4\pi, \quad \frac{\delta L}{\delta \dot{A}_i} = -E_i/4\pi c, \\ \Rightarrow \quad \vec{\nabla} \wedge \vec{B} &= \frac{1}{c} \partial_t \vec{E} + \frac{4\pi}{c} \vec{j}. \end{aligned} \quad (19.21)$$

The conjugate fields of ϕ and \vec{A} are

$$\begin{aligned} \Pi_\phi &= \frac{\partial L}{\partial \dot{\phi}} = 0, \\ \Pi_{A_i} &= \frac{\partial L}{\partial \dot{A}_i} = -E_i/4\pi c. \end{aligned} \quad (19.22)$$

Thus, the scalar potential ϕ has no dynamics and \vec{A} is conjugated to \vec{E} ; \vec{A} and \vec{E} cannot be sharp at the same time. In the Coulomb gauge, ϕ can be found by direct integration,

$$\phi(\vec{r}, t) = \int d^3 r' \frac{\rho(\vec{r}', t)}{|\vec{r}' - \vec{r}|}, \quad (19.23)$$

and can be eliminated from L ; the Lagrangian L then only contains the (dynamic) \vec{A} -field,

$$\begin{aligned} L = \int d^3 r \left[\frac{E_\perp^2(\vec{r}, t) - B^2(\vec{r}, t)}{8\pi} + \frac{1}{c} \vec{j}_\perp(\vec{r}, t) \cdot \vec{A}(\vec{r}, t) \right. \\ \left. - \frac{1}{2} \int d^3 r' \frac{\rho(\vec{r}, t) \rho(\vec{r}', t)}{|\vec{r}' - \vec{r}|} \right], \end{aligned} \quad (19.24)$$

where

$$\vec{E}_\perp = -\frac{1}{c} \partial_t \vec{A}, \quad \vec{j}_\perp = \vec{j} - \frac{1}{4\pi} \vec{\nabla} \partial_t \phi. \quad (19.25)$$

We can view the charges and currents as particle- and current densities of moving charged particles,

$$\rho = \sum_i e_i \delta(\vec{r} - \vec{r}_i), \quad \vec{j} = \sum_i e_i \vec{v}_i \delta(\vec{r} - \vec{r}_i). \quad (19.26)$$

The continuity equation follows from microscopic considerations,

$$\partial_t \rho = \sum_i e_i \vec{\nabla} \delta(\vec{r} - \vec{r}_i) \underbrace{(-\partial_t \vec{r}_i)}_{-\vec{v}_i} = -\vec{\nabla} \cdot \vec{j}. \quad (19.27)$$

If we add the kinetic energy of the particles to L , we obtain the Lagrangian function for the coupled system of charged particles in an electromagnetic field,

$$L = \frac{1}{2} \sum_i m v_i^2 + \int d^3r \frac{E^2(\vec{r}, t) - B^2(\vec{r}, t)}{8\pi} - \sum_i e_i \phi(\vec{r}_i, t) + \sum_i \frac{e_i}{c} \vec{v}_i \cdot \vec{A}(\vec{r}_i, t). \quad (19.28)$$

The variation with respect to the particle coordinates \vec{r}_i generates the equation of motion for the i^{th} particle,

$$\begin{aligned} \frac{\delta L}{\delta \vec{r}_i} &= -e_i \left(\vec{\nabla} \phi(\vec{r}_i, t) - \frac{1}{c} \vec{\nabla} [\vec{v}_i \cdot \vec{A}(\vec{r}_i, t)] \right), \\ \frac{\delta L}{\delta \vec{v}_i} &= \vec{p}_i = m \vec{v}_i + \frac{e_i}{c} \vec{A}(\vec{r}_i, t), \\ \Rightarrow 0 &= m \dot{\vec{v}}_i + \frac{e_i}{c} \dot{\vec{A}}(\vec{r}_i, t) + e_i \vec{\nabla} \phi(\vec{r}_i, t) - \frac{e_i}{c} \vec{\nabla} (\vec{v}_i \cdot \vec{A}(\vec{r}_i, t)). \end{aligned} \quad (19.29)$$

Using the identities ⁷

$$\begin{aligned} \frac{d\vec{A}(\vec{r}_i, t)}{dt} &= \frac{\partial \vec{A}(\vec{r}_i, t)}{\partial t} + [\vec{v}_i \cdot \vec{\nabla}] \vec{A}(\vec{r}_i, t), \\ \vec{\nabla}(\vec{v} \cdot \vec{A}) - (\vec{v} \cdot \vec{\nabla}) \vec{A} &= \vec{v} \wedge (\vec{\nabla} \wedge \vec{A}) = \vec{v} \wedge \vec{B}, \end{aligned} \quad (19.30)$$

we find the well-known equation of motion of charged particles in the electromagnetic field with the Lorentz force on the right hand side,

$$m \dot{\vec{v}}_i = e_i \left(\underbrace{-\vec{\nabla} \phi(\vec{r}_i, t) - \frac{1}{c} \frac{\partial \vec{A}(\vec{r}_i, t)}{\partial t}}_{\vec{E}(\vec{r}_i, t)} + \frac{1}{c} \vec{v}_i \wedge \vec{B}(\vec{r}_i, t) \right). \quad (19.31)$$

Finally, we obtain the Hamiltonian from the Lagrangian through a Legendre transformation $[\vec{\eta} = (\phi, \vec{A})]$

$$H = \sum_i \vec{p}_i \cdot \vec{v}_i + \int d^3r \sum_k \dot{\eta}_k \frac{\delta L}{\delta \dot{\eta}_k} - L, \quad (19.32)$$

⁷We use the relation

$$\vec{a} \wedge (\vec{\nabla} \wedge \vec{b}) = \vec{\nabla}(\vec{a} \cdot \vec{b}) - (\vec{a} \cdot \vec{\nabla}) \vec{b} - (\vec{b} \cdot \vec{\nabla}) \vec{a} - \vec{b} \wedge (\vec{\nabla} \wedge \vec{a})$$

for $\vec{a} = \vec{v}$ and $\vec{b} = \vec{A}$,

$$\vec{v} \wedge (\vec{\nabla} \wedge \vec{A}) = \vec{\nabla}(\vec{v} \cdot \vec{A}) - (\vec{v} \cdot \vec{\nabla}) \vec{A} - (\vec{A} \cdot \vec{\nabla}) \vec{v} - \vec{A} \wedge (\vec{\nabla} \wedge \vec{v}).$$

The last two terms disappear, $A_i \partial_i \partial_i x_j = 0$ and $(\vec{\nabla} \wedge \vec{v}) = \partial_t (\vec{\nabla} \wedge \vec{r}) = 0$.

$$\begin{aligned}
H &= \sum_i \vec{v}_i \cdot \underbrace{\left(m\vec{v}_i + \frac{e_i}{c}\vec{A}\right)}_{\sim \delta L / \delta \vec{v}_i} + \int d^3r \frac{1}{4\pi c} \underbrace{\dot{\vec{A}} \cdot (-\vec{E})}_{\sim \delta L / \delta \dot{\vec{A}}} \\
&\quad - \int d^3r \frac{E^2 - B^2}{8\pi} + \sum_i e_i \phi - \sum_i \frac{e_i}{c} \vec{v}_i \cdot \vec{A} - \sum_i \frac{m}{2} \underbrace{\left(\frac{\vec{p}_i - (e_i/c)\vec{A}}{m}\right)^2}_{\vec{v}_i} \\
&= \frac{1}{2m} \sum_i \left[\vec{p}_i - \frac{e_i}{c}\vec{A}(\vec{r}_i, t)\right]^2 + \sum_i e_i \phi(\vec{r}_i, t) + \int \frac{d^3r}{8\pi} [E^2 + B^2 + 2\vec{\nabla}\phi \cdot \vec{E}].
\end{aligned} \tag{19.33}$$

A partial integration in the last term gives

$$\int \frac{d^3r}{4\pi} \vec{\nabla}\phi \cdot \vec{E} = - \int d^3r \phi \rho = - \sum_i e_i \phi(\vec{r}_i),$$

which compensates for the second term in (19.33). This leads us the result

$$H = \frac{1}{2m} \sum_i \left[\vec{p}_i - \frac{e_i}{c}\vec{A}(\vec{r}_i, t)\right]^2 + \int d^3r \frac{E^2 + B^2}{8\pi} \tag{19.34}$$

$$\begin{aligned}
&= \frac{1}{2m} \sum_i \left[\vec{p}_i - \frac{e_i}{c}\vec{A}(\vec{r}_i, t)\right]^2 + \int d^3r \frac{E_{\perp}^2 + B^2}{8\pi} \\
&\quad + \frac{1}{2} \sum_{i \neq j} \frac{e_i e_j}{|\vec{r}_i - \vec{r}_j|},
\end{aligned} \tag{19.35}$$

where the first equation is generally valid and the second holds in the Coulomb gauge. This completes our classical analysis.

Before quantizing the theory, we repeat some of the above formulas in tensor notation.⁸ We write the time-position vector in the form

$$x^\mu = (x^0, \vec{r}) = (t, x^i), \tag{19.36}$$

a contravariant four-vector; the associated covariant form is obtained from the contraction with the metric tensor $g_{\mu\nu}$,

$$x_\mu = g_{\mu\nu} x^\nu, \tag{19.37}$$

where

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \tag{19.38}$$

⁸We continue with Heaviside-Lorentz units, $F = k_1 e^2 / r^2$ with $[k_1]_{\text{Gauss}} = 1 \rightarrow [k_1]_{\text{HL}} = 1/4\pi$ and set $c = 1$.

is the Minkowski metric in \mathbb{M}^4 . The change between inertial systems is done by the Lorentz transformation Λ_ν^μ

$$x'^\mu = \Lambda_\nu^\mu x^\nu, \quad (19.39)$$

$$x'_\mu = (\Lambda^{-1})_\mu^\nu x_\nu. \quad (19.40)$$

Lorentz transformations conserve the metric,

$$g_{\alpha\beta} = \Lambda_\alpha^\mu \Lambda_\beta^\nu g_{\mu\nu}. \quad (19.41)$$

Thus, the four-length $x^2 = x_\mu x^\mu = t^2 - r^2$ is invariant under Lorentz transformation,

$$\begin{aligned} x'^2 &= x'_\mu x'^\mu = g_{\mu\nu} x'^\nu x'^\mu = g_{\mu\nu} \Lambda_\alpha^\nu \Lambda_\beta^\mu x^\alpha x^\beta \\ &= g_{\alpha\beta} x^\alpha x^\beta = x_\beta x^\beta = x^2. \end{aligned} \quad (19.42)$$

With x^μ contravariant, the associated derivative $\partial_\mu = \partial/\partial x^\mu$ is covariant, since it transform according to (19.40),

$$\partial'_\mu = \frac{\partial}{\partial x'^\mu} = \frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial}{\partial x^\alpha} = (\Lambda^{-1})_\mu^\alpha \partial_\alpha. \quad (19.43)$$

As a consequence, different signs appear in the contractions with derivatives:

$$\begin{aligned} x_\mu x^\mu &= t^2 - \vec{r} \cdot \vec{r}, & \text{but} \\ \partial_\mu A^\mu &= \partial_t A + \vec{\nabla} \cdot \vec{A}, & \text{because} \\ x_\mu &= (t, -\vec{r}), & \text{but} \\ \partial_\mu &= (\partial_t, \vec{\nabla}). \end{aligned} \quad (19.44)$$

The four-potential is $A^\mu = (\phi, \vec{A})$, the four-source is $j^\mu = (\rho, \vec{j})$, the electromagnetic field tensor $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ has the matrix form⁹

$$F = \begin{pmatrix} 0 & -E^x & -E^y & -E^z \\ E^x & 0 & -B^z & B^y \\ E^y & B^z & 0 & -B^x \\ E^z & -B^y & B^x & 0 \end{pmatrix}, \quad (19.45)$$

and the Maxwell equations are

$$\partial_\mu F^{\mu\nu} = j^\nu. \quad (19.46)$$

⁹The electromagnetic field tensor is anti-symmetric, $F^{\mu\nu} = -F^{\nu\mu}$ with $F^{i0} = E_i$, $F^{ij} = -\epsilon_{ijk} B_k$. Objects with latin indices $i, j, k, \dots = 1, 2, 3$ over the spatial coordinates are not classified according to co-/contravariance.

The Lagrangian density is

$$L = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - j_\mu A^\mu \quad (19.47)$$

and the continuity equation reads

$$\partial_\mu j^\mu = 0. \quad (19.48)$$

A^0 is not dynamical because $\Pi^0 = \delta L/(\delta \partial_t A^0) = 0$ and gauge invariance holds in the form

$$A'_\mu = A_\mu - \partial_\mu \chi. \quad (19.49)$$

The (non-covariant) Coulomb-gauge is $\partial_i A_i = 0$, the Lorentz gauge has the covariant form $\partial_\mu A^\mu = 0$.

In the Coulomb gauge,

$$A^0(\vec{r}, t) = \frac{1}{4\pi} \int d^3 r' \frac{\rho(\vec{r}', t)}{|\vec{r}' - \vec{r}|} \quad \text{and} \quad \square A^i = j_\perp^i, \quad (19.50)$$

with the D'Alembert operator $\square = \partial_\mu \partial^\mu = \partial_t^2 - \nabla^2$. After elimination of A^0 , we obtain the expression (19.24) for the Lagrangian¹⁰ and the Hamiltonian (19.35), including the particle sector. The latter are no longer manifestly covariant due to our choice of the Coulomb gauge. Why don't we keep the manifestly covariant expressions—after all, they are much more beautiful. The reason is, that here, for a correct quantization, we need a form that is reduced to the minimum number of degrees of freedom. If we want to quantize the theory in a 'manifestly covariant' way¹¹, we have to fix the gauge with better tricks¹². Also note, that a covariant form requires A^0 and A^i to assume equivalent roles. Accordingly, A^0 and A^i must both occur dynamically in L , which can be achieved by adding a term $f(\partial_t A^0, \dots)$; this is done with the help of the gauge freedom.

19.3 Quantization

In the Coulomb gauge, the transverse gauge or vector field \vec{A} is the only remaining dynamical field that has to be quantized (note that ϕ is and

¹⁰except for c and 4π due to different definitions

¹¹This is advantageous in high energy physics where c is relevant.

¹²This is the so-called gauging-fixing problem, see Bogoliubov-Shirkov.

remains only a \mathbb{C} -number). In source-free space, i.e., for $\rho = 0$ and $\vec{j} = 0$, we have to solve the wave equation combined with the gauge condition,

$$\left(\frac{1}{c^2} \partial_t^2 - \nabla^2\right) \vec{A} = 0, \quad \vec{\nabla} \cdot \vec{A} = 0. \quad (19.51)$$

The formula (19.51) is a somewhat more complicated (because vector-valued) form of (18.2)¹³. The normal modes of (19.51) are [see (18.8)],

$$\vec{A}(\vec{r}, t) = \sum_{\vec{k}, \lambda} \left(\frac{hc^2}{V\omega_{\vec{k}}}\right)^{1/2} \left(\vec{\varepsilon}_{\vec{k}}^{\lambda} a_{\vec{k}, \lambda} e^{i(\vec{k} \cdot \vec{r} - \omega_{\vec{k}} t)} + \vec{\varepsilon}_{\vec{k}}^{\lambda*} a_{\vec{k}, \lambda}^* e^{-i(\vec{k} \cdot \vec{r} - \omega_{\vec{k}} t)}\right), \quad (19.52)$$

where $\vec{k} = 2\pi\vec{n}/L$, $n_i = 0, \pm 1, \pm 2, \dots$, $V = L^3$, $\omega_{\vec{k}} = c|\vec{k}| > 0$, and the polarization vectors $\vec{\varepsilon}_{\vec{k}}^{\lambda}$ are orthogonal to the propagation vector \vec{k} , $\vec{k} \cdot \vec{\varepsilon}_{\vec{k}}^{\lambda} = 0$. The latter condition derives from the gauge condition $\vec{\nabla} \cdot \vec{A} = 0 \Rightarrow \vec{k} \cdot \vec{A}_{\vec{k}} = 0$. We can construct normal modes as waves with either linear or circular polarization. The linearly polarized waves form a positively oriented orthonormal basis with \hat{k} ,

$$\begin{aligned} \vec{\varepsilon}_{\vec{k}}^1 \wedge \vec{\varepsilon}_{\vec{k}}^2 &= \hat{k}, \\ \vec{\varepsilon}_{\vec{k}}^2 \wedge \hat{k} &= \vec{\varepsilon}_{\vec{k}}^1, \\ \hat{k} \wedge \vec{\varepsilon}_{\vec{k}}^1 &= \vec{\varepsilon}_{\vec{k}}^2. \end{aligned} \quad (19.53)$$

The circularly polarized waves result from the linear combination

$$\begin{aligned} \vec{\varepsilon}_{\vec{k}}^+ &= -(\vec{\varepsilon}_{\vec{k}}^1 + i\vec{\varepsilon}_{\vec{k}}^2)/\sqrt{2}, \\ \vec{\varepsilon}_{\vec{k}}^- &= (\vec{\varepsilon}_{\vec{k}}^1 - i\vec{\varepsilon}_{\vec{k}}^2)/\sqrt{2}. \end{aligned} \quad (19.54)$$

The polarization vectors are orthonormal in both cases,

$$\vec{\varepsilon}_{\vec{k}}^{\lambda'} \cdot \vec{\varepsilon}_{\vec{k}}^{\lambda} = \delta_{\lambda\lambda'}. \quad (19.55)$$

Again, we have introduced the action h in such a way that \vec{A} has the dimension of a vector potential,

$$\int d^3r \left(\vec{E} \sim \frac{1}{c} \partial_t \vec{A}\right)^2 \sim \frac{V\omega^2}{c^2} \left(\frac{hc^2}{V\omega}\right) \sim h\omega. \quad (19.56)$$

¹³In (18.2) the speed of sound $c_s = \sqrt{\varepsilon/m}$ replaces the speed of light and the field is a scalar one, ϕ instead of \vec{A} .

We quantize (19.52) by replacing the amplitudes $a_{\vec{k},\lambda}$ and $a_{\vec{k},\lambda}^*$ with operators $a_{\vec{k},\lambda}$ and $a_{\vec{k},\lambda}^\dagger$ and replacing \hbar with $2\pi\hbar$. We request these operators to satisfy bosonic commutation relations,

$$\begin{aligned} [a_{\vec{k},\lambda}, a_{\vec{k}',\lambda'}^\dagger] &= \delta_{\vec{k}\vec{k}'} \delta_{\lambda\lambda'}, \\ [a_{\vec{k},\lambda}, a_{\vec{k}',\lambda'}] &= 0, \\ [a_{\vec{k},\lambda}^\dagger, a_{\vec{k}',\lambda'}^\dagger] &= 0. \end{aligned} \quad (19.57)$$

As usual, we introduce the vacuum state $|0\rangle$ which is annihilated by the destruction operators $a_{\vec{k},\lambda}$,

$$a_{\vec{k},\lambda}|0\rangle = 0. \quad (19.58)$$

Alternatively, we can quantize the fields

$$\vec{A}(\vec{r}, t) \quad \text{and} \quad \frac{\delta L}{\delta \vec{A}} = -\vec{E}_\perp(\vec{r}, t)/4\pi c. \quad (19.59)$$

However, this quantization procedure is somewhat more complicated due to the gauge condition: with $\Pi_{\vec{A}} = \delta L / \delta \dot{\vec{A}} = -\vec{E}_\perp/4\pi c = (1/4\pi c^2) \dot{\vec{A}}$ and

$$\dot{\vec{A}}(\vec{r}, t) = -i \sum_{\vec{k},\lambda} \left(\frac{2\pi\hbar c^2 \omega_{\vec{k}}}{V} \right)^{1/2} \left(\vec{\varepsilon}_{\vec{k}}^\lambda a_{\vec{k}\lambda} e^{i(\vec{k}\cdot\vec{r} - \omega_{\vec{k}}t)} - \text{c.c.} \right), \quad (19.60)$$

we obtain

$$[A^j(\vec{r}, t), \Pi_A^l(\vec{r}', t)] = \frac{i\hbar}{2V} \sum_{\vec{k}\lambda} \left(\vec{\varepsilon}_{\vec{k}}^{\lambda j} \vec{\varepsilon}_{\vec{k}}^{\lambda l*} e^{i\vec{k}\cdot(\vec{r}-\vec{r}')} + \text{c.c.} \right). \quad (19.61)$$

From (19.55), it follows that

$$\sum_{\lambda} \varepsilon_{\vec{k}}^{\lambda j} \varepsilon_{\vec{k}}^{\lambda l*} + k^j k^l / k^2 = \delta_{jl}, \quad (19.62)$$

such that $\vec{\varepsilon}_{\vec{k}}^1$, $\vec{\varepsilon}_{\vec{k}}^2$, and \hat{k} form a basis of \mathbb{R}^3 . The relation (19.62) can alternatively be formulated as

$$\sum_{\lambda} \varepsilon_{\vec{k}}^{\lambda j} \varepsilon_{\vec{k}}^{\lambda l*} = \delta_{jl} - \frac{k^j k^l}{k^2} = \delta_{jl}^T, \quad (19.63)$$

and we obtain commutators for the fields in the form

$$[A^j(\vec{r}, t), \Pi_A^l(\vec{r}', t)] = i\hbar \left(\delta^{jl} - \frac{\partial_j \partial_l}{\nabla^2} \right) \delta^3(\vec{r} - \vec{r}'), \quad (19.64)$$

while a naive straightforward field quantization would have missed the transverse nature of the Kronecker.

Finally, we calculate

- the energy ($\rightarrow \hbar\omega_k$),
- the momentum ($\rightarrow \hbar\vec{k}$),
- the angular momentum ($\rightarrow \text{spin } \pm\hbar$)

of the field (\rightarrow of the modes). This will allow us to interpret the operators

$$\begin{aligned} a_{\vec{k},+}^\dagger &= -(a_{\vec{k},1}^\dagger + ia_{\vec{k},2}^\dagger)/\sqrt{2}, \\ a_{\vec{k},-}^\dagger &= (a_{\vec{k},1}^\dagger - ia_{\vec{k},2}^\dagger)/\sqrt{2}, \end{aligned} \quad (19.65)$$

as ‘creators’ of photons with energy $\hbar\omega_k$, momentum $\hbar\vec{k}$, and helicity $+$ (or right-handed, spin 1 along $+\hat{k}$) and helicity $-$ (or left-handed, spin 1 along $-\hat{k}$), as sketched in 19.2.¹⁴

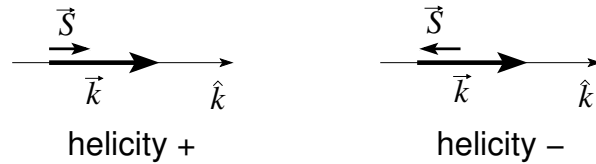


Fig. 19.2: Photons with helicities $+$ and $-$.

¹⁴The helicity is to be distinguished from the chirality which is a transformation property (particles transform according to a left- or right-handed representation of the Poincaré group). For massless particles (photon, gluon, graviton) the helicity is a relativistic invariant and helicity and chirality are identical. This does not apply to massive particles, since the helicity changes if the observer overtakes the particle.

19.3.1 Energy

In analogy to (18.10), we calculate ¹⁵

$$\begin{aligned}
H &= \int \frac{d^3r}{8\pi} \{ : E_{\perp}^2(\vec{r}, t) : + : B^2(\vec{r}, t) : \} \\
&= \frac{\hbar}{4} \int \frac{d^3r}{V} \sum_{\vec{k}\vec{k}'\lambda\lambda'} \left\{ \left(\sqrt{\omega_{\vec{k}}\omega_{\vec{k}'}} \vec{\varepsilon}_{\vec{k}}^{\lambda} \cdot \vec{\varepsilon}_{\vec{k}'}^{\lambda'} + (\vec{k} \wedge \vec{\varepsilon}_{\vec{k}}^{\lambda}) \cdot (\vec{k}' \wedge \vec{\varepsilon}_{\vec{k}'}^{\lambda'}) \frac{c^2}{\sqrt{\omega_{\vec{k}}\omega_{\vec{k}'}}} \right) \right. \\
&\quad \times \left(a_{\vec{k}\lambda}^{\dagger} a_{\vec{k}'\lambda'} e^{-i(\vec{k}-\vec{k}')\cdot\vec{r}} e^{i(\omega_{\vec{k}}-\omega_{\vec{k}'})t} + a_{\vec{k}'\lambda'}^{\dagger} a_{\vec{k}\lambda} e^{i(\vec{k}-\vec{k}')\cdot\vec{r}} e^{-i(\omega_{\vec{k}}-\omega_{\vec{k}'})t} \right) \\
&\quad - \left(\sqrt{\omega_{\vec{k}}\omega_{\vec{k}'}} \vec{\varepsilon}_{\vec{k}}^{\lambda} \cdot \vec{\varepsilon}_{\vec{k}'}^{\lambda'} + (\vec{k} \wedge \vec{\varepsilon}_{\vec{k}}^{\lambda}) \cdot (\vec{k}' \wedge \vec{\varepsilon}_{\vec{k}'}^{\lambda'}) \frac{c^2}{\sqrt{\omega_{\vec{k}}\omega_{\vec{k}'}}} \right) \\
&\quad \times \left(a_{\vec{k}\lambda} a_{\vec{k}'\lambda'} e^{i(\vec{k}+\vec{k}')\cdot\vec{r}} e^{-i(\omega_{\vec{k}}+\omega_{\vec{k}'})t} + a_{\vec{k}\lambda'}^{\dagger} a_{\vec{k}\lambda}^{\dagger} e^{-i(\vec{k}+\vec{k}')\cdot\vec{r}} e^{i(\omega_{\vec{k}}+\omega_{\vec{k}'})t} \right) \Big\}.
\end{aligned} \tag{19.66}$$

The relations

$$\begin{aligned}
\int \frac{d^3r}{V} e^{\pm i(\vec{k}-\vec{k}')\cdot\vec{r}} &= \delta_{\vec{k},\vec{k}'}, \\
\int \frac{d^3r}{V} e^{\pm i(\vec{k}+\vec{k}')\cdot\vec{r}} &= \delta_{\vec{k},-\vec{k}'}, \\
\vec{\varepsilon}_{\vec{k}}^{\lambda} \cdot \vec{\varepsilon}_{\vec{k}}^{\lambda'} &= \delta_{\lambda\lambda'}, \\
(\vec{k} \wedge \vec{\varepsilon}_{\vec{k}}^{\lambda}) \cdot (\vec{k} \wedge \vec{\varepsilon}_{-\vec{k}}^{\lambda'}) &= \frac{\omega_{\vec{k}}^2}{c^2} \vec{\varepsilon}_{\vec{k}}^{\lambda} \cdot \vec{\varepsilon}_{-\vec{k}}^{\lambda'}, \\
(\vec{k} \wedge \vec{\varepsilon}_{\vec{k}}^{\lambda}) \cdot (\vec{k} \wedge \vec{\varepsilon}_{\vec{k}}^{\lambda'}) &= \frac{\omega_{\vec{k}}^2}{c^2} \delta_{\lambda\lambda'},
\end{aligned} \tag{19.67}$$

help us to drastically simplify the result and write it in the form of a sum over harmonic oscillator modes,

$$H = \sum_{\vec{k}\lambda} \hbar\omega_{\vec{k}} a_{\vec{k}\lambda}^{\dagger} a_{\vec{k}\lambda}, \tag{19.68}$$

where each (photon) mode comes with an energy $\hbar\omega_{\vec{k}} = \hbar ck$.

¹⁵By calculating the normally ordered expression, we eliminate a divergent contribution due to the zero-point fluctuations of the vacuum, such that $\langle 0|H|0\rangle = 0$. We assume linear, real polarization vectors. Complex valued $\vec{\varepsilon}_{\vec{k}}^{\lambda}$ increase the effort in the derivation.

19.3.2 Momentum

The momentum is given by the integral over the momentum density $\vec{p} = \vec{E} \wedge \vec{B}/4\pi c$,

$$\begin{aligned}
P &= \int \frac{d^3r}{4\pi c} \{ : \vec{E}_\perp(\vec{r}, t) \wedge \vec{B}(\vec{r}, t) : \} \\
&= \frac{\hbar}{2} \int \frac{d^3r}{V} \sum_{\vec{k}\vec{k}', \lambda\lambda'} \left\{ \sqrt{\omega_{\vec{k}}/\omega_{\vec{k}'}} \left(\vec{\varepsilon}_{\vec{k}}^\lambda \wedge (\vec{k}' \wedge \vec{\varepsilon}_{\vec{k}'}^{\lambda'}) \right) \right. \\
&\quad \times \left(a_{\vec{k}\lambda}^\dagger a_{\vec{k}'\lambda'} e^{-i(\vec{k}-\vec{k}')\cdot\vec{r}} e^{i(\omega_{\vec{k}}-\omega_{\vec{k}'})t} + a_{\vec{k}'\lambda'}^\dagger a_{\vec{k}\lambda} e^{i(\vec{k}-\vec{k}')\cdot\vec{r}} e^{-i(\omega_{\vec{k}}-\omega_{\vec{k}'})t} \right) \\
&\quad - \sqrt{\omega_{\vec{k}}/\omega_{\vec{k}'}} \left(\vec{\varepsilon}_{\vec{k}}^\lambda \wedge (\vec{k}' \wedge \vec{\varepsilon}_{\vec{k}'}^{\lambda'}) \right) \\
&\quad \times \left(a_{\vec{k}\lambda} a_{\vec{k}'\lambda'} e^{i(\vec{k}+\vec{k}')\cdot\vec{r}} e^{-i(\omega_{\vec{k}}+\omega_{\vec{k}'})t} + a_{\vec{k}\lambda}^\dagger a_{\vec{k}'\lambda'}^\dagger e^{-i(\vec{k}+\vec{k}')\cdot\vec{r}} e^{i(\omega_{\vec{k}}+\omega_{\vec{k}'})t} \right) \Big\} \\
&= \sum_{\vec{k}\lambda} \hbar \vec{k} a_{\vec{k}\lambda}^\dagger a_{\vec{k}\lambda}, \tag{19.69}
\end{aligned}$$

where we used that

$$\begin{aligned}
\vec{\varepsilon}_{\vec{k}}^\lambda \wedge (\vec{k} \wedge \vec{\varepsilon}_{\vec{k}}^{\lambda'}) &= (\vec{\varepsilon}_{\vec{k}}^\lambda \cdot \vec{\varepsilon}_{\vec{k}}^{\lambda'}) \vec{k} - (\vec{\varepsilon}_{\vec{k}}^\lambda \cdot \vec{k}) \vec{\varepsilon}_{\vec{k}}^{\lambda'} = \vec{k} \delta_{\lambda\lambda'}, \\
\vec{\varepsilon}_{\vec{k}}^\lambda \wedge (\vec{k} \wedge \vec{\varepsilon}_{-\vec{k}}^{\lambda'}) - \vec{\varepsilon}_{-\vec{k}}^{\lambda'} \wedge (\vec{k} \wedge \vec{\varepsilon}_{\vec{k}}^\lambda) \\
&= \vec{\varepsilon}_{-\vec{k}}^{\lambda'} \wedge (\vec{k} \wedge \vec{\varepsilon}_{\vec{k}}^\lambda) + \vec{k} \wedge (\vec{\varepsilon}_{\vec{k}}^\lambda \wedge \vec{\varepsilon}_{-\vec{k}}^{\lambda'}) - \vec{\varepsilon}_{-\vec{k}}^{\lambda'} \wedge (\vec{k} \wedge \vec{\varepsilon}_{\vec{k}}^\lambda) \\
&= \vec{k} \wedge (\vec{\varepsilon}_{\vec{k}}^\lambda \wedge \vec{\varepsilon}_{-\vec{k}}^{\lambda'}) = 0.
\end{aligned}$$

that follow from the vector identities $\vec{a} \wedge (\vec{b} \wedge \vec{c}) = (\vec{a} \cdot \vec{c}) \vec{b} - (\vec{a} \cdot \vec{b}) \vec{c}$ and $\vec{a} \wedge (\vec{b} \wedge \vec{c}) + \vec{b} \wedge (\vec{c} \wedge \vec{a}) + \vec{c} \wedge (\vec{a} \wedge \vec{b}) = 0$.

19.3.3 Angular momentum

The calculation of the spin and angular momentum associated with $\vec{A}(\vec{r}, t)$ is more involved; here is a brief overview of the procedure. We start with the definition of the rotation operator $U_{\vec{\omega}}$ for a vector field (here \vec{A}) associated to the rotation $R_{\vec{\omega}}$, as described in chapter 11,

$$\begin{aligned}
U_{\vec{\omega}} \vec{A}(\vec{r}, t) &= R_{\vec{\omega}} \vec{A}(R_{-\vec{\omega}} \vec{r}, t) \\
&= [\mathbb{1} - (i/\hbar) \vec{\omega} \cdot (\vec{L} \mathbb{1} + \vec{S})] \vec{A}(\vec{r}, t). \tag{19.70}
\end{aligned}$$

The relation (19.70) expresses the action of $U_{\vec{\omega}}$ in the photonic Fock space that corresponds to the rotation $R_{\vec{\omega}}$ in real space. The rotation increment $\delta\vec{A}$ associated to an infinitesimal rotation $\delta\vec{\omega}$ has an orbital and a spin component

$$\begin{aligned}
 A_{\vec{\omega}}^j(\vec{r}, t) &= A^j - \varepsilon_{iml}\omega_i r^m \partial_l A^j + \varepsilon_{imj}\omega_i A^m \\
 \delta A^j(\vec{r}, t) &= [-\varepsilon_{ijl}A^l - \underbrace{\varepsilon_{iml}r^m \partial_l A^j}_{i(\vec{r}\wedge\vec{p})^i/\hbar=(i/\hbar)L^i}] \delta\omega_i \\
 &= -[\underbrace{\varepsilon_{ijl}A^l}_{\text{spin}} + \underbrace{(i/\hbar)L^i A^j}_{\text{orbit}}] \delta\omega_i.
 \end{aligned} \tag{19.71}$$

The task is to find the operator \vec{J} acting in the photonic Fock space that generates this rotation. A good ansatz for the angular momentum operator \vec{J} of the electromagnetic field follows from the classical formula for the angular momentum, expressed by the quantized field \vec{A} (correspondence principle),

$$\vec{J} = \int \frac{d^3r}{4\pi c} \vec{r} \wedge \{ : \vec{E}_{\perp} \wedge \vec{B} : \}. \tag{19.72}$$

Demonstrating that the expression \vec{J} produces an infinitesimal rotation of $\vec{A}(\vec{r}, t)$,

$$[J^i, A^j] = -i\hbar \frac{\delta A^j}{\delta\omega^i} = i\hbar \varepsilon_{ijk} A^k - L^i A^j, \tag{19.73}$$

we verify that \vec{J} indeed is the generalization of the classical angular momentum, or rather, the generator of the rotation in the field theory. After finding the expression (19.74) (and thus the representation of \vec{J} through fields \vec{A}), we show that the commutator $[J^i, A^j]$, see (19.75), actually matches up with (19.73). This allows us to identify the orbital angular momentum and spin components of \vec{J} in (19.74) and express the latter by creation and annihilation operators. By converting from creation/annihilation operators of linearly to circularly polarized waves, we find the spin operator (19.78) which counts the spin quanta $\pm\hbar$ in the direction \hat{k} .

In following this program, we first simplify the expression for \vec{J} ,

$$\begin{aligned}
 J^i &= \int \frac{d^3r}{4\pi c} \varepsilon_{ijl} r^j \varepsilon_{lkm} E_{\perp}^k \underbrace{\varepsilon_{mab} \partial_a A^b}_{B^m} \\
 &\quad \underbrace{(\vec{E}_{\perp} \wedge \vec{B})_l}_{[\vec{r} \wedge (\vec{E}_{\perp} \wedge \vec{B})]_i}
 \end{aligned}$$

$$\begin{aligned}
&= \int \frac{d^3r}{4\pi c} \varepsilon_{ijl} r^j (\delta_{la} \delta_{kb} - \delta_{ka} \delta_{lb}) E_{\perp}^k \partial_a A^b \\
&= \int \frac{d^3r}{4\pi c} \varepsilon_{ijl} r^j \left(E_{\perp}^b \partial_l A^b - \underbrace{E_{\perp}^a \partial_a A^l}_{\vec{E}_{\perp} \cdot \vec{\nabla}} \right); \\
\text{we use } & - \int d^3r \varepsilon_{ijl} r^j \vec{E}_{\perp} \cdot \vec{\nabla} A^l = \int d^3r \varepsilon_{ijl} \left(\underbrace{r^j \vec{\nabla} \cdot \vec{E}_{\perp}}_0 A^l + \underbrace{\delta^{jb}}_{\partial_b r^j} E_{\perp}^b A^l \right) \\
&= - \int \frac{d^3r}{4\pi c^2} \left(\varepsilon_{ijl} r^j \partial_t A^b \partial_l A^b + \varepsilon_{ibl} \partial_t A^b A^l \right) \\
&= \frac{-1}{4\pi c^2} \int d^3r : \partial_t A^b [(i/\hbar) L^i A^b + \varepsilon_{ibl} A^l] : . \tag{19.74}
\end{aligned}$$

For the commutator $[J^i, A^j]$, we then find (using $[\partial_t A^b, A^j] = -4\pi i \hbar c^2 (\delta^{jb} - (\partial_j \partial_b)/\nabla^2) \delta^3(\vec{r} - \vec{r}')$)

$$[J^i, A^j] = - \int d^3r \left[\left(\delta^{jb} - \frac{\partial_j \partial_b}{\nabla^2} \right) \delta(\vec{r} - \vec{r}') \right] (L^i A^b - i \hbar \varepsilon_{ibl} A^l). \tag{19.75}$$

The term $\propto \partial_j \partial_b / \nabla^2$ gives no contribution as the partial integration of ∂_b vanishes,

$$\begin{aligned}
&\partial_b (-i \hbar \varepsilon_{ilm} r^l \partial_m A^b - i \hbar \varepsilon_{ibl} A^l) \\
&= -i \hbar \varepsilon_{ilm} \left(\underbrace{\partial_b r^l}_{\delta^{bl}} \partial_m A^b + r^l \partial_m \underbrace{\partial_b A^b}_0 \right) - i \hbar \varepsilon_{ibl} \partial_b A^l \\
&= -i \hbar \varepsilon_{ibm} \partial_m A^b - i \hbar \varepsilon_{ibm} \partial_b A^m = 0.
\end{aligned}$$

The remaining term $\propto \delta^{jb}$ gives the commutator

$$\begin{aligned}
[J^i, A^j] &= -(L^i A^j - i \hbar \varepsilon_{ijl} A^l) \\
&\stackrel{(19.71)}{=} -i \hbar \frac{\delta A^j}{\delta \omega^i}, \tag{19.76}
\end{aligned}$$

in agreement with (19.73). Hence, \vec{J} indeed is the generator of the rotations and the spin part of the field is (see (19.74))

$$\begin{aligned}
S^i &= -\frac{1}{4\pi c^2} \int d^3r : \partial_t A^j \varepsilon_{ijl} A^l : \\
&= \frac{\hbar}{2} \int \frac{d^3r}{V} \sum_{\vec{k}\vec{k}', \lambda\lambda'} \sqrt{\omega_{\vec{k}}/\omega_{\vec{k}'}} i \varepsilon_{ijl} \vec{\varepsilon}_{\vec{k}}^{\lambda j} \vec{\varepsilon}_{\vec{k}'}^{\lambda' l}
\end{aligned}$$

$$\begin{aligned}
& \times \left(a_{\vec{k}\lambda} a_{\vec{k}'\lambda'} e^{i(\vec{k}+\vec{k}')\cdot\vec{r}} e^{-i(\omega_{\vec{k}}+\omega_{\vec{k}'})t} \right. \\
& \quad - a_{\vec{k}\lambda}^\dagger a_{\vec{k}'\lambda'}^\dagger e^{-i(\vec{k}+\vec{k}')\cdot\vec{r}} e^{i(\omega_{\vec{k}}+\omega_{\vec{k}'})t} \\
& \quad - a_{\vec{k}\lambda}^\dagger a_{\vec{k}'\lambda'} e^{-i(\vec{k}-\vec{k}')\cdot\vec{r}} e^{i(\omega_{\vec{k}}-\omega_{\vec{k}'})t} \\
& \quad \left. + a_{\vec{k}'\lambda'}^\dagger a_{\vec{k}\lambda} e^{i(\vec{k}-\vec{k}')\cdot\vec{r}} e^{-i(\omega_{\vec{k}}-\omega_{\vec{k}'})t} \right) \\
& = \frac{i\hbar}{2} \sum_{\vec{k}, \lambda\lambda'} \varepsilon_{ijl} \left[\varepsilon_{\vec{k}}^{\lambda j} \varepsilon_{-\vec{k}}^{\lambda' l} \left(a_{\vec{k}\lambda} a_{-\vec{k}\lambda'} e^{-2i\omega_{\vec{k}} t} - a_{\vec{k}\lambda}^\dagger a_{-\vec{k}\lambda'}^\dagger e^{2i\omega_{\vec{k}} t} \right) \right. \\
& \quad \left. - \left(\varepsilon_{\vec{k}}^{\lambda j} \varepsilon_{\vec{k}}^{\lambda' l} - \varepsilon_{\vec{k}}^{\lambda' j} \varepsilon_{\vec{k}}^{\lambda l} \right) a_{\vec{k}\lambda}^\dagger a_{\vec{k}\lambda'} \right]. \tag{19.77}
\end{aligned}$$

The first term is symmetric under exchange of j and l (the replacements $\vec{k} \rightarrow -\vec{k}$ and $\lambda \leftrightarrow \lambda'$ give the same term), the second term is anti-symmetric and survives the contraction with ε_{ijl} . In the result, we finally replace the operators associated with the linearly polarized waves with those that create/annihilate circularly polarized waves (note that this changes $i\hbar \rightarrow \hbar$) and find the final expression for the spin operator in the Fock space,

$$\begin{aligned}
\vec{S} &= i\hbar \sum_{\vec{k}, \lambda \neq \lambda'} \underbrace{\left(\vec{\varepsilon}_{\vec{k}}^{\lambda} \wedge \vec{\varepsilon}_{\vec{k}}^{\lambda'} \right)}_{\hat{k}} a_{\vec{k}\lambda'}^\dagger a_{\vec{k}\lambda}, \\
&= i\hbar \sum_{\vec{k}} \hat{k} (a_{\vec{k},2}^\dagger a_{\vec{k},1} - a_{\vec{k},1}^\dagger a_{\vec{k},2}) \\
&\quad \text{with: } a_1 = -(a_+ - a_-)/\sqrt{2}, \quad a_2 = -i(a_+ + a_-)/\sqrt{2}, \\
&\quad [a_+, a_+^\dagger] = 1 = [a_-, a_-^\dagger], \quad \text{all other } [\cdot, \cdot] \text{ are } 0, \\
&= \hbar \sum_{\vec{k}} \hat{k} [(a_{\vec{k},+}^\dagger + a_{\vec{k},-}^\dagger)(a_{\vec{k},+} - a_{\vec{k},-}) + (a_{\vec{k},+}^\dagger - a_{\vec{k},-}^\dagger)(a_{\vec{k},+} + a_{\vec{k},-})]/2, \\
&= \sum_{\vec{k}} \hbar \hat{k} [a_{\vec{k},+}^\dagger a_{\vec{k},+} - a_{\vec{k},-}^\dagger a_{\vec{k},-}]. \tag{19.78}
\end{aligned}$$

The results

– (19.68) for the energy

– (19.69) for the momentum

– (19.78) for the spin

allow us to interpret the quantized electromagnetic field as a photon field with the following properties:

Photons are quanta with the properties of particles as they feature,

- a momentum $\hbar\vec{k}$,
- an energy $\hbar\omega_{\vec{k}} = \hbar c|\vec{k}|$,
- a spin $\pm\hbar$ parallel to \hat{k} .

Photons have spin 1, but $|S_z| = 1$. Where did the spin $S_z = 0$ go?

19.4 Massive photons

The Maxwell theory of electromagnetic radiation is based on a massless carrier of the interaction between charged particles, the photon. What would the theory look like if the photon had mass, as is the case, for example, with the weak interaction (with the massive mediators W^\pm and Z)? The Lagrangian function for such a ‘massive’ theory has the form

$$L = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}M^2A_\mu A^\mu - j_\mu A^\mu, \quad (19.79)$$

with the additional mass term $M^2A_\mu A^\mu/2$. As before, $F^{\mu\nu}$ is the electromagnetic field tensor, A^μ is the four-potential, and j_μ is the four-vector of sources; the charge should be conserved, $\partial_\mu j^\mu = 0$. The Maxwell equations are replaced by Proca equations

$$\partial_\mu F^{\mu\nu} + M^2 A^\nu = j^\nu. \quad (19.80)$$

The four-divergence of (19.80) is (we use the antisymmetry $F^{\mu\nu} = -F^{\nu\mu}$)

$$M^2 \partial_\nu A^\nu = \partial_\nu j^\nu = 0 \quad (19.81)$$

and the Lorenz gauge becomes a condition (see (19.48), from which no gauge condition resulted): since the mass term is not gauge invariant, we lose the gauge invariance of the theory,

$$\begin{aligned} M^2 A_\mu A^\mu &\rightarrow M^2 A'_\mu A'^\mu = M^2 (A_\mu A^\mu - \partial_\mu \chi A^\mu - A_\mu \partial^\mu \chi + \partial_\mu \chi \partial^\mu \chi) \\ &\neq M^2 A_\mu A^\mu. \end{aligned} \quad (19.82)$$

The equations for the fields become massive,

$$(\square + M^2)A^\mu = j^\mu. \quad (19.83)$$

For example, for a stationary charge distribution $\rho(\vec{r}) = \delta(\vec{r})$, we find

$$\begin{aligned} (\Delta - M^2)\phi &= -\rho(\vec{r}), \\ \Rightarrow \phi &= \frac{1}{4\pi} \frac{e^{-rM}}{r}, \end{aligned} \quad (19.84)$$

which means that the interaction becomes short-range (what are the consequences of Cavendish experiment?) The dispersion for the electromagnetic waves becomes massive, i.e., the lowest energy is different from 0 by an energy gap,

$$\begin{aligned} A^\nu &\sim \varepsilon_k^{\lambda\nu} e^{i(\vec{k}\cdot\vec{r} - \omega_k t)} \\ \text{where } \omega_k &= \sqrt{M^2 + k^2}, \end{aligned} \quad (19.85)$$

see Fig 19.3.

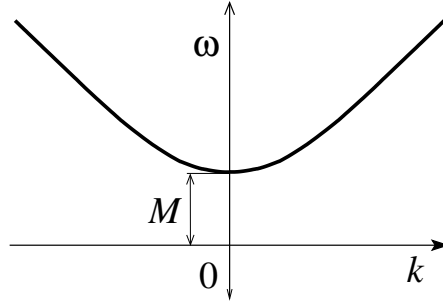


Fig. 19.3: Dispersion of electromagnetic waves in the Proca theory. The energy gap is a consequence of the finite photon mass in the Proca theory.

The polarisation vectors must satisfy the condition (see (19.80))

$$k_\mu \varepsilon_k^{\lambda\mu} = 0 \quad (19.86)$$

and we obtain 3 polarisations; the 3rd one is longitudinal and corresponds to the spin $S_z = 0$. Note that $\delta L / \delta \dot{A}^0 = 0$ still holds, i.e., A^0 is not dynamic. With the Proca theory, we have 3 dynamic modes $S_z = \pm 1, 0$.

The restriction $|S_z| = \max S$ generally applies to massless particles. This can be understood by reducing the Poincaré group¹⁶ (Wigner): one finds

¹⁶The Poincaré group is the symmetry group for Lorentz transformations combined with position and time translations. It replaces the Galileo group in the relativistic context.

that the irreducible representations are characterized by mass and spin. For $M \neq 0$ one has a $2S + 1$ -dimensional representation for S , see QM I, where we reduced $SO(3)$ and $SU(2)$. For $M = 0$, there are only 2-dimensional irreducible representations with $|S_z| = \max S$ and two helicities. This means that all massless long-range mediators appear with only two helicities.

$$\begin{array}{lll} \text{Photon,} & \text{Spin } 1, & S_z = \pm 1, \\ \text{Graviton,} & \text{Spin } 2, & S_z = \pm 2. \end{array} \quad (19.87)$$

Chapter 20

Radiation and Matter

Further material on the topics in this chapter can be found in the books by Sakurai and Gross.

The interaction between radiation and matter is obtained from the Hamilton¹ (19.34),

$$H = \frac{1}{2m} \sum_i \left[\vec{p}_i - \frac{q_i}{c} \vec{A}(\vec{r}_i, t) \right]^2 + \frac{1}{2} \sum_{i \neq j} \frac{q_i q_j}{|\vec{r}_i - \vec{r}_j|} + \int d^3r \frac{E_{\perp}^2 + B^2}{8\pi}.$$

This classical Hamilton function has to be transferred to quantum mechanics. To do this, we replace the particle coordinates and momenta \vec{r}_i and \vec{p}_i by operators which satisfy the commutation relations $[r_{i,j}, p_{k,l}] = i\hbar \delta_{ik} \delta_{jl}$ and quantize the electromagnetic field. The field quantization implies the substitution of the vector potential \vec{A} by its second quantized form with the amplitudes $a_{\vec{k},\lambda}$ replaced by operators with Bose exchange rules, in accordance with (19.57). This provides us with the Hamilton operator ²

$$\tilde{H} = \tilde{H}_{\text{particle}} + H_{\text{radiation}} + \tilde{H}_{\text{int}}, \quad (20.1)$$

$$\tilde{H}_{\text{particle}} = \sum_i \frac{p_i^2}{2m_i} + U(\vec{r}_i), \quad (20.2)$$

$$H_{\text{radiation}} = \sum_{\vec{k},\lambda} \hbar \omega_k a_{\vec{k},\lambda}^{\dagger} a_{\vec{k},\lambda}, \quad (20.3)$$

¹We denote charges with q_i to rule out sign problems; the electron has a charge $q = -e$ with $e > 0$ the unit charge.

²The tilde \tilde{H} stands for the representation in the Schrödinger picture—we will change to the interaction picture later.

$$\tilde{H}_{\text{int}} = \sum_i \left[\frac{-q_i}{2m_i c} \left(\vec{p}_i \cdot \vec{A}_i(\vec{r}_i, t) + \vec{A}_i(\vec{r}_i, t) \cdot \vec{p}_i \right) + \frac{q_i^2}{2m_i c^2} A^2(\vec{r}_i, t) \right]. \quad (20.4)$$

We have extended $\tilde{H}_{\text{particle}}$ by including a static potential $U(\vec{r}_i)$; it describes an external potential acting on the particles, as well as possible Coulomb interactions via an effective molecular field (for example in the Thomas-Fermi approximation). For a set of electrons, the charges q_i are given by $q_i = -e$, with the unit charge $e > 0$. The radiation field is described by

$$\vec{A}(\vec{r}, t) = \sum_{\vec{k}, \lambda} \left(\frac{2\pi\hbar c^2}{V\omega_k} \right)^{1/2} \left(\vec{\varepsilon}_{\vec{k}}^{\lambda} a_{\vec{k}\lambda} e^{i(\vec{k} \cdot \vec{r} - \omega_k t)} + \vec{\varepsilon}_{\vec{k}}^{\lambda*} a_{\vec{k}\lambda}^{\dagger} e^{-i(\vec{k} \cdot \vec{r} - \omega_k t)} \right). \quad (20.5)$$

Here, \vec{r} is a c valued variable and a, a^{\dagger} are operators on the photon Fock space. In formula (20.4), \vec{r} is replaced with the operator \vec{r}_i , which acts on the Hilbert space of the particle wave functions.

20.1 Consistent formulation

The Hamiltonian (20.1) is useful but is a combination of different formalisms. The electromagnetic field is relativistic and second-quantized—nothing else is possible because there is no first-quantized form of Maxwell’s theory. The particles are non-relativistic and first-quantized—we have coordinates and momenta instead of fields. In a better and more complete theory, we should consistently bring the formalism into a relativistic and second-quantized form. To do this, we have to represent the particles as fields, see (18.25), however, this is not as easily done for electrons as for spinless bosons. In fact, relativistic electrons are described by the Dirac theory, which will only be introduced in chapter 26; this path would lead us directly to quantum electrodynamics (QED), one of the most beautiful and perfect known theories. However, the non-relativistic first-quantized form (20.1) is quite useful for a first discussion.

Second, the time dependencies in (20.1) are inconsistent: The field operator $\vec{A}(\vec{r}, t)$ is time-dependent while the particle operators are time-independent. This discrepancy is due to the fact that we effectively use different pictures (the Heisenberg and the Schrödinger pictures). We first discuss the Hilbert space and then the three parts $\tilde{H}_{\text{particle}}$, $H_{\text{radiation}}$, and \tilde{H}_{int} of the Hamiltonian.

20.1.1 Hilbert space

Particles:

The particle Hilbert space \mathcal{H} is determined by the particle states Ψ_a , $\{\Psi_a\}$ a CONS (complete orthonormal system),

$$\bigoplus_a \Psi_a = \mathcal{H}. \quad (20.6)$$

Examples of suitable CONS are

- For a one-particle problem: Let $U(\vec{r}) = -Ze^2/r$ be the electronic potential generated by a nucleus³; then the Ψ_{nlm} including both bound and scattering states form an eigenbasis of $\hat{H}_{\text{particle}}$.
- For a many-body problem, we first construct a one-body (e.g., plane waves) basis and choose the product wave functions with suitable symmetrization for the⁴ $\Psi_a(x_1, \dots, x_N)$.

Radiation:

The Hilbert space for the electromagnetic radiation field is the Fock space of photon states $|n_1, \dots, n_k, \dots, n_\infty\rangle$,

$$\mathcal{F} = \bigoplus_N \left\{ \underbrace{\bigoplus_{\{n_1, \dots, n_k, \dots, n_\infty \mid \sum n_k = N\}} |n_1, \dots, n_k, \dots, n_\infty\rangle}_{N\text{-photon space}} \right\}. \quad (20.7)$$

Radiation and matter

We build the total Hilbert space from \mathcal{H} and \mathcal{F} ,

$$\mathcal{H}_{\text{RM}} = \mathcal{H} \otimes \mathcal{F} \quad (20.8)$$

with basis states $|a; n_1, \dots, n_k, \dots, n_\infty\rangle \in \mathcal{H}_{\text{RM}}$. The interaction Hamiltonian \mathcal{H}_{int} then induces transitions between the states $|a; n_1, \dots, n_k, \dots, n_\infty\rangle$ in the product basis.

³This leads us to the description of the H atom in the radiation field.

⁴Going one step further, you second-quantize the theory with field operators and arrive at quantum electrodynamics, QED.

20.1.2 Hamiltonian of the radiation field

The radiation field is described by (20.3) and (20.5). In the Heisenberg picture, we find the time evolution of the creation and annihilation operators $a_{\vec{k}\lambda}^\dagger$ and $a_{\vec{k}\lambda}$ using the equation of motion

$$\begin{aligned} i\hbar\partial_t a_{\vec{k}\lambda} &= [a_{\vec{k}\lambda}, H_{\text{radiation}}], \\ i\hbar\partial_t a_{\vec{k}\lambda}^\dagger &= [a_{\vec{k}\lambda}^\dagger, H_{\text{radiation}}], \end{aligned} \quad (20.9)$$

which shows that H is the infinitesimal generator of time translation.⁵ The calculation of the commutator in (20.9) yields the differential equation

$$i\hbar\partial_t a_{\vec{k}\lambda} = \sum_{\vec{k}', \lambda'} \hbar\omega_{k'} [a_{\vec{k}\lambda}, a_{\vec{k}'\lambda'}^\dagger a_{\vec{k}'\lambda'}] = \hbar\omega_k a_{\vec{k}\lambda}$$

and we find the time dependence of the creation and annihilation operators in the form

$$\begin{aligned} a_{\vec{k}\lambda}(t) &= a_{\vec{k}\lambda}(0) e^{-i\omega_k t}, \\ a_{\vec{k}\lambda}^\dagger(t) &= a_{\vec{k}\lambda}^\dagger(0) e^{i\omega_k t}. \end{aligned} \quad (20.11)$$

This time dependence is already taken into account in $\vec{A}(\vec{r}, t)$, see (20.5), that is, $\vec{A}(\vec{r}, t)$ is already formulated in the Heisenberg picture when we write

$$\vec{A}(\vec{r}, t) = \sum_{\vec{k}, \lambda} \left(\frac{2\pi\hbar c^2}{V\omega_k} \right)^{1/2} \left(\vec{\varepsilon}_k^\lambda a_{\vec{k}\lambda}(t) e^{i\vec{k}\cdot\vec{r}} + \vec{\varepsilon}_k^{\lambda*} a_{\vec{k}\lambda}^\dagger(t) e^{-i\vec{k}\cdot\vec{r}} \right). \quad (20.12)$$

20.1.3 Particle Hamiltonian

The description of the particles in $\tilde{H}_{\text{particle}}$, see (20.2), is based on time-independent operators \vec{p}_i and \vec{r}_i in the Schrödinger picture. The dynamics is therefore incorporated in the wave functions,

$$\begin{aligned} \Psi_a(\vec{r}, t) &= e^{-i\tilde{H}_{\text{particle}}t/\hbar} \Psi_a(\vec{r}, 0) \\ &= U_{\text{particle}}(t) \Psi_a(\vec{r}, 0). \end{aligned} \quad (20.13)$$

⁵Compare, for example, with the behavior of \vec{A} under rotations in the previous chapter, where we found that

$$i\hbar\partial_{\omega^i} A^j = [A^j, J^i]; \quad (20.10)$$

accordingly, we identified \vec{J} as the infinitesimal generator of the rotation.

In formula (20.13), only the ‘free’ particle dynamics is taken into account, without coupling to the electromagnetic field, in analogy to (20.11) and (20.12) where only the electromagnetic field determines the photonic dynamics without coupling to the particles.

In a consistent notation, we replace the Schrödinger picture for the particles with the Heisenberg picture,

$$\begin{aligned} \tilde{H}_{\text{particle}} \rightarrow H_{\text{particle}} &= U_{\text{particle}}^{-1}(t) \tilde{H}_{\text{particle}} U_{\text{particle}}(t) \\ \partial_t \tilde{H}_{\text{particle}} &\stackrel{=}{=} 0 \quad \tilde{H}_{\text{particle}}. \end{aligned} \quad (20.14)$$

20.1.4 Interaction Hamiltonian

Without an interaction between radiation and matter, the time evolution for the particles is independent of the time evolution of the radiation, with separate dynamics in the Hilbert and Fock spaces \mathcal{H} and \mathcal{F} . The interaction \mathcal{H}_{int} induces transitions between the states $|a; n_1, \dots, n_k, \dots, n_\infty\rangle$ and we want to examine these transitions as well as find the transition rates. We formulate this task in the interaction picture, where the operators carry the free evolution and the states carry the rest. The Hamilton operator in the interaction picture then has the form

$$H = U_{\text{particle}}^{-1} \tilde{H} U_{\text{particle}} = H_{\text{particle}} + H_{\text{radiation}} + H_{\text{int}} \quad (20.15)$$

with components

$$\begin{aligned} H_{\text{particle}} &= \tilde{H}_{\text{particle}}, & \text{since } \partial_t \tilde{H}_{\text{particle}} &= 0, \\ H_{\text{radiation}}, & & \text{unchanged,} \\ H_{\text{int}} &= U_{\text{particle}}^{-1} \tilde{H}_{\text{int}} U_{\text{particle}}. \end{aligned} \quad (20.16)$$

We assume that H_{int} is switched on and off adiabatically; in 2nd order perturbation theory, the dynamics of the state $|\Psi(t)\rangle$ is given by

$$\begin{aligned} |\Psi(t)\rangle &= e^{-iH_0(t-t_0)} \left[\mathbb{1} - \frac{i}{\hbar} \int_{t_0}^t dt_1 H_{\text{int}}(t_1) \right. \\ &\quad \left. - \frac{1}{2\hbar^2} \int_{t_0}^t dt_1 dt_2 T(H_{\text{int}}(t_1) H_{\text{int}}(t_2)) + \dots \right] |\Psi_0\rangle \Big|_{t_0 \rightarrow -\infty}. \end{aligned} \quad (20.17)$$

Here, $H_0 = H_{\text{particle}} + H_{\text{radiation}}$ describes the free evolution of the two systems without interaction and $|\Psi_0\rangle$ is the initial state at the time t_0 .

The interaction Hamiltonian H_{int} contains two terms, see (20.4), which we will briefly discuss in what follows (we suppress the tilde),

$$H_{\text{int}} = \sum_i \left[\frac{-q_i}{2m_i c} \left(\vec{p}_i \cdot \vec{A}_i(\vec{r}_i, t) + \vec{A}(\vec{r}_i, t) \cdot \vec{p}_i \right) + \frac{q_i^2}{2m_i c^2} A^2(\vec{r}_i, t) + q_i \phi(\vec{r}_i, t) \right].$$

Here, we have included a possible longitudinal field $\phi(\vec{r}_i, t)$ in the interaction (rather than the particle) term (for a pure radiation field, $\phi = 0$). For a uniform set of particles with charge q and mass m , we have $q_i = q$ and $m_i = m$ and can simplify

$$H_{\text{int}} = \int d^3r \left[-\frac{q}{c} \vec{j}(\vec{r}) \cdot \vec{A}(\vec{r}, t) + \frac{q^2}{2mc^2} \rho(\vec{r}) A^2(\vec{r}, t) + q \rho(\vec{r}) \phi(\vec{r}, t) \right]$$

with the particle and current densities defined by

$$\begin{aligned} \rho(\vec{r}) &= \sum_i \delta(\vec{r} - \vec{r}_i), \\ \vec{j}(\vec{r}) &= \frac{1}{2m} \sum_i [\vec{p}_i \delta(\vec{r} - \vec{r}_i) + \delta(\vec{r} - \vec{r}_i) \vec{p}_i]. \end{aligned} \quad (20.18)$$

Note that the velocity is given by

$$\vec{v}_i = \frac{\vec{p}_i}{m} - \frac{q}{mc} \vec{A}(\vec{r}_i, t), \quad (20.19)$$

and hence the real particle flow (or current density) is

$$\vec{J}(\vec{r}) = \underbrace{\vec{j}(\vec{r})}_{\text{para-}} - \underbrace{\frac{q}{mc} \rho(\vec{r}) \vec{A}(\vec{r}, t)}_{\text{dia-magnetic}} = -\frac{c}{q} \frac{\delta H_{\text{int}}}{\delta \vec{A}}. \quad (20.20)$$

In metals, the diamagnetic current is exactly canceled by a term in the paramagnetic current and the rest of the paramagnetic current generates Ohm's law⁶ $\vec{J} = \sigma \vec{E}$. In a superconductor, this cancellation is incomplete (leaving the superfluid density ρ_s uncompensated) and the diamagnetic current becomes the supercurrent, with the London equation $\vec{J} = -\rho_s \vec{A}$ replacing Ohm's law. Therefore

⁶Note that for a regular treatment of the conductor, scattering processes by impurities etc. have to be accounted for. Otherwise, we have to describe an ideal conductor, a singular object which is different from a superconductor, i.e, a superconductor \neq ideal conductor.

$$\begin{aligned} \text{in a metal} \quad \quad \quad \vec{J} &= \sigma \vec{E}, \\ \text{in a superconductor} \quad \vec{J} &= -\rho_s \vec{A}. \end{aligned}$$

The diamagnetic component is usually small, e.g., for metallic electrons in a radiation field (with $\hbar\omega$ of order eV, i.e., optical frequencies), we have

$$\frac{(q^2/mc^2)\rho A^2}{(q/c)jA} = \frac{eA}{mcv_F} \stackrel{\omega A/c=\mathcal{E}}{\sim} \frac{e\mathcal{E}}{\hbar\omega k_F} \sim \frac{\mathcal{E}}{10^8 \text{ eV/cm}} \ll 1,$$

because typical laboratory fields \mathcal{E} are small on the microscopic field scale 10^8 eV/cm (e.g., relatively large fields of order MV/cm appear in transistors that are still small on the microscopic field-scale, though; truly large fields of the order of microscopic fields appear in short-pulse lasers). A similar estimate applies to an electron bound in an atom, where $v \sim \hbar/ma_B$ and \hbar/a_B is of the same order as p_F .

20.2 H-Atom in the radiation field

To describe the absorption and emission of photons by the H (type) atom, see Fig. 20.1, we start from the following particle Hamiltonian,

$$H_{\text{particle}} = \frac{p^2}{2m} - \frac{Ze^2}{r}, \quad (20.21)$$

and consider transitions between the eigenstates $|a\rangle$ and $|b\rangle$ of H_{particle} ,

$$\begin{aligned} H_{\text{particle}} |a\rangle &= E_a |a\rangle, \\ H_{\text{particle}} |b\rangle &= E_b |b\rangle. \end{aligned} \quad (20.22)$$

We first consider the case of a photon emission (with wave vector \vec{k} and polarization λ) in a decay process from $|a\rangle$ to $|b\rangle$ with $E_a > E_b$. The initial⁷ and final states in the total Hilbert space of the particle in the radiation field are given by

$$\begin{aligned} |i\rangle &= |a; n_{\vec{k}\lambda}\rangle, \\ |f\rangle &= |b; n_{\vec{k}\lambda} + 1\rangle. \end{aligned} \quad (20.23)$$

In order to determine the transition rate, we have to calculate the time evolution (20.17) of the initial state and project it to the final state; in

⁷We denote the initial state with $|i\rangle$, the final state with $|f\rangle$

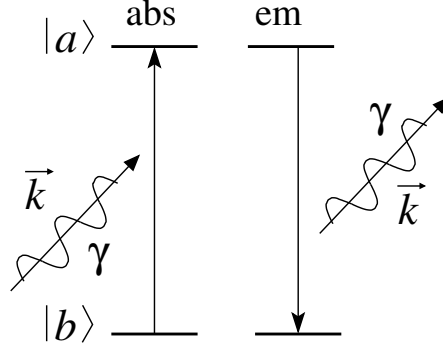


Fig. 20.1: Absorption and emission processes of a photon γ during transition between the electronic levels $|a\rangle$ and $|b\rangle$ in an atom.

doing so, it is sufficient to consider the first-order term in the perturbative expansion in the interaction Hamiltonian H_{int} ; with $|\Psi_0\rangle = |i\rangle$, we obtain the transition amplitude

$$\langle f|\Psi(t)\rangle \stackrel{1.\text{order}}{=} \frac{1}{i\hbar} e^{-i\mathcal{E}_f(t-t_0)/\hbar} \int_{t_0}^t dt_1 \langle f|H_{\text{int}}(t_1)|i\rangle \quad (20.24)$$

for $|a\rangle \neq |b\rangle$. Without loss of generality, we set all occupation numbers of uninvolved modes to zero, $n_{\vec{k}'\lambda'} = 0$. The matrix element $\langle f|H_{\text{int}}(t_1)|i\rangle$ can be written in the form

$$\langle f|H_{\text{int}}(t_1)|i\rangle = \langle b; n_{\vec{k}\lambda} + 1| U_{\text{particle}}^{-1} [(e\hbar/mci) \vec{A}(\vec{r}, t_1) \cdot \vec{\nabla}] U_{\text{particle}} |a; n_{\vec{k}\lambda}\rangle,$$

where we used the Coulomb gauge⁸ $\vec{\nabla} \cdot \vec{A}(\vec{r}, t) = 0$. We let the particle operators act on the particle coordinates, the field operators on the field state, and obtain (we choose the phase of $a_{\vec{k}'\lambda'}^\dagger$ shifted by $\exp(-i\omega_{\vec{k}'}t_0)$)

$$\begin{aligned} \langle f|H_{\text{int}}(t_1)|i\rangle &= e^{i(E_b - E_a)(t_1 - t_0)/\hbar} \frac{e\hbar}{mci} \int d^3r \left\{ \Psi_b^*(\vec{r}) [\vec{\nabla} \Psi_a(\vec{r})] \right. \\ &\quad \cdot \langle n_{\vec{k}\lambda} + 1 | \sum_{\vec{k}'\lambda'} \left(\frac{2\pi\hbar c^2}{V\omega_{\vec{k}'}} \right)^{1/2} \vec{\varepsilon}_{\vec{k}'}^{\lambda'*} a_{\vec{k}'\lambda'}^\dagger e^{-i\vec{k}' \cdot \vec{r} + i\omega_{\vec{k}'}(t_1 - t_0)} | n_{\vec{k}\lambda} \rangle \Big\} \\ &\quad \text{with } \langle n_{\vec{k}\lambda} + 1 | a_{\vec{k}'\lambda'}^\dagger | n_{\vec{k}\lambda} \rangle = \delta_{\vec{k}\vec{k}'} \delta_{\lambda\lambda'} (n_{\vec{k}\lambda} + 1)^{1/2} \\ &= e^{i(E_b - E_a + \hbar\omega_k)(t_1 - t_0)/\hbar} \frac{e\hbar}{mci} \left(\frac{2\pi\hbar c^2 (n_{\vec{k}\lambda} + 1)}{V\omega_k} \right)^{1/2} \\ &\quad \times \int d^3r e^{-i\vec{k} \cdot \vec{r}} \Psi_b^*(\vec{r}) \vec{\varepsilon}_{\vec{k}}^{\lambda*} \cdot \vec{\nabla} \Psi_a(\vec{r}). \end{aligned} \quad (20.25)$$

⁸Specifically, $[\vec{p} \cdot \vec{A} + \vec{p} \cdot \vec{A}] \Psi = [(\hbar/i)(\vec{\nabla} \cdot \vec{A}) + \vec{A} \cdot \vec{p}] \Psi + \vec{A} \cdot \vec{p} \Psi = 2\vec{A} \cdot \vec{p} \Psi$.

The transition amplitude $\langle f|\Psi(t)\rangle$ becomes

$$\begin{aligned} \langle f|\Psi(t)\rangle &= - \int_{t_0}^t dt_1 e^{-i[E_b + (n_{\vec{k}\lambda} + 1)\hbar\omega_k](t-t_1)/\hbar} \left(\frac{2\pi\hbar c^2(n_{\vec{k}\lambda} + 1)}{V\omega_k} \right)^{1/2} \frac{e}{mc} \\ &\times \int d^3r \left[e^{-i\vec{k}\cdot\vec{r}} \Psi_b^*(\vec{r}) \vec{\varepsilon}_{\vec{k}}^{\lambda*} \cdot \vec{\nabla} \Psi_a(\vec{r}) \right] e^{-i[E_a + n_{\vec{k}\lambda}\hbar\omega_k](t_1-t_0)/\hbar}. \end{aligned} \quad (20.26)$$

The phases in (20.26) reflect the free time evolution of $|i\rangle$ between t_0 and t_1 and of $|f\rangle$ between t_1 and t . The transition $|i\rangle \rightarrow |f\rangle$ can take place at any time t_1 between t_0 and t . All phases with t and t_0 are irrelevant for the transition probability $|\langle f|\Psi\rangle|^2$, only phases $\propto t_1$ are relevant. If we let $t_0 \rightarrow -\infty$ and $t \rightarrow \infty$, we find an expression for the differential transition rate⁹

$$\frac{d}{dt} |\langle f|\Psi(t)\rangle|^2 = \delta W_{ab} = \frac{2\pi}{\hbar} |f_{ab}|^2 \delta(E_b - E_a + \hbar\omega_k), \quad (20.27)$$

with

$$\begin{aligned} f_{ab} &= - \sqrt{\frac{2\pi\hbar c^2}{V\omega_k}} \frac{e\hbar\sqrt{n_{\vec{k}\lambda} + 1}}{mc} \int d^3r e^{-i\vec{k}\cdot\vec{r}} \Psi_b^*(\vec{r}) \vec{\varepsilon}_{\vec{k}}^{\lambda*} \cdot \vec{\nabla} \Psi_a(\vec{r}) \\ &= - \frac{e\hbar}{m} \sqrt{\frac{2\pi\hbar(n_{\vec{k}\lambda} + 1)}{V\omega_k}} \langle b | e^{-i\vec{k}\cdot\vec{r}} \vec{\varepsilon}_{\vec{k}}^{\lambda*} \cdot \vec{\nabla} | a \rangle. \end{aligned} \quad (20.28)$$

Next, we drop the specification \vec{k}, λ and let the initial state decay into any (allowed) mode; the summation over the continuum of photon states,

$$\sum_{\vec{k}\lambda} = V \int \frac{d^3k}{(2\pi)^3} \sum_{\lambda} = \frac{V}{2\pi^2} \int \frac{d\Omega_{\vec{k}}}{4\pi} \frac{\omega^2 d\omega}{c^3} \sum_{\lambda}, \quad (20.29)$$

then adds all the decay channels that are compatible with the energy conservation $E_b - E_a + \hbar\omega_k = 0$. In the following, we consider an isotropic unpolarized radiation field with $n_{\vec{k}\lambda} \equiv n_k$ and obtain for the decay rate of the $a \rightarrow b$ process,

$$W_{ab}^{\text{em}} = \frac{2e^2\hbar\omega}{m^2c^3} (n_k + 1) \oint_{\mathcal{X}} \frac{d\Omega_{\vec{k}}}{4\pi} |\langle b | e^{-i\vec{k}\cdot\vec{r}} \vec{\varepsilon}_{\vec{k}}^{\lambda*} \cdot \vec{\nabla} | a \rangle|^2. \quad (20.30)$$

⁹The differential transition rate is $\propto \partial_t |\int dt_1 \exp[i(E_b - E_a + \hbar\omega_k)t_1/\hbar]|^2 \sim 2\pi\hbar\delta(E_b - E_a + \hbar\omega_k)$.

The analogous result for the absorption of radiation \vec{k}, λ due to a transition from b to a (with $E_b < E_a$) can be obtained by calculating $\partial_t |\langle f | \Psi(t) \rangle|^2$ with

$$\begin{aligned}\Psi(t) &= U(t)|i\rangle, \\ |i\rangle &= |b; n_{\vec{k}\lambda}\rangle, \\ |f\rangle &= |a; n_{\vec{k}\lambda} - 1\rangle,\end{aligned}\tag{20.31}$$

and one finds (with $\langle n_{\vec{k}\lambda} - 1 | a_{\vec{k}'\lambda} | n_{\vec{k}\lambda} \rangle = \sqrt{n_{\vec{k}\lambda}} \delta_{\vec{k}\vec{k}'} \delta_{\lambda\lambda'}$ and again assuming an unpolarized, isotropic radiation field)

$$W_{ab}^{\text{abs}} = \frac{2e^2 \hbar \omega}{m^2 c^3} n_k \sum_{\lambda} \frac{d\Omega_{\vec{k}}}{4\pi} |\langle a | e^{i\vec{k}\cdot\vec{r}} \vec{\varepsilon}_{\vec{k}}^{\lambda} \cdot \vec{\nabla} | b \rangle|^2.\tag{20.32}$$

In order to get quantitative results, we have to calculate the two (identical) matrix elements. As the integration

$$\int d^3r e^{-i\vec{k}\cdot\vec{r}} \Psi_b^*(\vec{r}) \vec{\nabla} \Psi_a(\vec{r})\tag{20.33}$$

with the exponential factor $\exp(-i\vec{k}\cdot\vec{r})$ is too difficult, we simplify the problem. With the energy difference $|E_a - E_b| \sim$ a few eV in a typical atom (the energy of yellow light is about 2 eV), we have a k vector of the radiation of the order of 1000 \AA^{-1} ,

$$\begin{aligned}k = \frac{\omega}{c} &= \frac{|E_b - E_a|}{\hbar c} \sim \left(\frac{2\text{eV}}{6.58 \cdot 10^{-16} \text{eVs}} \right) \cdot \frac{1}{3 \cdot 10^{18} \text{\AA/s}} \\ &\sim \frac{1}{1000 \text{ \AA}},\end{aligned}\tag{20.34}$$

while the size of an atom is of order 1 \AA ; accordingly, we can set the factor $\exp(-i\vec{k}\cdot\vec{r})$ in the volume of the atom (where $r \sim a_B \sim 1 \text{ \AA}$) to the constant $= 1$ and calculate the simpler matrix element on the right hand side

$$|\langle a | e^{i\vec{k}\cdot\vec{r}} \vec{\varepsilon}_{\vec{k}}^{\lambda} \cdot \vec{\nabla} | b \rangle|^2 \approx \frac{1}{\hbar^2} |\langle b | \vec{p} | a \rangle \cdot \vec{\varepsilon}_{\vec{k}}^{\lambda*}|^2.\tag{20.35}$$

In a last step, we make use of the commutation relation $[H_{\text{particle}}, \vec{r}] = -i\hbar\vec{p}/m$ that allows us to replace the momentum \vec{p} with the position vector \vec{r} (we make use of $H_{\text{particle}}|a\rangle = E_a|a\rangle$ and similar for $|b\rangle$),

$$|\langle a | e^{i\vec{k}\cdot\vec{r}} \vec{\varepsilon}_{\vec{k}}^{\lambda} \cdot \vec{\nabla} | b \rangle|^2 \approx \left(\frac{m\omega}{\hbar} \right)^2 |\langle b | \vec{r} | a \rangle \cdot \vec{\varepsilon}_{\vec{k}}^{\lambda*}|^2.\tag{20.36}$$

The following (classical) estimate gives the (qualitatively) same result: we replace the momentum \vec{p} in the coupling H_{int} with the velocity to find $-(e/c)\vec{A} \cdot \partial_t \vec{r}$ and transfer the time derivative to the vector field via partial integration, that leads us to $\sim e[(1/c)\partial_t \vec{A}] \cdot \vec{r} = -e\vec{E} \cdot \vec{r}$. The expression $e\vec{r} \cdot \vec{E}$ is just the energy of the dipole $e\vec{r}$ in the field¹⁰ \vec{E} . This approximation leads us to the electrical dipole transition E1 in the H atom.

Finally, we can rewrite the expression (20.30) as (we define the fine structure constant $\alpha = e^2/\hbar c \approx 1/137$, the small parameter in the perturbation series of QED and write $\vec{r}_{ab} \equiv \langle b|\vec{r}|a\rangle$)

$$W_{ab}^{\text{em,abs}} = 2\alpha \frac{\omega^3}{c^2} \binom{n_k+1}{n_k} \sum_{\lambda} \frac{d\Omega_{\vec{k}}}{4\pi} |\vec{r}_{ab} \cdot \vec{\varepsilon}_{\vec{k}}^{\lambda*}|^2. \quad (20.37)$$

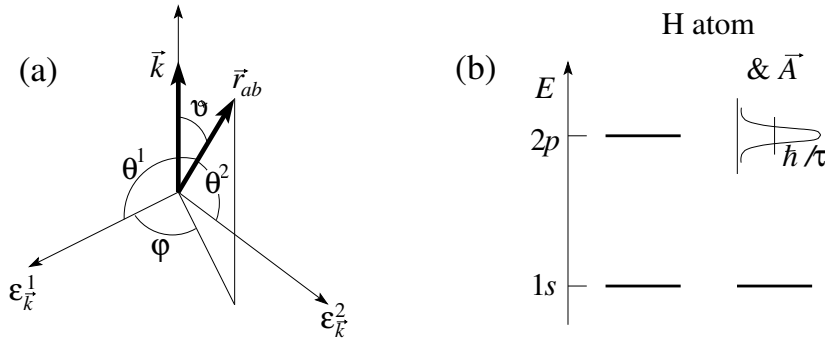


Fig. 20.2: (a) Geometry for the \vec{k} integration in (20.37) with the definition of the angles θ^1 , θ^2 , ϑ , and φ . (b) Broadening of the $2p$ level due to the coupling to the electromagnetic field \vec{A} . The broadening is given (via Heisenberg's uncertainty) by the lifetime τ of the state.

Remains to carry out the integration over the photonic degree of freedom \vec{k}, λ in (20.37). By choosing $|a\rangle$ and $|b\rangle$, \vec{r}_{ab} is fixed in space¹¹. We define the angles θ^1 and θ^2 via (see Fig. 20.2)

$$\begin{aligned} \cos \theta^1 &= \frac{|\vec{r}_{ab} \cdot \vec{\varepsilon}_{\vec{k}}^1|}{r_{ab}}, & \cos \theta^2 &= \frac{|\vec{r}_{ab} \cdot \vec{\varepsilon}_{\vec{k}}^2|}{r_{ab}}, \\ \cos \theta^1 &= \sin \vartheta \cos \varphi, & \cos \theta^2 &= \sin \vartheta \sin \varphi. \end{aligned} \quad (20.38)$$

¹⁰The electric field strength of the photon field can be estimated by $E^2 \sim (\omega/c)^2 [2\pi\hbar c^2/\omega V] n_{\vec{k}\lambda} \sim 2\pi\hbar\omega n_{\vec{k}\lambda}/V$.

¹¹Note that $\vec{r}_{ab} = \exp(i\chi)\vec{R}_{ab}$ with $\vec{R}_{ab} \in \mathbb{R}^3$; we can choose $\vec{r}_{ab} \in \mathbb{R}^3$ and ignore the phase.

The sum over polarizations λ gives the weight $\sin^2 \vartheta$ of a \vec{k} vector and with the integration

$$\int \frac{d\Omega_{\vec{k}}}{4\pi} \sin^2 \vartheta = \frac{2\pi}{4\pi} \int_{-1}^1 dz (1 - z^2) = \frac{2}{3}, \quad (20.39)$$

we arrive at the final formula

$$W_{ab}^{\text{em,abs}} = \frac{4}{3} \alpha \frac{\omega^3}{c^2} |r_{ab}|^2 \begin{pmatrix} n_k + 1 \\ n_k \end{pmatrix}. \quad (20.40)$$

These transition rates define finite lifetimes for the excited states of the (hydrogen) atom, e.g., the $2p$ state decays into the ground state $1s$ (the groundstate remains sharp as it cannot decay). To find a rough estimate of the lifetime of the $2p$ state, we use $\hbar\omega = E_{2p} - E_{1s} = E_R(1 - 1/4) = (3/4)E_R \sim 10 \text{ eV}$, $r_{ab} \sim a_B \approx 0.5 \text{ \AA}$, $\omega/c = k \sim 1/200 \text{ \AA}$, $\hbar = 6.58 \cdot 10^{-16} \text{ eVs}$, $n_k = 0$ for spontaneous emission and using $\alpha = 1/137$, we find

$$W_{ab}^{2p \rightarrow 1s} \simeq \alpha \frac{(10 \text{ eV})}{6.58 \cdot 10^{-16} \text{ eVs}} 6 \cdot 10^{-6} \sim 10^9 \text{ s}^{-1}. \quad (20.41)$$

The finite lifetime $\tau_{2p} = 1/W_{ab} \sim 10^{-9} \text{ s}$ leads to a small relative level broadening,¹²

$$\frac{\Delta E_{2p}}{E_{2p}} \sim \frac{\hbar}{\tau_{2p} E_{2p}} \sim 10^{-7}. \quad (20.42)$$

hence, the $2p$ level still remains sharply defined.

20.2.1 Selection rules

Above, we choose the $2p \rightarrow 1s$ transition on purpose: since the transitions depend on matrix elements $\langle b | \vec{r} | a \rangle$, they are subject to selection rules: The states a and b of the atom are characterized by their orbital angular momenta L_a and L_b ; with the help of the representation of the coordinate \vec{r} in terms of spherical harmonics,

$$\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = r \sqrt{\frac{4\pi}{3}} \begin{pmatrix} -(Y_{1,1} - Y_{1,-1})/\sqrt{2} \\ i(Y_{1,1} + Y_{1,-1})/\sqrt{2} \\ Y_{1,0} \end{pmatrix}, \quad (20.43)$$

¹²This broadening results from the energy smearing $\Delta E \Delta t \sim \hbar$ due to Heisenberg's uncertainty principle.

we define the tensor V_q^1 of rank 1 (a vector)

$$\begin{aligned} V_{q=\pm 1}^1 &= \mp(x \pm iy)/\sqrt{2} = r\sqrt{4\pi/3} Y_{1,\pm 1}, \\ V_{q=0}^1 &= z = r\sqrt{4\pi/3} Y_{1,0}, \end{aligned} \quad (20.44)$$

and rewrite the matrix elements as linear combinations of

$$\langle b; L_b, m_b | V_q^1 | a; L_a, m_a \rangle \stackrel{\text{WET}}{=} \frac{\langle b; L_b || V^1 || a; L_a \rangle}{\sqrt{2L_a + 1}} \underbrace{\langle 1L_a q m_a | 1L_a L_b m_b \rangle}_{\text{CG-coefficient}}. \quad (20.45)$$

The Wigner Eckart Theorem (WET) then allows us to express the dependencies on the angular momentum quantum numbers through Clebsch-Gordon (CG) coefficients $\langle j_1 j_2 m_1 m_2 | j_1 j_2 j m \rangle$. These CG coefficients differ from zero only if the total angular momentum $j = L_b$ of the final state results from a combination of the angular momentum $j_1 = 1$ of the electromagnetic field (that appears via the vector V_q^1 and corresponds to the spin = 1 of the Photon) and the initial state $j_2 = L_a$ of the atom. This gives us the selection rules

$$\begin{aligned} \Delta m &= m_a - m_b = 0, \pm 1, \\ \Delta L &= L_a - L_b = 0, \pm 1, \\ \text{but not: } L_a &= L_b = 0. \end{aligned} \quad (20.46)$$

The formulas (20.46) represent the selection rules for vector operators. In addition, a and b must have different parities

$$\pi_a = -\pi_b, \quad (20.47)$$

which excludes $\Delta L = 0$.

20.2.2 Detailed balance and Bose occupation

We analyze the transition rates $W_{ab}^{\text{em,abs}}$ given by (20.40) and focus on the difference in the factors $n_k + 1$ for emission and n_k for absorption that turns out to be highly relevant. It seems clear that the absorption probability should be proportional to n_k —this is a classically expected result. The emission probability, however, involves two terms, one we call ‘induced emission’ which is proportional to n_k and just compensates for the absorption,

and an additional *spontaneous* emission proportional to 1. An excited state decays with the emission of radiation even if it is not perturbed accordingly! Consider an ensemble of atoms (with levels A above B) in equilibrium with the radiation field γ_k ,

$$A \leftrightarrow B + \gamma_k, \quad E_A - E_B = \hbar\omega_k. \quad (20.48)$$

By the principle of detailed balance, the populations $N(A)$ and $N(B)$ of the two levels are given by

$$N(B)W_{\text{abs}} = N(A)W_{\text{em}} \quad (20.49)$$

and using the expression (20.40) for the rates $W_{ab}^{\text{em,abs}}$, we find the ratio of occupations

$$\frac{N(B)}{N(A)} = \frac{n_k + 1}{n_k}. \quad (20.50)$$

On the other hand, according to Boltzmann

$$\frac{N(B)}{N(A)} = e^{\hbar\omega_k/k_B T} \quad (20.51)$$

and the combination of the result (20.50) from quantum mechanics with the ratio (20.51) from statistical mechanics gives us the occupation number n_k for electromagnetic mode with wavenumber k , the so-called Bose factor

$$n_k = \frac{1}{e^{\hbar\omega_k/k_B T} - 1}. \quad (20.52)$$

A container with reflecting walls and Planck's coal grain is nothing but an experimental implementation of a system in which all electromagnetic modes are in thermodynamic equilibrium¹³. The energy density $u(\omega)d\omega$ in the spectral range $d\omega$ then is given by

$$u(\omega)d\omega = \underbrace{\frac{1}{V}}_{\text{density}} \underbrace{\frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1}}_{\text{energy in mode } \omega} \underbrace{2\left(\frac{L}{2\pi}\right)^3 4\pi k^2 dk}_{\text{\#modes in } d\omega}; \quad (20.53)$$

the factor in the mode number (# modes) takes into account the two polarizations. The expression (20.53) provides us with the spectral energy density $du/d\nu = u(\omega) d\omega/d\nu$ (Planck's formula)

$$u(\nu) = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{h\nu/k_B T} - 1}. \quad (20.54)$$

¹³A coal grain is a body that contains all possible two-level systems with arbitrary energy separation $E_A - E_B \in \mathbb{R}$; 'black charcoal' absorbs and re-emits every mode.

20.2.3 Higher-order transitions

If the E1 dipole-transition is forbidden by symmetry, $\vec{r}_{ab} = 0$, then we have to consider further terms in the expansion of $e^{-i\vec{k}\cdot\vec{r}}$ in the matrix element of (20.30),

$$e^{-i\vec{k}\cdot\vec{r}} \approx 1 - i\vec{k}\cdot\vec{r} - \frac{(\vec{k}\cdot\vec{r})^2}{2} + \dots \quad (20.55)$$

The 1 in this expansion generates the E1-transition. The next term $-i\vec{k}\cdot\vec{r}$ generates electrical quadrupole and magnetic dipole transitions. To find them, we have to calculate matrix elements of the form

$$\langle b | (\vec{k}\cdot\vec{r})(\vec{\varepsilon}_k^{\lambda*}\cdot\vec{p}) | a \rangle, \quad (20.56)$$

which we can decompose as follows,

$$\begin{aligned} \langle b | (\vec{k}\cdot\vec{r})(\vec{\varepsilon}_k^{\lambda*}\cdot\vec{p}) | a \rangle &= \frac{1}{2} \langle b | (\vec{k}\cdot\vec{r})(\vec{\varepsilon}\cdot\vec{p}) + (\vec{k}\cdot\vec{p})(\vec{\varepsilon}\cdot\vec{r}) | a \rangle \\ &+ \frac{1}{2} \langle b | (\vec{k}\cdot\vec{r})(\vec{\varepsilon}\cdot\vec{p}) - (\vec{k}\cdot\vec{p})(\vec{\varepsilon}\cdot\vec{r}) | a \rangle \\ &= \frac{1}{2} \langle b | \underbrace{\vec{k}\cdot(\vec{r}\vec{p} + \vec{p}\vec{r})}_{\text{symm. tensor}} \vec{\varepsilon} | a \rangle + \frac{1}{2} \langle b | \underbrace{(\vec{k} \wedge \vec{\varepsilon})}_{\rightarrow \vec{\nabla} \wedge \vec{A} = \vec{B}} \cdot \underbrace{(\vec{r} \wedge \vec{p})}_{\vec{L}} | a \rangle. \end{aligned}$$

The term $\propto \vec{B} \cdot \vec{L}$ has to be combined with the spin-term $g(e/2mc)\vec{B} \cdot \vec{S}$, hence, we have to substitute $\vec{L} \rightarrow \vec{L} + 2\vec{S}$ and obtain the magnetic-dipole transition M1 with the selection rules $|J_b - J_a| \leq 1$ for the angular momentum and $\pi_a = \pi_b$ for the parity (the transition with $J_b = J_a = 0$ is prohibited).

For the tensorial component $\propto (\vec{r}\vec{p} + \vec{p}\vec{r})$, we rewrite

$$\begin{aligned} (\vec{r}\vec{p} + \vec{p}\vec{r}) &= \frac{im}{\hbar} [H_0, \vec{r}\vec{r}] \\ \rightarrow \frac{1}{2} \langle b | \vec{k} \cdot (\vec{r}\vec{p} + \vec{p}\vec{r}) \vec{\varepsilon} | a \rangle &= -\frac{im\omega}{2} \vec{k} \cdot \langle b | \vec{r}\vec{r} | a \rangle \cdot \vec{\varepsilon}, \quad (20.57) \end{aligned}$$

and obtain the electronic quadrupole transition E2 with the selection rules $|L_b - L_a| \leq 2 \leq L_a + L_b$ and $\pi_a = \pi_b$.

The lifetime of states a with forbidden E1-transitions is much longer, e.g., the factor $\vec{k}\cdot\vec{r}$ produces a reduction in the E2-transition

$$(\vec{k}\cdot\vec{r})^2 \sim \left(\frac{a_B}{\lambda}\right)^2 \sim 10^{-6}, \quad (20.58)$$

which gives us a lifetime in the ms range. An extreme example is the $2s \rightarrow 1s$ decay which is highly forbidden by symmetries: E1 is prohibited by parity, M1 makes no contribution to non-relativistic wave functions, E2 is prohibited by angular momentum. The decay then occurs via 2-photon emission and results in a long lifetime $\tau \approx 0.15$ s.

20.3 Light scattering

Photons per se do not interact and hence cannot scatter—in order to have light scattering, we need the coupling to matter. In discussing the scattering of light, we consider scattering process of photons $\vec{k}\lambda \rightarrow \vec{k}'\lambda'$. We therefore choose, without loss of generality

$$\begin{aligned} |i\rangle &= |a; \vec{k}\lambda\rangle, \\ |f\rangle &= |b; \vec{k}'\lambda'\rangle, \end{aligned} \quad (20.59)$$

which means $n_{\vec{k}\lambda} = 1$ or $= 0$, with the value 1 assumed only for the scattering photons. The quantum numbers a and b are those of the hydrogen-like atom (the matter providing the coupling) involved in the scattering process. The interaction Hamiltonian H_{int} as described in (20.15) can connect $|i\rangle$ and $|f\rangle$ in the photon sector via the diamagnetic term A^2 in first-order and via the paramagnetic term $\vec{A} \cdot \vec{p}$ in second-order: pairings $a_{\vec{k}\lambda} a_{\vec{k}'\lambda'}^\dagger$ and $a_{\vec{k}'\lambda'}^\dagger a_{\vec{k}\lambda}$ appear in both cases and transfer $|\vec{k}\lambda\rangle$ to $|\vec{k}'\lambda'\rangle$. There are three relevant processes which can be represented by Feynman graphs as shown in Fig. 20.3. Process A is to be calculated in first-order perturbation theory in A^2 , while processes B and C are second-order in $\vec{A} \cdot \vec{p}$.

First-order A^2 term

In first-order in A^2 , we obtain the following expression for the amplitude $\langle f | \Psi(t) \rangle$ [see (20.25) and (20.26)]

$$\begin{aligned} \langle f | \Psi(t) \rangle &\approx \frac{1}{i\hbar} \int_{t_0}^t dt_1 e^{-i(E_b + \hbar\omega_{k'})(t-t_1)/\hbar} \left\{ \frac{e^2}{2mc^2} \left(\frac{2\pi\hbar c^2}{V\sqrt{\omega_k\omega_{k'}}} \right) \right. \\ &\quad \times \langle b; \vec{k}'\lambda' | (a_{\vec{k}\lambda} a_{\vec{k}'\lambda'}^\dagger + a_{\vec{k}'\lambda'}^\dagger a_{\vec{k}\lambda}) \vec{\varepsilon}_{\vec{k}}^\lambda \cdot \vec{\varepsilon}_{\vec{k}'}^{\lambda'*} e^{i(\vec{k}-\vec{k}') \cdot \vec{r}} | a; \vec{k}\lambda \rangle \Big\} \\ &\quad \times e^{-i(E_a + \hbar\omega_k)(t_1-t_0)/\hbar}. \end{aligned} \quad (20.60)$$

The matrix element in (20.60) factorises with a particle component $\langle a | b \rangle = \delta_{ab}$ and a photon component $\propto \vec{\varepsilon}_{\vec{k}}^\lambda \cdot \vec{\varepsilon}_{\vec{k}'}^{\lambda'*}$.

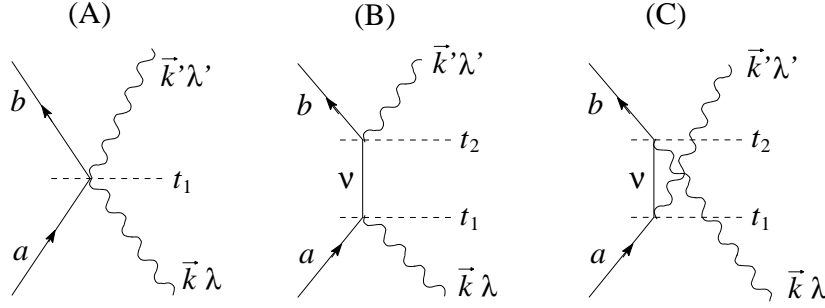


Fig. 20.3: Feynman diagrams for light scattering: (A): first-order process in A^2 , ‘seagull graph’, with a 4-vertex with two electron lines and two photon lines. (B) and (C): second-order processes in $\vec{A} \cdot \vec{p}$ with two 3-vertices and only one photon line. (B) absorption before emission, (C) emission before absorption. In general, the shape of the vertices (here, with $2_{e^-} + 2_{h\nu} = 4$ legs and with $2_{e^-} + 1_{h\nu} = 3$ legs) is determined by the operators/wavefunctions appearing in the interaction term.

Second-order in $\vec{A} \cdot \vec{p}$

The second-order perturbation theory in $\vec{A} \cdot \vec{p}$ is somewhat more complex. The product $(\vec{A} \cdot \vec{p})_1 (\vec{A} \cdot \vec{p})_2$ describes processes in which $(\vec{A} \cdot \vec{p})_1$ first annihilates the photon $|\vec{k}\lambda\rangle$ at t_1 , and $(\vec{A} \cdot \vec{p})_2$ creates the photon $|\vec{k}'\lambda'\rangle$ at $t_2 > t_1$. Alternatively, $(\vec{A} \cdot \vec{p})_1$ can first create the photon $|\vec{k}'\lambda'\rangle$ at t_1 and the factor $(\vec{A} \cdot \vec{p})_2$ annihilates the photon $|\vec{k}\lambda\rangle$ at the later time $t_2 > t_1$. The amplitude for these processes is given by

$$\begin{aligned}
 \langle f | \Psi(t) \rangle &\approx \frac{1}{(i\hbar)^2} \int_{t_0}^t dt_2 \int_{t_0}^{t_2} dt_1 e^{-i(E_b + \hbar\omega_{k'})(t-t_2)/\hbar} \frac{e^2}{m^2 c^2} \frac{2\pi\hbar c^2}{V\sqrt{\omega_k \omega_{k'}}} \sum_{\nu} \\
 &\left[\langle b; \vec{k}'\lambda' | a_{\vec{k}'\lambda'}^\dagger \vec{\varepsilon}_{\vec{k}'}^{\lambda'*} \cdot \vec{p} e^{-i\vec{k}' \cdot \vec{r}} | \nu; 0 \rangle e^{-iE_\nu(t_2-t_1)/\hbar} \langle \nu; 0 | a_{\vec{k}\lambda} \vec{\varepsilon}_{\vec{k}}^\lambda \cdot \vec{p} e^{i\vec{k} \cdot \vec{r}} | a; \vec{k}\lambda \rangle \right. \\
 &+ \langle b; \vec{k}'\lambda' | a_{\vec{k}\lambda} \vec{\varepsilon}_{\vec{k}}^\lambda \cdot \vec{p} e^{i\vec{k} \cdot \vec{r}} | \nu; \vec{k}\lambda, \vec{k}'\lambda' \rangle e^{-i(E_\nu + \hbar\omega_k + \hbar\omega_{k'})(t_2-t_1)/\hbar} \\
 &\quad \left. \times \langle \nu; \vec{k}\lambda, \vec{k}'\lambda' | a_{\vec{k}'\lambda'}^\dagger \vec{\varepsilon}_{\vec{k}'}^{\lambda'*} \cdot \vec{p} e^{-i\vec{k}' \cdot \vec{r}} | a; \vec{k}\lambda \rangle \right] \\
 &\times e^{-i(E_a + \hbar\omega_k)(t_1-t_0)/\hbar}. \quad (20.61)
 \end{aligned}$$

Formula (20.61) has a simple structure, with a free propagation between t_0 and t_1 followed by a process, then again a free propagation with an excited state ν between t_1 and t_2 , then another process and finally a free propagation

between t_2 and t , see the Feynman graphs in Fig. 20.3 and the following schematic representation,

$$\begin{aligned} \langle f | \Psi(t) \rangle \approx & \int_{t_0}^t dt_2 \int_{t_0}^{t_2} dt_1 e^{-i(t-t_2)\cdots} \sum_{\nu} \left(\underbrace{\langle \cdot | \cdot | \nu \rangle e^{-i(t_2-t_1)\cdots} \langle \nu | \cdot | \cdot \rangle}_{(B)} \right. \\ & \left. + \underbrace{\langle \cdot | \cdot | \nu \rangle e^{-i(t_2-t_1)\cdots} \langle \nu | \cdot | \cdot \rangle}_{(C)} \right) e^{-i(t_1-t_0)\cdots}. \end{aligned} \quad (20.62)$$

The integration over t_1 in (20.61) generates the energy denominations $(E_{\nu} - E_a - \hbar\omega_k)^{-1}$ and $(E_{\nu} - E_a + \hbar\omega_{k'})^{-1}$, and with $\exp(-i\vec{k}' \cdot \vec{r}) \approx 1$, $\exp(-i\vec{k} \cdot \vec{r}) \approx 1$, we obtain

$$\begin{aligned} \langle f | \Psi(t) \rangle \approx & -\frac{1}{i\hbar} \int_{t_0}^t dt_2 e^{-i(E_b + \hbar\omega_{k'})(t-t_2)/\hbar} \frac{e^2}{m^2 c^2} \frac{2\pi\hbar c^2}{V\sqrt{\omega_k\omega_{k'}}} \\ & \times \sum_{\nu} \left[\frac{\langle b | \vec{\varepsilon}_{\vec{k}'}^{\lambda'*} \cdot \vec{p} | \nu \rangle \langle \nu | \vec{\varepsilon}_{\vec{k}}^{\lambda} \cdot \vec{p} | a \rangle}{E_{\nu} - E_a - \hbar\omega_k} + \frac{\langle b | \vec{\varepsilon}_{\vec{k}}^{\lambda} \cdot \vec{p} | \nu \rangle \langle \nu | \vec{\varepsilon}_{\vec{k}'}^{\lambda'*} \cdot \vec{p} | a \rangle}{E_{\nu} - E_a + \hbar\omega_{k'}} \right] \\ & \times e^{-i(E_a + \hbar\omega_k)(t_2-t_0)/\hbar}. \end{aligned} \quad (20.63)$$

We have to add the amplitudes (20.60) and (20.63) coherently (we replace t_2 in (20.63) with t_1 and write both expressions under one time integral over t_1 , adding both amplitudes). Then, we square to find the probability of the three interfering processes and divide by t to obtain the transition rate. We integrate the resulting δ -function $\delta(E_b + \hbar\omega_{k'} - E_a - \hbar\omega_k)$ over the continuum of the photons in the final state [accordingly, energy conservation enforces that $\hbar\omega_{k'} \rightarrow \hbar\omega - (E_b - E_a)$] and find for the states scattered into $d\Omega_{\vec{k}'}$ (check that all terms have the same dimensionality)

$$\begin{aligned} W_{ab,d\Omega} = & \frac{2\pi}{\hbar} \left(\frac{2\pi\hbar c^2}{V\sqrt{\omega\omega'}} \right)^2 \left(\frac{e^2}{mc^2} \right)^2 \frac{V}{2\pi^2} \frac{\omega'^2}{\hbar c^3} \frac{d\Omega_{\vec{k}'}}{4\pi} \\ & \times \left| \delta_{ab} \vec{\varepsilon}_{\vec{k}}^{\lambda} \cdot \vec{\varepsilon}_{\vec{k}'}^{\lambda'*} - \frac{1}{m} \sum_{\nu} \left[\frac{\langle b | \vec{\varepsilon}_{\vec{k}'}^{\lambda'*} \cdot \vec{p} | \nu \rangle \langle \nu | \vec{\varepsilon}_{\vec{k}}^{\lambda} \cdot \vec{p} | a \rangle}{E_{\nu} - E_a - \hbar\omega} \right. \right. \\ & \left. \left. + \frac{\langle b | \vec{\varepsilon}_{\vec{k}}^{\lambda} \cdot \vec{p} | \nu \rangle \langle \nu | \vec{\varepsilon}_{\vec{k}'}^{\lambda'*} \cdot \vec{p} | a \rangle}{E_{\nu} - E_a + \hbar\omega'} \right] \right|^2. \end{aligned} \quad (20.64)$$

The incident photon current density is $j = nc = c/V$ and we simplify the the pre-factor according to

$$\frac{2\pi}{\hbar} \left(\frac{2\pi\hbar c^2}{V\sqrt{\omega\omega'}} \right)^2 \left(\frac{e^2}{mc^2} \right)^2 \frac{V}{2\pi^2} \frac{\omega'^2}{\hbar c^3} \frac{1}{4\pi} \frac{V}{c} = \frac{e^4}{m^2 c^4} \frac{\omega'}{\omega} \quad (20.65)$$

using the classic electron radius

$$r_0 = \frac{e^2}{mc^2} \approx 2.8 \cdot 10^{-13} \text{ cm.} \quad (20.66)$$

The final result for the differential cross-section is the Kramers-Heisenberg formula for the light scattering,

$$\begin{aligned} \frac{d\sigma}{d\Omega} = r_0^2 \frac{\omega'}{\omega} \left| \delta_{ab} \vec{\varepsilon}_k^\lambda \cdot \vec{\varepsilon}_{k'}^{\lambda'*} - \frac{1}{m} \sum_{\nu} \left[\frac{\langle b | \vec{\varepsilon}_{k'}^{\lambda'*} \cdot \vec{p} | \nu \rangle \langle \nu | \vec{\varepsilon}_k^\lambda \cdot \vec{p} | a \rangle}{E_\nu - E_a - \hbar\omega_k} \right. \right. \\ \left. \left. + \frac{\langle b | \vec{\varepsilon}_k^\lambda \cdot \vec{p} | \nu \rangle \langle \nu | \vec{\varepsilon}_{k'}^{\lambda'*} \cdot \vec{p} | a \rangle}{E_\nu - E_a + \hbar\omega_{k'}} \right] \right|^2. \end{aligned} \quad (20.67)$$

As an exercise, calculate the Rayleigh, Thomson, and Raman scattering with the help of (20.67). We only give a brief discussion here.

20.3.1 Rayleigh scattering

Rayleigh scattering deals with the elastic scattering of long wavelength light from atoms; thus the energies of the initial and final states are fixed, with $a = b$, $\hbar\omega = \hbar\omega' \ll \hbar\omega_{\nu a} = E_\nu - E_a > 0$. All three terms in (20.67) are relevant (the first term is exactly compensated by part of the second term) and we obtain the famous ω^4 law (we use the relation $[H, \vec{r}] = -i\hbar\vec{p}/m$, see (20.36))

$$\frac{d\sigma}{d\Omega} = \left(\frac{r_0 m}{\hbar} \right)^2 \omega^4 \left| \sum_{\nu} \frac{(\vec{r} \cdot \vec{\varepsilon}_{k'}^{\lambda'*})_{a\nu} (\vec{r} \cdot \vec{\varepsilon}_k^\lambda)_{\nu a} + (\vec{r} \cdot \vec{\varepsilon}_k^\lambda)_{a\nu} (\vec{r} \cdot \vec{\varepsilon}_{k'}^{\lambda'*})_{\nu a}}{\omega_{\nu a}} \right|^2 \quad (20.68)$$

which explains why the sky is blue and the evening sun is red. With a dimensional estimate, we find that (with $mr_{\nu a}^2/\hbar^2 \sim 1/\hbar\omega_{\nu a}$)

$$\frac{d\sigma}{d\Omega} \sim r_0^2 \frac{\omega^4}{\omega_{\nu a}^4}.$$

20.3.2 Thomson scattering

Thomson scattering involves high-energy photons that elastically scatter from electrons. Since $\hbar\omega \gg$ binding energy, the binding of the electrons to the atoms is irrelevant and the scatterers can be described as free electrons.

Since $\hbar\omega \gg (\vec{\varepsilon}_{\vec{k}'}^{\lambda'*} \cdot \vec{p})_{a\nu} (\vec{\varepsilon}_{\vec{k}}^{\lambda} \cdot \vec{p})_{\nu a} / m$, only the first term is relevant (regardless of the atomic structure and the binding of the electrons) and we obtain

$$\frac{d\sigma}{d\Omega} = r_0^2 \left| \vec{\varepsilon}_{\vec{k}}^{\lambda} \cdot \vec{\varepsilon}_{\vec{k}'}^{\lambda'*} \right|^2. \quad (20.69)$$

The integration over the angles and the sum over polarizations gives the total cross section

$$\sigma_{\text{Thomson}} = \frac{8\pi r_0^2}{3} \approx 6.65 \cdot 10^{-25} \text{ cm}^2 \sim 1 \text{ b}. \quad (20.70)$$

Note that $\hbar\omega \ll mc^2$ must be satisfied, because we treated the electrons nonrelativistically.

20.3.3 Raman scattering

Raman scattering describes the inelastic scattering of light on atoms (in general on solids). The light can lose energy by exciting atoms (this leads to the Stokes line with the light shifted towards the red side in the spectrum) or gain energy as the atoms transfer energy to the photon by relaxing an excited state (the anti-Stokes line with blue-shifted light). In a solid, phonons play the role of the transitions in the atoms (the phonon generation leads to the Stokes line, phonon absorption to the anti-Stokes line). The intensities and scattering diagrams in the atom/solid body are sketched in figure 20.4.

20.4 Lamb-shift

We have already treated some of the additional terms appearing in the Hamiltonian H of the H atom in perturbation theory within the quantum mechanics I lecture. Here, we reconsider this Hamiltonian and further add the coupling to the electromagnetic field H_L^F ,

$$H = \underbrace{H_0}_{\frac{p^2}{2m} - \frac{e^2}{r}} + \underbrace{H_{\text{kin}}^R}_{\frac{1}{2m^2c^2} \frac{1}{r} \partial_r V \vec{L} \cdot \vec{S}} + \underbrace{H_{\text{SO}}^R + H_D^R}_{\frac{\hbar^2}{8m^2c^2} \nabla^2 V} + \underbrace{H_L^F}_{-\vec{\mu}_I \cdot \vec{B}_J} + \underbrace{H_{\text{IJ}}^N}_{\frac{e}{mc} \vec{A}(\vec{r}, t) \cdot \vec{p}}. \quad (20.71)$$

The individual terms can each be assigned to a shift in the energy level, as sketched in Fig. 20.5. The indices stand for

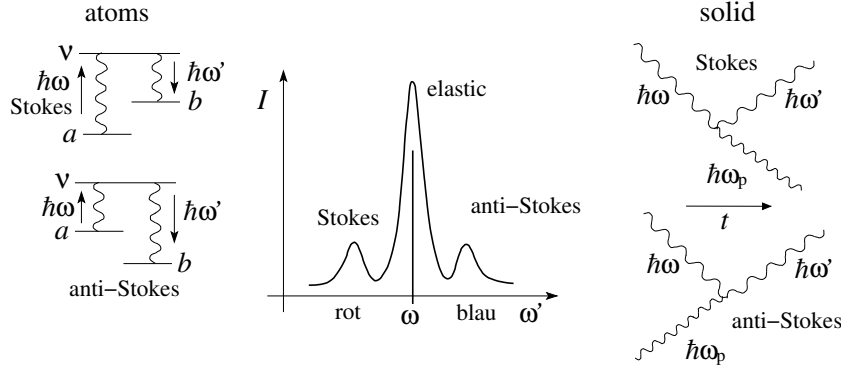


Fig. 20.4: Light scattering via the Raman process. On the left the processes of light scattering on atoms, on the right the Feynman graphs for light scattering on the solid with phonon emission (Stokes) and phonon absorption (anti-Stokes) are shown.

R: relativistic corrections (SO = spin-orbit, D = Darwin),

F: field theoretical corrections due to the coupling of the electron field to the photon field (L = Lamb-shift),

N: Correction based on the nuclear magnetic field = hyperfine interaction. Here, \vec{B}_J denotes the field which is generated by the electrons at the location of the nucleus (with \vec{B}_J including a spin and an angular momentum component) and $\vec{\mu}_I$ is the magnetic moment of the nucleus. Given the coupling between \vec{I} and \vec{J} , only the total angular momentum $\vec{F} = \vec{J} + \vec{I}$ of the atom is sharply defined. In the hydrogen atom, the splitting is maximal for s states whose orbitals have finite weight at the location of the nucleus.

The Lamb-shift was discovered in the late forties. It was precisely measured by means of the then new (due to the invention of the radar) microwave technology by Lamb and Retherford in 1947; the theoretical description in terms of electromagnetic field fluctuations was carried out in the same year by Hans Bethe. In second-order perturbation theory, we have to calculate the following correction to the energy due to the interaction between radiation and matter,

$$\Delta E_\alpha = \langle \alpha | H_{\text{int}} | \alpha \rangle + \sum_{\nu \neq \alpha} \frac{|\langle \nu | H_{\text{int}} | \alpha \rangle|^2}{E_\alpha - E_\nu}. \quad (20.72)$$

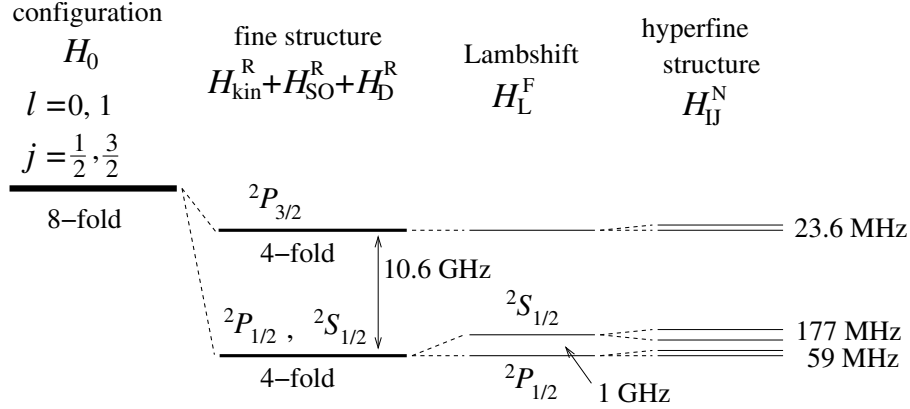


Fig. 20.5: Electronic level shifts in the $n = 2$ level (with $2s, 2p, 2 \cdot 2^2 = 8$ -fold degenerate, including spin) due to relativistic corrections (kin = kinetic correction, SO = spin-orbit coupling, D = Darwinterm), field theoretical corrections (L = Lamb-shift), and corrections due to the magnetic field of the nucleus. The experimentally measured Lamb-shift measures ≈ 1 GHz.

Here (with $\beta = \alpha, \nu$), $|\beta\rangle = |b; n_1, \dots, n_\infty\rangle$ is a particle (b) field (n_1, \dots, n_∞) state with associated energy E_β (more precisely, the energy of particle and radiation without interaction) and H_{int} is the interaction Hamiltonian (20.4). The first-order correction $\langle \alpha | H_{\text{int}} | \alpha \rangle$ vanishes: Since only the A^2 part of H_{int} has terms that preserve the number of photons it is the only one contributing. But as H_{int} contains the normal-ordered expression of A^2 , its expected vacuum value is zero (we are interested in states $|\alpha\rangle$ where the electromagnetic field is not excited, $|\alpha\rangle = |a; n_1, \dots, n_\infty\rangle = |a; 0\rangle$),

$$\langle a; 0 | : A^2(\vec{r}, t) : | a; 0 \rangle = 0. \quad (20.73)$$

The 2nd order term with $|\nu\rangle = |v; 1_{\vec{k}\lambda}\rangle$ remains, a 1 photon state which couples to $|\alpha\rangle = |a; 0\rangle$ via the $\vec{A} \cdot \vec{p}$ -term in H_{int} . Both terms can be represented graphically as shown in Figs. 20.6 (a) and (b), whereby only the process for the Lamb shift 20.6 (b) is relevant.

We already encountered the matrix elements $\langle a; 0 | H_{\text{int}} | v; 1_{\vec{k}\lambda} \rangle$ on page 200; up to a trivial phase $\exp(i\varphi)$, it is

$$\langle a; 0 | H_{\text{int}} | v; 1_{\vec{k}\lambda} \rangle = e^{i\varphi} \frac{e\hbar}{m} \sqrt{\frac{2\pi\hbar}{V\omega_k}} \langle a | e^{-i\vec{k} \cdot \vec{r}} \vec{\varepsilon}_{\vec{k}}^{\lambda*} \cdot \vec{\nabla} | v \rangle. \quad (20.74)$$

The associated energy denominators are $E_a - E_v - \hbar\omega_k$ for photon emission

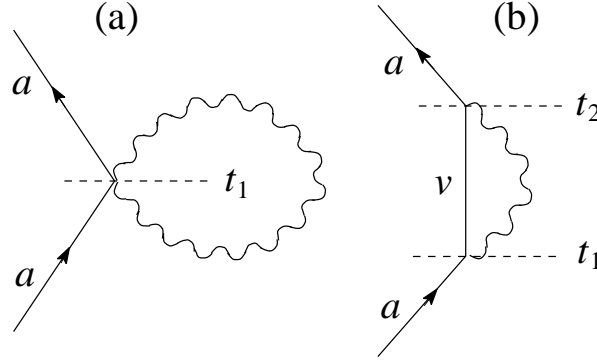


Fig. 20.6: Feynman diagrams for the first-order A^2 Term (a) and the second-order $\vec{A} \cdot \vec{p}$ term in H_{int} (b). The latter is responsible for the Lamb-shift.

and with $\exp(-i\vec{k} \cdot \vec{r}) \sim 1$, we obtain (with $|\alpha\rangle = |a; 0\rangle$)

$$\Delta E_a = \sum_{v\vec{k}\lambda} \frac{2\pi\hbar}{V\omega_k} \frac{e^2\hbar^2}{m^2} \frac{|\vec{\varepsilon}_k^{\lambda*} \cdot \vec{\nabla}|_{av}^2}{E_a - E_v - \hbar\omega_k} \quad (20.75)$$

$$\begin{aligned} &= \oint_v \frac{d\Omega_{\vec{k}}}{4\pi} \frac{\omega^2 d\omega}{2\pi^2 c^3} \frac{2\pi\hbar}{\omega} \frac{e^2\hbar^2}{m^2} \sum_{\lambda} \frac{|\vec{\varepsilon}_k^{\lambda*} \cdot \vec{\nabla}|_{av}^2}{E_a - E_v - \hbar\omega} \\ &= \oint_v \frac{d\hbar\omega}{2\pi} \frac{\hbar W_{av}(\omega)}{E_a - E_v - \hbar\omega}, \end{aligned} \quad (20.76)$$

with the transition rate $W_{av} \propto \omega$ from a to v for spontaneous emission.¹⁴ In formula (20.76), we have to specify how the singularity at $E_a = E_v + \hbar\omega$ is to be treated. By adiabatically turning H_{int} on and off using a factor $\exp(-\varepsilon|t|)$, we obtain the causal regularization $E_a - E_v - \hbar\omega \rightarrow E_a - E_v - \hbar\omega + i\varepsilon$ with $\varepsilon \rightarrow 0^+$. Using the Sokhotski formula $1/(x + i\varepsilon) = P(1/x) - i\pi\delta(x)$ one obtains

$$\Delta E_a = \underbrace{\oint_v \frac{d\hbar\omega}{2\pi} P \frac{\hbar W_{av}(\omega)}{E_a - E_v - \hbar\omega}}_{\Delta E_a^R} - i \underbrace{\frac{\hbar}{2} \sum_{v, E_v < E_a} W_{av}(\omega = (E_a - E_v)/\hbar > 0)}_{\Gamma/2}. \quad (20.77)$$

¹⁴We keep the matrix element in the form $|\vec{\varepsilon}_k^{\lambda*} \cdot \vec{\nabla}|_{av}^2$ and do not replace $\hbar\vec{\nabla}$ with $\sim m\omega\vec{r}$.

The formula (20.77) leads to a very elegant interpretation: Due to the interaction with the field, two types of transitions are possible. If the energy is conserved, that is, for every combination ω , E_v with $E_a - E_v = \hbar\omega$, a transition $a \rightarrow v + \gamma$ is allowed and the state a decays into $v + \gamma$. This gives a a finite lifetime $\tau = 1/\Gamma$, which is determined by the total rate $\sum_v W_{av}$. In fact, we find for the dynamics of a

$$e^{-i(E_a + \Delta E_a)t/\hbar} = e^{-i(E_a + \Delta E_a^R)t/\hbar} e^{-\Gamma t/2} \quad (20.78)$$

and the probability of finding the state a decreases according to $\exp(-\Gamma t)$. This process is graphically illustrated in Fig. 20.7(a).

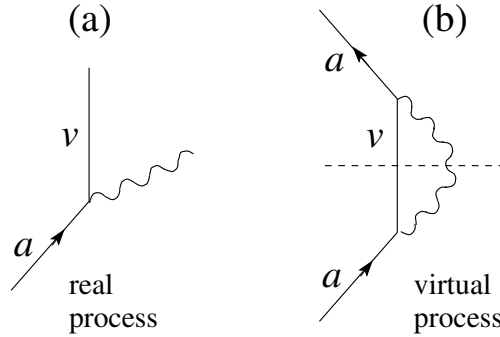


Fig. 20.7: Real (a) and virtual process (b) of a photon emission. Both processes result from second-order perturbation theory with the energy denominator $1/(\Delta E + i\varepsilon)$. Sokhotsky's formula generates two terms from this factor, the principal value P associated with the virtual process that creates an energy shift and the δ -function that belongs to the real and energy-preserving process that defines the lifetime. The graph of the real process (a) results from the virtual process (b) by halving (dashed line).

In addition to the decay of a , we also find a shift ΔE_a^R in its energy; this is the Lamb-shift. The processes that generate ΔE_a^R need not be energy-conserving: All $E_a - E_v - \hbar\omega$ are relevant. Processes that do not conserve the energy¹⁵ are called virtual and can be represented graphically as shown in Fig. 20.7 (b). The diagram 20.7(a) is just 'half' of the diagram 20.7(b).

Next, we determine the Lamb-shift ΔE_a^R . This is a tricky and interesting calculation because the principal value (20.77) with $W_{av}(\omega) \propto \omega$ is linearly

¹⁵According to Heisenberg uncertainty principle, energy conservation can be violated during the time $\tau = \hbar/(E_a - E_b - \hbar\omega)$.

divergent in the cutoff Ω ,

$$\Delta E_a^R \propto \oint_0^\Omega \frac{d\omega}{2\pi} \frac{\omega}{E_a - E_v - \hbar\omega} \sim \Omega. \quad (20.79)$$

We can argue that our non-relativistic description of the electron is no longer applicable at high energies and thus $\hbar\Omega \sim mc^2$ generates a natural cutoff. What is unfortunate about this result, however, is that it depends linearly on the selected cutoff. Bethe solved the problem (at Kramers' suggestion) as follows: observable as an energy shift is not directly ΔE_a^R , but the difference between ΔE_a^R and ΔE_f with ΔE_f the corresponding shift for a free electron. Because also a free electron surrounds itself with a virtual photon cloud and experiences (as we will see) a mass renormalization. The process can be illustrated as in Fig. 20.8.

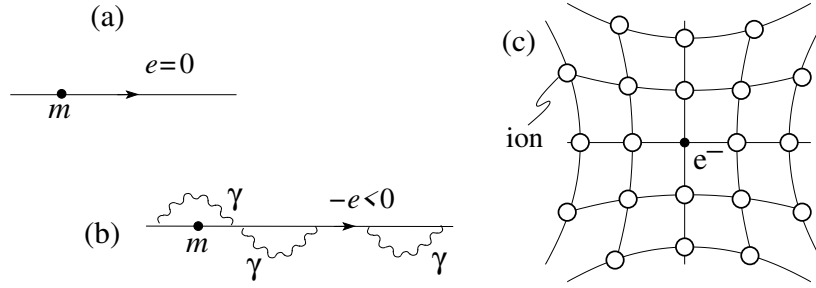


Fig. 20.8: (a) ‘Naked’ electron in a world with charge $e = 0$, so that there is no coupling to the radiation field. (b) ‘Dressed’ (with a photon cloud) electron. (c) The same process produces the polarons in the solid, electrons surrounded by virtual phonons. This phonon cloud corresponds to an entrained deformation of the ions and increases the mass of the polaron as compared to the mass of the electron.

The same effect also occurs in solids: instead of virtual photons, the electron generates a cloud of virtual phonons in the solid and thus becomes heavier (polaron): the electron gets mass from the ions by deformation of the lattice and, because it must drag this deformation along, heavier¹⁶. The interaction which produces the effect is the electron–phonon interaction which takes the place of the electron–photon interaction in the electromagnetic vacuum, so there is the following analogy: the photon cloud around an electron,

¹⁶With λ , the dimensionless electron-phonon coupling constant, the mass increases by λm , $m^* = m(1 + \lambda)$.

generated by the coupling to the radiation field, corresponds to a phonon cloud (a lattice deformation) around the electron, generated by the electron-ion coupling.

Next, we show that electrons are heavier due to the photon cloud carried along. For this, we consider the energy shift of a free electron due to its coupling to the radiation field (we change the parameters in order to adapted them to the free-particle situation)

$$\Delta E_f = \frac{4\alpha}{3} \lambda_c^2 \sum_v \frac{d\hbar\omega}{2\pi} P \frac{\hbar\omega |(e^{-i\vec{k}\cdot\vec{r}} \vec{\nabla})_{va}|^2}{(p_a^2 - p_v^2)/2m - \hbar\omega}, \quad (20.80)$$

with the Compton wavelength $\lambda_c = \lambda/2\pi = \hbar/mc \approx 3.863 \cdot 10^{-3} \text{ \AA}$. The expression for the matrix element takes the form

$$(e^{-i\vec{k}\cdot\vec{r}} \vec{\nabla})_{va} = \int \frac{d^3r}{V} e^{-i\vec{k}_v\cdot\vec{r}} e^{-i\vec{k}\cdot\vec{r}} \vec{\nabla} e^{i\vec{k}_a\cdot\vec{r}} = i\vec{k}_a \delta_{\vec{k}_a, \vec{k} + \vec{k}_v} \quad (20.81)$$

and the energy denominator can be rewritten using momentum conservation

$$\begin{aligned} p_a^2 - p_v^2 &= \hbar^2 [k_a^2 - (\vec{k}_a - \vec{k})^2] = 2\hbar^2 \vec{k}_a \cdot \vec{k} - \hbar^2 k^2, \\ (p_a^2 - p_v^2)/2m - \hbar\omega &= \underbrace{\frac{\hbar^2 \vec{k}_a \cdot \vec{k}}{m}}_{\sim v_a \hbar\omega/c} - \hbar\omega \left(1 + \underbrace{\frac{\hbar\omega}{2mc^2}}_{\ll 1}\right) \approx -\hbar\omega. \end{aligned} \quad (20.82)$$

If we combine (20.80) – (20.82), we find (with $\sum_v \delta_{\vec{k}_a, \vec{k} + \vec{k}_v} = 1$)

$$\begin{aligned} \Delta E_f &\approx -\frac{4\alpha}{3} (\lambda_c k_a)^2 \underbrace{\int \frac{d\hbar\omega}{2\pi} \frac{\hbar\omega}{\hbar\omega}}_{\int_0^{\hbar\Omega} \frac{d\hbar\omega}{2\pi} = \frac{\hbar\Omega}{2\pi}} \\ &= -C p^2 \end{aligned} \quad (20.83)$$

with

$$C \stackrel{\hbar\Omega \sim mc^2}{=} \frac{4\alpha}{3} \frac{mc^2}{2\pi} \frac{\lambda_c^2}{\hbar^2} = \frac{2}{3\pi} \frac{\alpha}{m}.$$

The energy shift is quadratic in the momentum and thus defines a mass renormalization according to $E_f = p^2/2m + \Delta E_f \equiv p^2/2m^*$ with

$$\frac{1}{m^*} = \frac{1}{m} \left(1 - \underbrace{\frac{4\alpha}{3\pi}}_{0.005}\right)$$

$$\approx \frac{1}{m(1 + 4\alpha/3\pi)},$$

$$m^* = m \left(1 + \frac{4\alpha}{3\pi} \right). \quad (20.84)$$

The ‘naked’ mass m of the electron cannot be observed, since we cannot eliminate the coupling of the electron to the electromagnetic field. Only the *renormalized* mass m^* increased by the photon cloud is observable. Note: It is the strong electric field around the electron that leads to spontaneous emission of virtual photons (and their re-absorption) and thus to an increase in mass. The appearance of the cutoff in a linear form does not bother us here, since it is lost in a linear order when the same cutoff is used in other calculations.

The calculation of the Lamb shift is now easy (we subtract the result for the free electron and choose a corresponding momentum $\hbar k_a$, $k_a^2 = |(\vec{\nabla})_{aa}|^2 = \sum_v |(\vec{\nabla})_{va}|^2$; the superscript ‘r, obs’ stands for renormalized, observable):

$$\begin{aligned} \Delta E_a^{\text{r, obs}} &= \frac{4\alpha}{3} \lambda_c^2 \int \frac{d\hbar\omega}{2\pi} \left\{ \sum_\nu \frac{\hbar\omega |(\vec{\nabla})_{va}|^2}{E_a - E_v - \hbar\omega} - \frac{\hbar\omega k_a^2}{-\hbar\omega} \right\}, \\ &= \frac{4\alpha}{3} \lambda_c^2 \int \frac{d\hbar\omega}{2\pi} \left\{ \sum_\nu |(\vec{\nabla})_{va}|^2 \hbar\omega \left(\frac{1}{E_a - E_v - \hbar\omega} + \frac{1}{\hbar\omega} \right) \right\}, \\ &= \frac{4\alpha}{3} \lambda_c^2 \sum_\nu \int \frac{d\hbar\omega}{2\pi} \frac{(E_a - E_v) |(\vec{\nabla})_{va}|^2}{E_a - E_v - \hbar\omega}. \end{aligned} \quad (20.85)$$

The remaining integral is only log -divergent, with $x > 0$,

$$\begin{aligned} P \int_0^\Omega d\omega \frac{1}{x - \omega} &= \int_0^{x-\varepsilon} \frac{d\omega}{x - \omega} + \int_{x+\varepsilon}^\Omega \frac{d\omega}{x - \omega} \\ &= -\log(x - \omega) \Big|_0^{x-\varepsilon} - \log(x - \omega) \Big|_{x+\varepsilon}^\Omega = -\log \frac{\Omega}{x} \\ &\rightarrow \text{and with } x < 0, \quad = -\log \frac{\Omega}{|x|}. \end{aligned} \quad (20.86)$$

This results in¹⁷

$$\begin{aligned} \Delta E_a^{\text{r, obs}} &= -\frac{2\alpha}{3\pi} \lambda_c^2 \sum_v (E_a - E_v) |(\vec{\nabla})_{va}|^2 \log \frac{\hbar\Omega}{|E_a - E_v|} \\ &= \frac{8\alpha}{3} E_R \lambda_c^2 a_B |\Psi_a(0)|^2 \log \frac{\hbar\Omega}{\langle |E_a - E_v| \rangle}. \end{aligned} \quad (20.87)$$

¹⁷We replace $|E_a - E_v|$ under the log with the expectation value $\langle |E_a - E_v| \rangle$ (slow variation of the argument) and calculate the remaining sum $\sum_v (E_a - E_v) |(\vec{\nabla})_{va}|^2$ as

For the Lamb-shift ${}^2S_{1/2} - {}^2P_{1/2}$ in the $n = 2$ level, the resulting energy is (with $|\Psi_a(0)|^2 = 1/\pi a_B^3 (n=2)^3$ and $\lambda_C^2/a_B^2 \approx 5.328 \cdot 10^{-5}$)

$$\begin{aligned}
 E_{2S_{1/2}, 2P_{1/2}} &= \frac{\alpha}{3\pi} E_R \frac{\lambda_C^2}{a_B^2} \log \frac{\hbar\Omega}{\langle |E_a - E_v| \rangle} \\
 &\downarrow \text{ and with } \langle |E_a - E_v| \rangle \Big|_{a=2s} \cong 17E_R = 231 \text{ eV}, \\
 &\qquad \qquad \qquad \hbar\Omega = mc^2 \approx 511 \text{ keV}, \\
 &= \frac{\alpha}{3\pi} 13.6 \text{ eV} \underbrace{5.33 \cdot 10^{-5} \log \frac{5.11 \cdot 10^5}{231}}_{7.7} \\
 &= 4.32 \cdot 10^{-6} \text{ eV}.
 \end{aligned} \tag{20.88}$$

This corresponds to a frequency shift (with $E = 2\pi\hbar\nu$)

$$\nu \approx 1.05 \cdot 10^9 \text{ s}^{-1} \approx 1 \text{ GHz}. \tag{20.89}$$

Note: Virtual states far above the ionization energy are relevant in the determination of the average ¹⁸ $\langle |E_a - E_v| \rangle$, see Bethe & Salpeter, ‘QM of one and two electrons’, Academic Press, New York, 1957.

Today the Lamb-shift is used to determine the proton radius; experiments with $p + e^-$ and $p + \mu^-$ (muonic hydrogen) have shown different proton radii, see the ‘Proton Structure from the Measurement of 2S-2P Transition Frequencies of Muonic Hydrogen’, in Science **339**, 417 (2013). Meanwhile, this discrepancy (‘the proton radius puzzle’) has been (partly) resolved, see Nature **575**, 61 (2019).

follows (with $H_P \equiv H_{\text{particle}}$),

$$\begin{aligned}
 \sum_v (E_a - E_v) |\langle \vec{p} \rangle_{va}|^2 &= \sum_v \langle a | H_P p_i - p_i H_P | v \rangle \langle v | p_i | a \rangle \\
 &= - \sum_v \langle a | p_i | v \rangle \langle v | H_P p_i - p_i H_P | a \rangle \\
 &= \frac{1}{2} \langle a | [[H_P, p_i], p_i] | a \rangle.
 \end{aligned}$$

Calculating the commutator

$$[[H_P, p_i], p_i] = -\hbar^2 [[H_P, \partial_i], \partial_i] = -\hbar^2 [\partial_i V, \partial_i] = -\hbar^2 \nabla^2 V = \hbar^2 e^2 \nabla^2 (1/r) = -4\pi\hbar^2 e^2 \delta^3(\vec{r}).$$

The wave function Ψ_a thus is to be evaluated at $\vec{r} = 0$; only s states with $\Psi(0) \neq 0$ contribute to the Lamb-shift.

¹⁸This results in the relevant energy $\langle |E_a - E_v| \rangle \approx 17E_R$.

Chapter 21

Quantum Statistical Mechanics

Material for this chapter can be found in the book by Kerson Huang. We consider a system characterized by the Hilbert space \mathcal{H} and the Hamiltonian H . Let $\{\Phi_n\}$ be a basis in \mathcal{H} ,

$$\begin{aligned}\langle \Phi_n | \Phi_m \rangle &= \delta_{mn} \quad (\text{orthonormal}), \\ \sum_n |\Phi_n\rangle \langle \Phi_n| &= \mathbb{1} \quad (\text{complete}).\end{aligned}\tag{21.1}$$

Often, one chooses $\{\Phi_n\}$ to be the eigenvectors of H ,

$$H\Phi_n = E_n\Phi_n.\tag{21.2}$$

Any state Ψ of the system can be written as a superposition

$$\Psi = \sum_n \langle \Phi_n | \Psi \rangle \Phi_n = \sum_n c_n \Phi_n.\tag{21.3}$$

21.1 Expectation values

With $q = q_1, \dots, q_{3N}$ the position vector of a N -particle system, $\Psi(q) = \langle q | \Psi \rangle$ represents the complex wave function in the position representation. The expectation values of observables \mathcal{M} are given by

$$\langle \mathcal{M} \rangle = \frac{\langle \Psi | \mathcal{M} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\sum_{n,m} c_n^* c_m \langle \Phi_n | \mathcal{M} | \Phi_m \rangle}{\sum_n |c_n|^2}.\tag{21.4}$$

In statistical physics, we divide our system into a subsystem and a reservoir,

$$\Psi = \sum_n c_n \Phi_n \quad (21.5)$$

with c_n the amplitude of the reservoir, Φ_n the state of the subsystem (we choose a time-independent basis $\{\Phi_n\}$ of the subsystem).¹ Since we do not know the state of the reservoir, we replace the coefficients $c_n^* c_m$ with expectation values $\langle c_n | c_m \rangle_{\text{res}}$; likewise $|c_n|^2$ is replaced by $\langle c_n | c_n \rangle_{\text{res}}$.

In an experiment, the time average of \mathcal{M} is measured. With $\langle \Psi | \Psi \rangle = \sum_n \langle c_n | c_n \rangle_{\text{res}}$ time-independent in the complete system = system + reservoir, this average is given by

$$\overline{\langle \mathcal{M} \rangle}^T = \frac{\sum_{m,n} \overline{\langle c_n | c_m \rangle_{\text{res}}}^T \langle \Phi_n | \mathcal{M} | \Phi_m \rangle}{\langle \Psi | \Psi \rangle}. \quad (21.6)$$

In a first step, we consider the non-diagonal elements $\overline{\langle c_n | c_m \rangle_{\text{res}}}^T$, $n \neq m$. The dynamics of the reservoir adds *random phases* to the amplitudes c_n , such that the time averages of the matrix elements of different amplitudes disappear,

$$\overline{\langle c_n | c_m \rangle_{\text{res}}}^T = 0 \quad \text{for } n \neq m. \quad (21.7)$$

Eq. (21.7) is the *postulate of random phases*. Note that inelastic collisions between particles are sufficient to randomize the phases.² As a result of phase randomization, only the diagonal weights $\overline{\langle c_n | c_n \rangle_{\text{res}}}^T$ remain in the average (21.6). In a next step, we have to make a statistical assumption about the system in equilibrium, as is usually done with the assumption of molecular chaos in kinetic theory or the equal *a priori* probabilities in phase-space Γ for the microcanonical ensemble (at fixed extensive variables E , V , N) in classical statistical mechanics. The latter is needed to replace the time-average by the ensemble-average and brings us to the ergodic hypothesis:

¹Here, the reservoir serves as a generator of random phases and not for the exchange of energy or particles—the latter reservoirs are introduced in the context of the canonical and grand canonical ensembles.

²The assumption (21.7) breaks down when the reservoir does not properly randomize phases. This can happen in small (mesoscopic) systems when the phase is preserved over distances that are of similar length or longer than the sample. Also, phase information may be kept in a small reservoir as is often discussed in the context of quantum thermodynamics. The important element entering the discussion then is the phenomenon of entanglement between the system and the reservoir, preventing the system from being factorized.

The trajectories of particles explore equally often all parts of phase-space Γ that are compatible with fixed E , V , and N . Thus, in the present case, we also define the *microcanonical ensemble* by giving all states³ Φ_n that are compatible with the constraints of fixed energy E , volume V , and particle number N an *equal a priori probability*⁴

$$\overline{\langle c_n | c_n \rangle_{\text{Res}}}^T = \begin{cases} 1, & \text{if } E < E_n < E + \Delta, \\ 0, & \text{else.} \end{cases} \quad (21.8)$$

This gives the expression for the experimentally measured time average

$$\langle \mathcal{M} \rangle^T = \frac{\sum_n w_n \langle \Phi_n | \mathcal{M} | \Phi_n \rangle}{\sum_n w_n}, \quad (21.9)$$

$$w_n = \begin{cases} 1, & E < E_n < E + \Delta, \\ 0, & \text{else.} \end{cases} \quad (21.10)$$

21.1.1 Density matrix

In chapter 7 of the QM I lecture, we have learned about a formalism that is optimally adapted to the above problem: The quantum mechanics of a system whose state has to be described by an *incoherent* superposition of states is given by the *density matrix*

$$\rho = \sum_n w_n |\Phi_n\rangle \langle \Phi_n|. \quad (21.11)$$

The density matrix ρ then generalizes the notion of a coherent wave function Ψ or state $|\Psi\rangle$ to an incoherent superposition of states with weights w_n . These weights w_n are usually normalized according to $\sum_n w_n = 1$ ⁵. Here, we take the normalization into account explicitly and write for the expectation values

$$\overline{\langle \mathcal{M} \rangle}^T = \frac{\text{Tr}(\rho \mathcal{M})}{\text{Tr} \rho}, \quad \text{Tr}(A) = \text{Trace}(A) = \sum_n \langle \Phi_n | A | \Phi_n \rangle. \quad (21.12)$$

³Remember that a classical state is a vector $(p, q) \in \Gamma$ in phase space.

⁴In classical mechanics, we count the states in phase-space by using the integral $\int d^{3N}p d^{3N}q / h^{3N}$ where h is an action (plus Gibbs factor $1/N!$ for the correct ‘Boltzmann counting’ motivated by the mixing (entropy) of indistinguishable particles); in quantum statistical mechanics every state obtains a weight 1. If we take Planck’s constants to be the action $h = 2\pi\hbar$ and use the Heisenberg uncertainty principle, we obtain a weight of 1 for every independent state both classically and quantum mechanically. The indistinguishability of particles is inherent to quantum mechanics and the Gibbs factor appears straightforwardly.

⁵If $w_n = \delta_{n0}$, we have a pure state.

The equivalence of (21.9) and (21.12) is trivial.

The time evolution of ρ is given by the Neumann differential equation, the Liouville theorem in its quantum mechanical version,

$$i\hbar\partial_t\rho = [H, \rho]. \quad (21.13)$$

Using $\rho = \rho(H)$, see (21.9) and (21.10), we obtain

$$[H, \rho] = 0 \quad (21.14)$$

and the density matrix is time-independent. In the formalism of the density matrices, the (phase) reservoir causes the density matrix to collapse onto the diagonal and the ergodicity gives those diagonal elements which belong to states compatible with the given parameters E , V and N the same weight.

A useful property of the trace Tr is

$$\text{Tr}(AB) = \text{Tr}(BA). \quad (21.15)$$

This renders to trace Tr independent of the representation,

$$\text{Tr}(UAU^{-1}) = \text{Tr}(A). \quad (21.16)$$

Using the concept of density matrices ρ from the beginning, one may proceed in a more straightforward way. Consider a system that is isolated but still coupled to the rest of the world (a reservoir), with the total Hilbert space $\mathcal{H}_{\text{sys}} \otimes \mathcal{H}_{\text{res}}$. Note that this reservoir neither exchanges energy nor particles with our isolated system, hence E , V , N are fixed (alternatively, we look at a subsystem in a larger system). In quantum mechanics, the measurement of a system observable \mathcal{M} provides us with the expectation value $\langle \mathcal{M} \rangle = \text{Tr}(\rho\mathcal{M})$. Since we do not know anything about the rest of the world (or reservoir), we trace it out, $\langle \mathcal{M} \rangle = \text{Tr}_{\text{sys}}\text{Tr}_{\text{res}}(\rho\mathcal{M}) = \text{Tr}_{\text{sys}}[\text{Tr}_{\text{res}}(\rho)\mathcal{M}] \equiv \text{Tr}_{\text{sys}}(\rho_{\text{sys}}\mathcal{M})$, since \mathcal{M} only acts in \mathcal{H}_{sys} . In tracing ρ over the reservoir, we have eliminated the entanglement with the reservoir/the rest of the world, what corresponds to the phase randomization above. Requiring ρ_{sys} to describe a time-independent equilibrium state, we have $\partial_t\rho_{\text{sys}} = 0$ and hence $[H_{\text{sys}}, \rho_{\text{sys}}] = 0$. As a result, ρ_{sys} is diagonal in the eigenstates $|\Phi_n\rangle$ of \mathcal{H}_{sys} and hence $\rho_{\text{sys}} = \sum_i w_i |\Phi_n\rangle\langle\Phi_n|$ with real coefficients w_i (since $\rho = \rho^\dagger$). Finally, the *a priori* assumption of equal probabilities $w_i = 1/\Gamma(E)$ provides the final result for the properly normalized (i.e., $\text{Tr}(\rho_{\text{sys}}) = 1$, cf. Eq. (21.18)) density matrix of the microcanonical ensemble, $\rho_{\text{sys}} = \mathbb{1}/\Gamma(E)$.

21.2 Ensembles

We recapitulate the three ensembles of classical statistical physics in a formulation adapted to quantum statistics.⁶ As our basis $\{\Phi_n\}$, we choose the eigenstates of the Hamiltonian, $H\Phi_n = E_n\Phi_n$. The states Φ_n^N belong to the N -particle sector.

Micro canonical ensemble, E, V, N fixed:

$$\text{density matrix: } \rho = \sum_n w_n |\Phi_n\rangle\langle\Phi_n|, \quad (21.18)$$

$$w_n = \begin{cases} 1, & E < E_n < E + \Delta, \\ 0, & \text{else.} \end{cases}$$

$$\Gamma(E) = \text{Tr}(\rho) = \Delta \omega(E).$$

$$\text{density of states: } \omega(E) = \sum_n \delta(E - E_n).$$

$$\text{thdyn potential: } S(E, V, N) = k_B \log \Gamma(E).$$

Canonical ensemble, T, V, N fixed:

$$\text{density matrix: } \rho = e^{-\beta H}, \quad (21.19)$$

$$= \sum_n w_n |\Phi_n\rangle\langle\Phi_n|$$

$$w_n = e^{-\beta E_n},$$

$$\beta = \frac{1}{k_B T}.$$

$$\text{partition function: } Z_N(T, V) = \text{Tr}(e^{-\beta H}),$$

$$\text{thdyn potential: } F(T, V, N) = -k_B T \ln Z_N.$$

Grand canonical ensemble, T, V, μ fixed:

$$\text{density matrix: } \rho = e^{-pV/k_B T} e^{-\beta(H - \mu N)}, \quad (21.20)$$

$$\text{partition function: } \mathcal{Z} = \text{Tr}(e^{-\beta(H - \mu N)})$$

$$= \sum_{N=0}^{\infty} z^N Z_N \quad (z = e^{\beta\mu} \text{ the fugacity})$$

⁶Note the normalization

$$\text{Tr}(\rho) = 1 \quad (21.17)$$

(trace in the Fock space) for the density function ρ of the grand canonical ensemble. With $\text{Tr}(\rho) = \exp(-pV/k_B T) \mathcal{Z}$, we obtain an expression of the grand potential, $\Omega = -pV$.

$$= \sum_{N,n} \langle \Phi_n^N | e^{-\beta(H-\mu N)} | \Phi_n^N \rangle$$

where the trace is over Fock space.

$$\text{thdyn potential: } \Omega(T, V, z) = -pV = -k_B T \log \mathcal{Z}(T, V, z)$$

or

$$F(T, V, N) = N \log z - k_B T \log \mathcal{Z},$$

$$N = z \partial_z \log \mathcal{Z}.$$

In modern solid-state physics and in quantum engineered systems, one often reaches areas where the traditional quantum statistical mechanics collapses: For example, currents in small metallic/semiconducting rings (continuous currents) are described by the canonical ensemble (N is fixed). As another example, we mention coherence effects in small (mesoscopic) structures; here there is no reservoir that creates random phases and thus (21.7) is not guaranteed.

21.3 Concerning the third law of thermodynamics

With $S = k_B \log \Gamma$, the entropy counts the number of available states. According to the Boltzmann factor $\rho \sim e^{-H/T}$, only the ground state of the system is available at $T = 0$ and therefore $\Gamma \rightarrow 1$ and $S \rightarrow 0$. If the ground state is $g \sim N$ -fold degenerate, $s = S/N \sim (\log N)/N \rightarrow 0$ and the entropy per particle vanishes. A more relevant problem is the continuity of the spectrum—the above argument is correct, but rather academic, since the ground state is only realized for $T < \Delta E \sim \hbar^2/mL^2 \sim \mu\text{K}$ (where we took $m = m_{\text{nucleon}}$ and $L = 1 \mu\text{m}$ for a trapped atom). The density of states $\omega(E)$ in the low-temperature phase of the substance is therefore relevant. At $T \rightarrow 0$ all substances transform into one of the following phases: crystals (almost all), glasses, superfluids (^3He , ^4He at normal pressure). The density of states in these systems disappears according to $\omega(E) \sim E^{d-1}$ for $E \rightarrow 0$ (with the exception of glasses) and accordingly, $S \rightarrow 0$ for $T \rightarrow 0$.

Chapter 22

Gases

Material for this chapter can be found in the book by Kerson Huang.

The topic of quantum gases has attracted huge interest during recent years. The ability to laser- and evaporatively cool atoms to amazingly low temperatures (μK and nK) using optical systems and the development of suitable (magneto-optic) traps allowed Eric Cornell and Carl Wiemann at JILA and Wolfgang Ketterle, Kendall David, and Marc-Oliver Mewes at MIT to produce a Bose-Einstein condensate in 1995; in 2001 Cornell, Wiemann, and Ketterle received the physics Nobel prize for this achievement. At the beginning of the millenium, the cooling of fermionic atoms into a quantum-degenerate state with $k_{\text{B}}T \ll \varepsilon_{\text{F}}$ followed. Today quantum optics is an ever blooming field. The focus is on interesting systems described by complicated Hamiltonians that generate non-trivial physics. The quantum gases are then used as ‘quantum simulators’ for such Hamiltonians. Elements of such Hamiltonians include non-trivial periodic lattices (\rightarrow simulations of graphene, topological insulators assisted by optical lattices), interactions (tunable by Feshbach resonances), interactions mediated by gauge fields (generated by structured laser fields), disorder, Bose-Fermi mixtures, multicomponent gases, etc.

Note the difference between quantum gases and strongly correlated quantum gases, i.e., gases with strong interactions, and quantum liquids that commonly interact strongly *and* have a high density. Typical examples of quantum liquids are fermionic ^3He and bosonic ^4He . The subject of fermionic quantum liquids brings one to the Landau theory of Fermi-liquids (encompassing the electron gas/liquid in a metal), that is renormalized due to the strong interactions but otherwise behaves phenomenologically like a

fermionic quantum gas. In what follows, we first discuss the statistics of classical and quantum gases before we study Fermi and Bose gases more thoroughly in the following two chapters, respectively.

22.0.1 The classical ideal gas

For later comparison, we first start with the classical or so called “Boltzmann” gas. We calculate the partition function in the canonical ensemble¹

$$Z_N = \sum_{\{n_{\vec{p}_i}\}_N} e^{-\beta \sum_{\vec{p}_i} n_{\vec{p}_i} \varepsilon_{p_i}} \times \underbrace{g_c}_{\text{configs}}, \quad (22.1)$$

$$n_{\vec{p}_i} = \# \text{ particles with momentum } \vec{p}_i, \quad \sum_{\vec{p}_i} n_{\vec{p}_i} = N,$$

$$\varepsilon_{p_i} = p_i^2/2m.$$

The correct Boltzmann counting provides the number of configurations, the possible ways to distribute the N particles among the momenta $\vec{p}_0, \dots, \vec{p}_i, \dots$,²

$$g_c = \frac{N!}{n_{\vec{p}_0}! n_{\vec{p}_1}! \dots n_{\vec{p}_i}! \dots}, \quad (22.2)$$

In addition, we have to take into account a Gibbs factor $1/N!$ in order to avoid the Gibbs’ paradox, hence $g_c \rightarrow g = g_c/N!$. We exchange the sum and product by using the multi-nomial theorem,³

$$Z_N = \sum_{\{n_{\vec{p}_i}\}_N} \prod_i \frac{e^{-\beta n_{\vec{p}_i} \varepsilon_{p_i}}}{n_{\vec{p}_i}!} = \frac{1}{N!} \left(\sum_i e^{-\beta \varepsilon_{p_i}} \right)^N, \quad (22.5)$$

¹Here, we discretize phase-space and work with discrete momenta \vec{p}_i and sums; you may want to redo the calculation with the standard classical partition integral $\int d^N p d^N q$ for Z_N .

²In the classical Boltzmann gas the particles are distinguishable resulting in configuration number $g_c > 1$; in quantum mechanics the particles are generically indistinguishable giving $g = 1$.

³The multinomial theorem says that

$$(x_1 + x_2 + \dots + x_n)^k = \sum_{k_1 + \dots + k_n = k} \binom{k}{k_1, \dots, k_n} x_1^{k_1} x_2^{k_2} \dots x_n^{k_n}, \quad (22.3)$$

where

$$\binom{k}{k_1, \dots, k_n} = \frac{k!}{k_1! \dots k_n!}. \quad (22.4)$$

and then replace the sum over momenta by integrals, counting one state per phase-space volume $h^3 \rightarrow (2\pi\hbar)^3$ (alternatively replace $d^3n \rightarrow (L/2\pi)^3 d^3k \rightarrow V/(2\pi\hbar)^3 d^3p$)

$$\sum_{\vec{p}_i} e^{-\beta\varepsilon_{p_i}} = V \int \frac{d^3p}{(2\pi\hbar)^3} \exp\left(-\frac{p^2}{2mk_{\text{B}}T}\right) = \frac{V}{\lambda^3}, \quad (22.6)$$

where

$$\lambda = \sqrt{\frac{2\pi\hbar^2}{mk_{\text{B}}T}} \quad (22.7)$$

is the thermal de Broglie wavelength. The partition function, the thermodynamic potential, the caloric equation of state, and the thermodynamic equation of state then become (we use Stirling's formula for $N!$)

$$\begin{aligned} Z_N &= \frac{1}{N!} \left(\frac{V}{\lambda^3}\right)^N, \\ \log Z_N &= N \log \frac{V}{N\lambda^3} + N = N \left[-\log(n\lambda^3) + 1\right], \\ F &= -k_{\text{B}}T \log Z_N = Nk_{\text{B}}T \left[\log(n\lambda^3) - 1\right], \\ S &= -\left.\frac{\partial F}{\partial T}\right|_V = -F/T + (3/2)Nk_{\text{B}} = (-F + U)/T, \\ p &= -\left.\frac{\partial F}{\partial V}\right|_T = Nk_{\text{B}}T/V, \end{aligned} \quad (22.8)$$

with the gas parameter $\delta \equiv n\lambda^3$. For an ideal classical gas $\delta \ll 1$, i.e., the gas is dilute (small density n) or hot (large temperature T , small λ). Using $U = (3/2)Nk_{\text{B}}T$, the above results produce the Sackur-Tedrode formula for the entropy with E replaced by T and h by $2\pi\hbar$, $S(T, V, N) = Nk_{\text{B}}[-\log(n\lambda^3) + 5/2]$.

22.0.2 Ideal quantum gases

The quantum gases are easiest to describe in the grand canonical ensemble (this removes the constraint on the particle number N , i.e., all values of N are admitted)

$$\mathcal{Z}(V, T, z) = \sum_{N=0}^{\infty} z^N Z_N(V, T)$$

$$\begin{aligned}
&= \sum_{N=0}^{\infty} \sum_{\{n_{\vec{p}}\}_N} z^N \exp\left(-\beta \sum_{\vec{p}} n_{\vec{p}} \varepsilon_p\right) \\
&= \sum_{n_{\vec{p}_0}} \sum_{n_{\vec{p}_1}} \dots \sum_{n_{\vec{p}_k}} \dots \left[(ze^{-\beta \varepsilon_{p_0}})^{n_{\vec{p}_0}} \dots (ze^{-\beta \varepsilon_{p_k}})^{n_{\vec{p}_k}} \dots \right] \\
&= \prod_{\vec{p}} \left[\sum_n (ze^{-\beta \varepsilon_p})^n \right]. \tag{22.9}
\end{aligned}$$

For bosons, the state \vec{p} can be populated with $n = 0, 1, \dots, k, \dots, \infty$ particles. This has to be contrasted with the fermions, where only occupations $n = 0, 1$ are allowed (Pauli exclusion principle, every state is occupied at most once). Thus (22.9) simplifies to the product form

$$\mathcal{Z}(V, T, z) = \begin{cases} \prod_{\vec{p}} 1/(1 - ze^{-\beta \varepsilon_p}), & \text{bosons,} \\ \prod_{\vec{p}} (1 + ze^{-\beta \varepsilon_p}), & \text{fermions.} \end{cases} \tag{22.10}$$

The grand potential $\Omega = -pV$ and the particle number N are given by

$$\begin{aligned}
pV &= k_B T \log \mathcal{Z} \\
&= k_B T \begin{cases} -\sum_{\vec{p}} \log(1 - ze^{-\beta \varepsilon_p}), & \text{bosons,} \\ \sum_{\vec{p}} \log(1 + ze^{-\beta \varepsilon_p}), & \text{fermions} \end{cases} \tag{22.11}
\end{aligned}$$

$$\begin{aligned}
N &= z \frac{\partial}{\partial z} \log \mathcal{Z} \\
&= \begin{cases} \sum_{\vec{p}} 1/(e^{\beta \varepsilon_p}/z - 1), & \text{bosons,} \\ \sum_{\vec{p}} 1/(e^{\beta \varepsilon_p}/z + 1), & \text{fermions.} \end{cases} \tag{22.12}
\end{aligned}$$

From the latter, we obtain the occupation numbers $\langle n_{\vec{p}} \rangle$ for the state $|\vec{p}\rangle$

$$\langle n_{\vec{p}} \rangle = \begin{cases} 1/(e^{\beta(\varepsilon_p - \mu)} - 1), & \text{bosons,} \\ 1/(e^{\beta(\varepsilon_p - \mu)} + 1), & \text{fermions,} \end{cases} \tag{22.13}$$

[by inspection of (22.12) and using $N = \sum_{\vec{p}} \langle n_{\vec{p}} \rangle$, or calculate

$$\langle n_{\vec{p}} \rangle = \frac{1}{\mathcal{Z}} \sum_N z^N \sum_{\{n_{\vec{p}}\}_N} n_{\vec{p}} e^{-\beta \sum_{\vec{p}} n_{\vec{p}} \varepsilon_p} = -\frac{1}{\beta} \frac{\partial}{\partial \varepsilon_p} \log \mathcal{Z}$$

and use (22.11).]

The sums in (22.11) and (22.12) can be simplified further. Using

$$\sum_{\vec{p}} = V \int_0^\infty \frac{4\pi p^2 dp}{(2\pi\hbar)^3} \stackrel{x^2 = \beta p^2/2m}{=} \frac{V}{\lambda^3} \frac{4}{\sqrt{\pi}} \int_0^\infty dx x^2, \quad (22.14)$$

we obtain

$$\begin{cases} p = (k_B T / \lambda^3) f_{5/2}(z), \\ N/V = n = (1/\lambda^3) f_{3/2}(z), \end{cases} \quad \text{fermions} \quad (22.15)$$

where the functions $f_{5/2}$ and $f_{3/2}$ derive from [we use the expansion $\log(1+s) = s - s^2/2 + s^3/3 - s^4/4 + \dots = \sum_\ell (-1)^{\ell+1} s^\ell / \ell$ and $s = ze^{-x^2}$]

$$\begin{aligned} f_{5/2}(z) &= \frac{4}{\sqrt{\pi}} \int_0^\infty dx x^2 \log(1 + ze^{-x^2}) \\ &= \frac{4}{\sqrt{\pi}} \int_0^\infty dx x^2 \sum_{\ell=1}^\infty \frac{(-1)^{\ell+1} z^\ell}{\ell} e^{-\ell x^2} = \sum_{\ell=1}^\infty \frac{(-1)^{\ell+1} z^\ell}{\ell^{5/2}}, \\ f_{3/2}(z) &= z \frac{\partial}{\partial z} f_{5/2}(z) = \sum_{\ell=1}^\infty \frac{(-1)^{\ell+1} z^\ell}{\ell^{3/2}}, \end{aligned} \quad (22.16)$$

replacing (22.11) and (22.12) by simpler expressions. In the second integral, we have replaced $\ell x^2 \rightarrow r^2$ (\rightarrow additional factor $\ell^{-3/2}$) and the remaining integral provides the compensating factor $\sqrt{\pi}/4$.

The analysis for bosons is slightly more complicated as the terms $\vec{p} = 0$ in (22.11) and (22.12) lead to divergences (Bose-Einstein condensation), see below. Thus, we first separate these terms from the rest before replacing the sum by an integral and obtain the expressions

$$\begin{cases} p = (k_B T / \lambda^3) g_{5/2}(z) - (k_B T / V) \log(1 - z), \\ N/V = n = (1/\lambda^3) g_{3/2}(z) + (1/V) z / (1 - z), \end{cases} \quad \text{bosons} \quad (22.17)$$

with

$$\begin{aligned} g_{5/2}(z) &= -\frac{4}{\sqrt{\pi}} \int_0^\infty dx x^2 \log(1 - ze^{-x^2}) = \sum_{\ell=1}^\infty \frac{z^\ell}{\ell^{5/2}}, \\ g_{3/2}(z) &= z \frac{\partial}{\partial z} g_{5/2}(z) = \sum_{\ell=1}^\infty \frac{z^\ell}{\ell^{3/2}}. \end{aligned} \quad (22.18)$$

For $z \rightarrow 1$, the term $(1/V) z / (1 - z) = \langle n_0 \rangle / V$ will single out a fraction of the total density as a k-space condensate, i.e., $\langle n_0 \rangle \propto N$; we will discuss in detail

later how the occupation number $\langle n_0 \rangle$ can assume an extensive/macroscopic value. The consequence is the appearance of quantum mechanical effects on a macroscopic scale: MQP = macroscopic quantum phenomena. The internal energy U of the quantum gas is obtained from (22.11),

$$\begin{aligned} U(V, T, z) &= \frac{1}{\mathcal{Z}} \sum_N z^N \sum_{\{n_{\vec{p}}\}_N} \left(\sum_{\vec{p}} n_{\vec{p}} \varepsilon_p \right) e^{-\beta \sum_{\vec{p}} n_{\vec{p}} \varepsilon_p} \\ &= -\frac{\partial}{\partial \beta} \log \mathcal{Z}(V, T, z) = -\frac{\partial}{\partial \beta} \left[\frac{V}{\lambda^3} \begin{cases} f_{5/2}(z) \\ g_{5/2}(z) \end{cases} \right] \\ &= \frac{3}{2} \frac{V k_B T}{\lambda^3} \begin{cases} f_{5/2}(z), & \text{fermions,} \\ g_{5/2}(z), & \text{bosons,} \end{cases} \end{aligned} \quad (22.19)$$

$$= \frac{3}{2} p(V, T, z) V. \quad (22.20)$$

In the last expression, valid for the pure Bose gas, we have ignored the condensation terms for the bosons. Free Fermi and Bose gases have a multitude of interesting properties and applications which we will come back to later.

22.0.3 Classical limit of ideal quantum gases

We calculate the *canonical partition function* of ideal quantum gases and investigate their classical (high-temperature) limit $\beta \rightarrow 0$, $T \rightarrow \infty$. We have to calculate the quantity

$$Z_N = \text{Tr}(e^{-\beta H_0}) = \sum_n \langle \Phi_n | e^{-\beta E_n} | \Phi_n \rangle. \quad (22.21)$$

The energy eigenvalue problem is solved by (note that n describes a $3N$ multi-index $n \rightarrow \{\vec{p}_1, \dots, \vec{p}_N\}$)

$$\begin{aligned} H_0 | \Phi_n \rangle &= E_n | \Phi_n \rangle, \\ E_n &= \sum_{i=1}^N \frac{p_i^2}{2m}, \quad \vec{p}_i = \frac{2\pi}{L} \hbar \vec{n}_i, \quad \vec{n}_i \in \mathbb{Z}^3, \\ \langle \vec{q}_1, \vec{q}_2, \dots, \vec{q}_N | \Phi_n \rangle &= \frac{1}{\sqrt{N!}} \sum_{\pi \in S_N} (\pm 1)^\pi \phi_{\vec{p}_1}(\vec{q}_{\pi_1}) \dots \phi_{\vec{p}_N}(\vec{q}_{\pi_N}) \\ \phi_{\vec{p}}(\vec{q}) &= \frac{e^{i\vec{p} \cdot \vec{q} / \hbar}}{\sqrt{V}}. \end{aligned}$$

Substituting the sum over states $| \Phi_n \rangle$ in (22.21) by an integral

$$\sum_n \rightarrow \frac{V^N}{N! (2\pi \hbar)^{3N}} \int d^{3N} p \quad (22.22)$$

and inserting the identity $\mathbb{1} = \int d^{3N}q |\vec{q}_1, \dots\rangle \langle \vec{q}_1 \dots|$, we obtain

$$Z_N = \frac{V^N}{N!(2\pi\hbar)^{3N}} \int d^{3N}p d^{3N}q |\Phi_n(\vec{q}_1, \dots, \vec{q}_N)|^2 e^{-\beta E_n}. \quad (22.23)$$

The factor $1/N!$ takes into account that permutations of the vector \vec{p}_i in $|\Phi_n\rangle$ does not generate any new states. We simplify the expression for the probability $|\Phi_n|^2$ according to

$$\begin{aligned} |\Phi_n|^2 &= \frac{1}{N!} \sum_{\pi, \pi'} (\pm 1)^{\pi + \pi'} \cdot \phi_{\vec{p}_1}^*(\vec{q}_{\pi_1}) \phi_{\vec{p}_1}(\vec{q}_{\pi'_1}) \dots \phi_{\vec{p}_N}^*(\vec{q}_{\pi_N}) \phi_{\vec{p}_N}(\vec{q}_{\pi'_N}) \\ &= \frac{1}{V^N} \sum_{\pi} (\pm 1)^{\pi} e^{i\vec{p}_1 \cdot (\vec{q}_1 - \vec{q}_{\pi_1})/\hbar} \dots, \end{aligned}$$

where we have used that every permutation π' under the integral $\int d^{3N}p$ gives the same contribution⁴; this allows us to replace $\pi' = \mathbb{1}$ and multiply by $N!$. The \vec{p} -integrations can be done by completing the square

$$\begin{aligned} \int \frac{d^3p}{(2\pi\hbar)^3} e^{-\beta p^2/2m + i\vec{p} \cdot \vec{q}/\hbar} &= \int \frac{d^3p}{(2\pi\hbar)^3} \exp\left[-\frac{\beta}{2m} \left(\vec{p} - \frac{im}{\hbar\beta} \vec{q}\right)^2\right] \exp\left[-\frac{mq^2}{2\hbar^2\beta}\right] \\ &= \frac{1}{\lambda^3} \exp\left[-\frac{\pi q^2}{\lambda^2}\right], \quad \lambda^2 = 2\pi\hbar^2/mk_B T. \end{aligned}$$

Making use of this result in the expression (22.22) for Z_N gives us the *exact* result

$$Z_N = \frac{1}{N!} \left(\frac{V}{\lambda^3}\right)^N \int \frac{d^{3N}q}{V^N} \sum_{\pi \in S^N} (\pm 1)^{\pi} \left[e^{-\frac{\pi}{\lambda^2}(\vec{q}_1 - \vec{q}_{\pi_1})^2} \dots e^{-\frac{\pi}{\lambda^2}(\vec{q}_N - \vec{q}_{\pi_N})^2} \right]. \quad (22.24)$$

For $T \rightarrow \infty$, $\lambda^2 \rightarrow 0$, the only term in the sum \sum_{π} that survives is the identity $\mathbb{1}$ and thus we obtain the classical result (22.8). For $T < \infty$ the terms in \sum_{π} become smaller the more particles are permuted by π . As long as $\langle (\vec{q}_i - \vec{q}_j)^2 \rangle > \lambda^2$ (i.e., at low densities), we can expand the sum in the number of permutations and obtain

$$\sum_{\pi} (\pm 1)^{\pi} [\dots] \approx 1 \pm \sum_{i < j} \exp\left[-\frac{2\pi}{\lambda^2}(\vec{q}_i - \vec{q}_j)^2\right] + 3 \text{ particle - terms.} \quad (22.25)$$

⁴Show this for $N = 2, 3$.

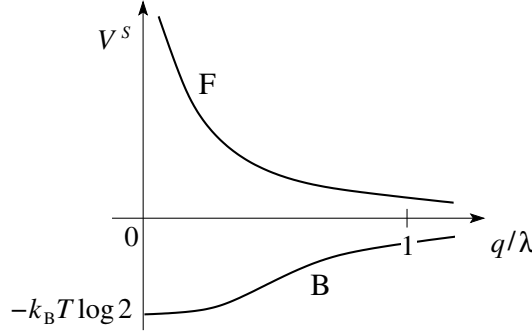


Fig. 22.1: Statistical interaction for bosons (B, attractive) and fermions (F, repulsive with a hard-core).

Using the identity $x = e^{\log x}$, we construct the potential

$$V^s(\vec{q}) \equiv -k_B T \log \left[1 \pm e^{-2\pi(q/\lambda)^2} \right], \quad (22.26)$$

allowing us to approximate (22.24) at large temperatures by the expression

$$Z_N \approx \frac{1}{N!} \left(\frac{V}{\lambda^3} \right)^N \int \frac{d^{3N}q}{V^N} \exp \left[-\beta \sum_{i < j} V^s(\vec{q}_i - \vec{q}_j) \right]. \quad (22.27)$$

Comparing (22.27) with the classical result (22.8), we see that the quantum statistics of the particles expresses itself as an effective potential under a configurational integral, the so called **statistical interaction**

$$V^s(q) = -k_B T \log(1 \pm e^{-2\pi(q/\lambda)^2}). \quad (22.28)$$

This interaction is short ranged (λ) and depends on the particle statistics, attractive for bosons and repulsive for fermions, see Fig. 22.1.

22.0.4 Interacting particles at $T \rightarrow 0$

For $T \rightarrow 0$, particles have small kinetic energies and hence $\vec{k} \rightarrow 0$; in this situation, particle-particle interaction can be described by *s*-wave scattering. The scattering phases $\delta_{\ell=0}(k \rightarrow 0)$ are given by (see the first lecture on quantum mechanics, scattering theory, partial wave analysis)

$$k \cot \delta_0 \approx -\frac{1}{a} + \frac{r_0}{2} k^2 + \dots \quad (22.29)$$

The *scattering length* is $a > 0$ for a *repulsive interaction* and $a < 0$ for an *attractive interaction*. The parameter r_0 gives the effective range of the

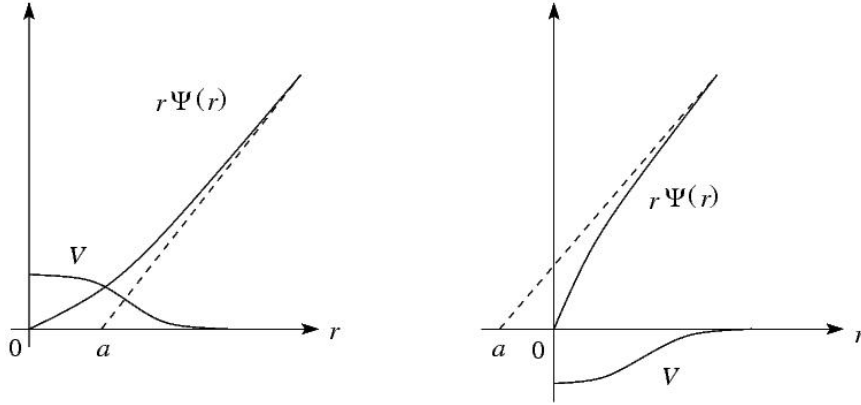


Fig. 22.2: Shape of the wave function $\Psi(r)$ for a repulsive (left) and an attractive (right) interaction $V(r)$.

potential. At small energies, $\cot \delta_0 \approx -1/ak$ is a valid approximation. For a repulsive interaction, the scattering between particles then corresponds to that of hard spheres of radius $a > 0$. The s -wave scattering of hard spheres can conveniently be described by the *pseudopotential*

$$V_{\ell=0}^a(\vec{r}) = \frac{4\pi a \hbar^2}{m} \delta(\vec{r}) \left(\frac{\partial}{\partial r} r \right) \quad (22.30)$$

i.e., the two Hamiltonians

$$\begin{aligned} H &= \frac{p^2}{2m} + V_{\text{hard sphere}, R=a}, \\ H &= \frac{p^2}{2m} + V_{\ell=0}^a(\vec{r}), \end{aligned} \quad (22.31)$$

generate the same scattering solutions $\Psi_{\ell=0,k}(\vec{r})$ in the exterior space $r > a$ for $\ell = 0$, $k \rightarrow 0$; correspondingly they generate the same eigenvalues $E_{\ell=0}(k)$.

A further simplification arises when the pseudopotential $V_{\ell=0}^a(\vec{r})$ is treated perturbatively. Within a perturbative analysis, one uses plane waves as the unperturbed solutions. As these solutions, say $\varphi(\vec{r})$, are regular at the origin $\vec{r} = 0$, we have $\delta(\vec{r})(\partial_r r \varphi) = \delta(\vec{r})\varphi(0)$ and one may replace⁵ (22.30)

⁵One may go beyond perturbation theory if one replaces the operator $g\delta(\vec{r})\partial_r r$ by $\bar{g}(\Lambda)\delta(\vec{r})$ with a renormalized coupling $\bar{g}(\Lambda)$ and Λ a short distance cutoff, see W. Zwerger's Varenna lectures for details and conditions.

by $V_{\ell=0}^a(\vec{r}) = (4\pi a\hbar^2/m)\delta(\vec{r})$. To treat repulsively interacting quantum gases (either fermions or bosons) in the limit of small temperatures $T \rightarrow 0$ the following strategy is suitable:

1. Calculate the spectrum of

$$H = \sum_i \frac{p_i^2}{2m} + \frac{4\pi a\hbar^2}{m} \sum_{i<j} \delta(\vec{q}_i - \vec{q}_j) \quad (22.32)$$

in *perturbation theory*. The result depends on the particle statistics. Note that divergencies at small scales $r \rightarrow 0$ or large momenta $k \rightarrow \infty$ have to be treated by reconsidering the full operator $\delta(\vec{r})(\partial_r r)$, e.g., in the context of the Bogoliubov spectrum for repulsive condensed bosons.

2. Calculate the canonical partition function from the spectrum.

The above strategy is succesfull for repulsive particles with

$$\begin{aligned} \lambda, n^{-1/3} = \langle \text{distance} \rangle &> a, \\ a/\lambda(T), an^{1/3}(\text{dilute gas}) &\ll 1. \end{aligned} \quad (22.33)$$

Chapter 23

Fermions

Material on this chapter can be found in the book of Kerson Huang and in many books on solid state physics, see chapters on the electron gas, Fermi sea, metals. We first summarize the results for the Fermi gas from the previous chapters:

Partition function:

$$\mathcal{Z}(V, T, z) = \prod_{\vec{p}} (1 + ze^{-\beta\epsilon_p}). \quad (23.1)$$

Equation of state:

$$p = \frac{k_B T}{V} \log \mathcal{Z} = \frac{k_B T}{\lambda^3} f_{5/2}(z). \quad (23.2)$$

Particle density:

$$n = \frac{z}{V} \partial_z \log \mathcal{Z} = \frac{1}{\lambda^3} f_{3/2}(z). \quad (23.3)$$

Occupation number:

$$\langle n_{\vec{p}} \rangle = \frac{1}{e^{\beta(\epsilon_p - \mu)} + 1}. \quad (23.4)$$

$$\begin{aligned} f_{5/2}(z) &= \frac{4}{\sqrt{\pi}} \int_0^\infty dx x^2 \log(1 + ze^{-x^2}) = \sum_{\ell=1}^\infty \frac{(-1)^{\ell+1} z^\ell}{\ell^{5/2}}, \\ f_{3/2}(z) &= \frac{4}{\sqrt{\pi}} \int_0^\infty dx \frac{x^2}{1 + e^{x^2}/z} = \sum_{\ell=1}^\infty \frac{(-1)^{\ell+1} z^\ell}{\ell^{3/2}}. \end{aligned}$$

We carry out the following program: We first solve the equation relating the gas parameter $n\lambda^3$ and the fugacity $z = e^{\beta\mu}$,

$$\delta = n\lambda^3 = f_{3/2}(z), \quad (23.5)$$

in order to find the chemical potential μ as a function of the density n

$$\mu(n) = k_B T \log z(n). \quad (23.6)$$

By calculating the internal energy (or caloric equation of state)

$$U = \sum_{\vec{p}} \langle n_{\vec{p}} \rangle \varepsilon_p, \quad (23.7)$$

we can then find the thermal equation of state via (see (22.19))

$$p = \frac{2}{3} \frac{U}{V} \quad (23.8)$$

and we are done.

The dimensionless **density parameter**

$$\delta = n\lambda^3 \quad (23.9)$$

characterizes the quantumness of the gas, see Fig. 23.1,

$$\begin{aligned} \delta \ll 1 & \rightarrow \text{classical gas,} \\ \delta \gg 1 & \rightarrow \text{quantum gas.} \end{aligned} \quad (23.10)$$

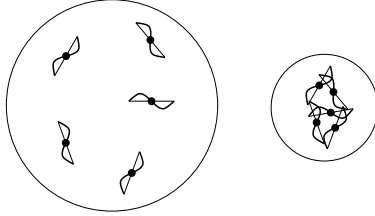


Fig. 23.1: Density parameter: left, a dilute classical gas with $\delta \ll 1$, right, a dense quantum gas with $\delta \gg 1$.

We expand $f_{3/2}(z)$ for large and small z using the series expansion for $z \ll 1$,

$$f_{3/2}(z) \approx z - \frac{z^2}{2^{3/2}} + \frac{z^3}{3^{3/2}} - \dots, \quad (23.11)$$

and the integral representation for $z \gg 1$,

$$\begin{aligned} f_{3/2}(z) &= \frac{4}{\sqrt{\pi}} \int_0^\infty dx \frac{x^2}{1 + e^{x^2 - \beta\mu}} \\ &= \frac{2}{\sqrt{\pi}} \int_0^\infty dy \frac{\sqrt{y}}{1 + e^{y - \beta\mu}} = \frac{4}{3\sqrt{\pi}} \int_0^\infty dy \frac{y^{3/2} e^{y - \beta\mu}}{(1 + e^{y - \beta\mu})^2} \end{aligned} \quad (23.12)$$

where $z = e^{\beta\mu}$. Above, we have replaced $y = x^2$ and performed a partial integration. The function

$$\frac{e^{y-\nu}}{(1+e^{y-\nu})^2} = -\frac{d}{dy} \frac{1}{1+e^{y-\nu}} \quad (23.13)$$

evaluated at fixed $\nu = \log z = \beta\mu$ is sharply peaked around ν motivating an expansion of $y^{3/2}$ around this point (Sommerfeld expansion),

$$y^{3/2} \approx \nu^{3/2} + \frac{3}{2}\nu^{1/2}(y-\nu) + \frac{3}{8\nu^{1/2}}(y-\nu)^2 + \dots \quad (23.14)$$

We then have to calculate the integral

$$I_n = \int_{-\infty}^{\infty} dt \frac{t^n e^t}{(1+e^t)^2}, \quad (23.15)$$

where we have replaced the lower bound $t = y - \nu \rightarrow -\nu$ at $y = 0$ by $-\infty$ at large values of $z \gg 1$ or ν . One then finds the coefficients

$$I_0 = 1, I_1 = I_{2n+1} = 0, I_2 = \frac{\pi^2}{3}, I_4 = \frac{7\pi^4}{15}, I_6 = \frac{31\pi^6}{21}, \quad (23.16)$$

providing us with the **Sommerfeld expansion** for $f_{3/2}(z)$,

$$f_{3/2}(z) \approx \frac{4}{3\sqrt{\pi}} \left[(\log z)^{3/2} + \frac{\pi^2}{8} (\log z)^{-1/2} + \dots \right] + \mathcal{O}(e^{-\nu}). \quad (23.17)$$

23.1 Dilute Fermi gas, $\delta = n\lambda^3 \ll 1$

Starting from $\delta = f_{3/2}(z)$, (23.5), and using (23.11), we find ($f_{\text{MB}}(\vec{p})$ is the classical Maxwell-Boltzmann distribution¹)

$$\begin{aligned} \delta &\approx z - \frac{z^2}{2\sqrt{2}} \rightarrow z \approx \delta + \frac{1}{2\sqrt{2}}\delta^2 > 0, \\ \langle n_{\vec{p}} \rangle &\approx ze^{-\beta\epsilon_p} = n \left(\frac{2\pi\hbar^2}{mk_{\text{B}}T} \right)^{3/2} \exp\left[-\frac{p^2}{2mk_{\text{B}}T}\right] \\ &\approx h^3 f_{\text{MB}}(\vec{p}). \end{aligned} \quad (23.18)$$

¹The classical distribution function $f_{\text{MB}}(\vec{p}) = n(2\pi mk_{\text{B}}T)^{-3/2} \exp(-p^2/2mk_{\text{B}}T)$, differs from the present result by its normalization, $\langle n_{\vec{p}} \rangle = (2\pi\hbar)^3 f_{\text{MB}}(\vec{p})$.

The series expansion of $f_{5/2}(z)$ provides us with the thermal equation of state

$$pV \approx \frac{V}{\lambda^3} k_B T \left(z - \frac{z^2}{2^{5/2}} + \dots \right) \quad (23.19)$$

$$\approx N k_B T \left(1 + \frac{n \lambda^3}{4\sqrt{2}} + \dots \right), \quad (23.20)$$

describing the dilute Fermi Gas as a *classical ideal gas* with quantum corrections $\propto \delta$. Note that the chemical potential μ and its dependence on n appears in the prefactor of the Maxwell-Boltzmann distribution (via the fugacity $z = \exp(\beta\mu)$).

23.2 Degenerate Fermi gas, $\delta = n\lambda^3 \gg 1$

Using (23.5) and (23.17), we find the relation

$$\delta = n\lambda^3 \approx \frac{4}{3\sqrt{\pi}} \left[(\beta\mu)^{3/2} + \frac{\pi^2}{8} \frac{1}{(\beta\mu)^{1/2}} \right] \quad (23.21)$$

between the density n and the chemical potential² μ , and thus to lowest order

$$\mu(T=0) = \frac{\hbar^2}{2m} (6\pi^2 n)^{2/3} = \varepsilon_F. \quad (23.22)$$

Including corrections to lowest order in T , one finds the chemical potential $\mu(T)$ to decrease with increasing temperature,³

$$\mu(T) \approx \varepsilon_F \left[1 - \frac{\pi^2}{12} \left(\frac{k_B T}{\varepsilon_F} \right)^2 \right]. \quad (23.24)$$

The occupation number

$$f_{\text{FD}}(\vec{p}) = \langle n_p \rangle \approx \frac{1}{e^{\beta(\varepsilon_p - \mu)} + 1} \quad (23.25)$$

²Note that ε_F is one of the characterising parameters of a Fermi gas, while μ describes the thermodynamic chemical potential.

³First, recast (23.21) into the form

$$\mu \left[1 + \frac{\pi^2}{8} \left(\frac{k_B T}{\mu} \right)^2 \right]^{2/3} \approx \frac{\hbar^2}{2m} (6\pi^2 n)^{2/3} \quad (23.23)$$

and then solve iteratively for small values of $k_B T / \mu \approx k_B T / \varepsilon_F$.

is 1 for $\varepsilon_p < \mu$ and rapidly decays to 0 in a window of width $k_B T$ around μ . Expressing the Fermi energy through the Fermi momentum, $\varepsilon_F = \hbar^2 k_F^2 / 2m$, we see that inside the sphere $k < k_F$ the states are occupied at $T = 0$, while outside they are unoccupied. The sharp boundary at k_F is called the Fermi surface. Using this insight, we can easily calculate k_F and ε_F ,

$$\begin{aligned} N &= \sum_{k < k_F} 1 = \int_{k < k_F} d^3 n, \quad \text{where } \vec{k} = \frac{2\pi}{L} \vec{n} \\ &= \frac{V}{(2\pi)^3} \int_0^{k_F} dk 4\pi k^2 = \frac{V}{6\pi^2} k_F^3, \\ \rightarrow k_F^3 &= 6\pi^2 n. \end{aligned} \quad (23.26)$$

For energies ε_p that are g -times degenerate, the corresponding result is

$$k_F^3 = \frac{6\pi^2}{g} n. \quad (23.27)$$

For example, for an **electron gas** with $s = 1/2$, $g = 2$, and one obtains

$$k_F^3 = 3\pi^2 n. \quad (23.28)$$

In order to find the *internal energy*, we calculate

$$\begin{aligned} U &= \sum_p \varepsilon_p \langle n_p \rangle = \frac{V}{4\pi^2 \hbar^3 m} \int_0^\infty dp p^4 \langle n_p \rangle \stackrel{\text{PI}}{=} \frac{V}{4\pi^2 m \hbar^3} \int_0^\infty dp \frac{p^5}{5} (-\partial_p \langle n_p \rangle) \\ &= \frac{Vm\sqrt{2m}}{5\pi^2 \hbar^3 \beta^{5/2}} \underbrace{\int_0^\infty dy \frac{y^{5/2} e^{y-\beta\mu}}{(1 + e^{y-\beta\mu})^2}}_{(\beta\mu)^{5/2} + \dots} \approx \frac{Vm\sqrt{2m}\mu^{5/2}}{5\pi^2 \hbar^3} (1 + \dots) \\ &\approx \frac{3}{5} N \varepsilon_F \left[1 + \frac{5\pi^2}{12} \left(\frac{k_B T}{\varepsilon_F} \right)^2 + \dots \right]. \end{aligned} \quad (23.29)$$

With the Sommerfeld expansion to order T^2 , we can obtain the *specific heat*

$$C_V = N k_B \frac{\pi^2 k_B T}{2 \varepsilon_F} \xrightarrow{T \rightarrow 0} 0, \quad (23.30)$$

in agreement with the third law.

Finally, we find the thermal equation of state (or the partition function) to be

$$p = \frac{2}{3} \frac{U}{V} = \frac{2}{5} n \varepsilon_F \left[1 + \frac{5\pi^2}{12} \left(\frac{k_B T}{\varepsilon_F} \right)^2 + \dots \right]. \quad (23.31)$$

Note that the pressure, different from the classical ideal gas, *does not vanish* at $T \rightarrow 0$ but stays large. It corresponds to the pressure of a system at the *degeneracy temperature*

$$k_B T_F = \varepsilon_F = \frac{\hbar^2 k_F^2}{2m} = \frac{\hbar^2}{2m} \left(\frac{6\pi^2}{g} n \right)^{2/3}. \quad (23.32)$$

For a *metallic electron system* with $n \approx 10^{22} \text{ cm}^{-3}$, we find a degeneracy temperature

$$\begin{aligned} k_B T_F &\approx \frac{1.055 \cdot 10^{-34} \text{ m}^2 \text{ kg/s} \cdot 6.58 \cdot 10^{-16} \text{ eV s}}{2 \cdot 0.911 \cdot 10^{-30} \text{ kg}} \cdot \left(\frac{6\pi^2}{2} \frac{10^{22}}{10^{-6} \text{ m}^3} \right)^{2/3} \\ &\approx 1.69 \text{ eV}, \\ T_F &\approx 2 \cdot 10^4 \text{ K}. \end{aligned} \quad (23.33)$$

23.2.1 Particle-hole asymmetry

The $\propto T^2$ correction in the expression for the chemical potential $\mu(T)$, see Eq. (23.24), is a consequence of the **particle-hole asymmetry**: The density of states $\rho(\varepsilon)$ of the free Fermi gas in 3D is (we use $\rho(k) = k^2/\pi^2$ for fermions with spin and $\partial_k \varepsilon = \hbar^2 k/m$)

$$\rho(\varepsilon) = \frac{dk}{d\varepsilon} \rho(k) = \frac{mk}{\pi^2 \hbar^2} = \frac{m\sqrt{2m\varepsilon}}{\pi^2 \hbar^3} = \rho(\varepsilon_F) \sqrt{\frac{\varepsilon}{\varepsilon_F}}, \quad (23.34)$$

with $\rho(\varepsilon_F) = 3n/2\varepsilon_F$ the density of states for an electron gas of total density n including both spin directions.⁴ We see that $\rho(\varepsilon) \propto \sqrt{\varepsilon}$, i.e., there are more states available for the electrons at energies $\varepsilon > \varepsilon_F$ than for the holes at $\varepsilon < \varepsilon_F$. Thus, for $T > 0$ the chemical potential μ has to decrease such as to conserve the particle number, see Fig. 23.2. Indeed, in the integral providing the fixed number of particles N ,

$$N = V \int_0^\infty d\varepsilon f_{\text{FD}}(\varepsilon) \rho(\varepsilon), \quad (23.35)$$

the distribution function $f_{\text{FD}}(\varepsilon)$ is symmetric with respect to μ , but $\rho(\varepsilon)$ is *not*. For a free massive particle, we find a constant density of states $\rho(\varepsilon)$ (and hence symmetric around ε_F) in dimension $d = 2$, as $\rho(k) \propto k$ and $\partial_k \varepsilon \propto k$. In general, with a dispersion $\varepsilon(k) \propto k^n$ and $\rho(k) \propto k^{d-1}$

⁴Sometimes one may encounter the formula $\rho_\sigma = 3n/4\varepsilon_F$ for the density of states per spin σ in an electron gas.

in d dimensions, we have a constant $\rho(\varepsilon)$ if $d = n$. Note that electrons moving in a periodic potential have complicated dispersions and particle-hole symmetry may appear in specific regions of the Brioullin zone. In 1D mesoscopic physics, the dispersion around the Fermi energy is often linearized, $\varepsilon_k - \varepsilon_F \approx v_F \hbar(k - k_F)$, implying particle-hole symmetry within this approximation.

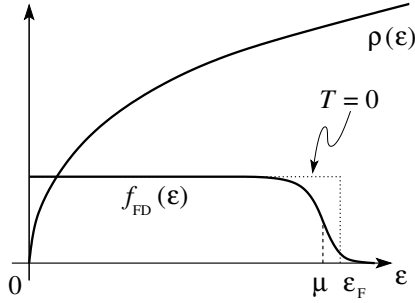


Fig. 23.2: Decrease of the chemical potential $\mu(T)$ in a free electron gas. For a free massive particle in 3D, the particle-hole asymmetry generates a negative shift in $\mu(T)$ when T increases.

23.2.2 Fermi gases in stars

The most prominent example of a Fermi gas is the electron gas in a metal. However, other interesting examples are known from astronomy, i.e., the matter from which stars are made. Depending on the properties of such matter, gravitational properties (i.e., mass density) versus pressure (i.e., equation of state), a star may be stable or not.

In *white dwarves* (made from a ${}^4\text{He}^{2+}$, e^- plasma), the pressure is generated by the Fermi gas of the electrons. At a mass density of $\rho \sim 10^6 \text{ g/cm}^3$, the electron density is of the order of $n \sim 10^{29} \text{ cm}^{-3}$ ($m_{{}^4\text{He}} \approx 7 \cdot 10^{-24} \text{ g}$) and constitutes a non-relativistic degenerate electron gas with a degeneracy temperature

$$T_F \sim 10^9 \text{ K}, \quad (23.36)$$

much larger than the effective temperature $T \sim 10^7 \text{ K}$ in the core of the star. We may notice a neat division of tasks between the nuclei (the ${}^4\text{He}$ -cores) and the electrons: The nucleons generate the attractive gravitational force while the electrons, through their degeneracy pressure, generate the stabilizing pressure. In the non-relativistic case, we obtain the polytrope (or equation of state)

$$p = K\rho^{1+2/3} = K\rho^{5/3}. \quad (23.37)$$

The exponent with the additional $2/3$ -contribution from ε_F coincides with the adiabat in the classical ideal gas. In the relativistic case, where $cp_F > m_e c^2$, one finds the exponent $\gamma = 4/3$. Using the hydrostatic equation for a stable star $\vec{\nabla} \cdot (\vec{\nabla} p / \rho) = -4\pi G \rho$ ($G = 6.674 \cdot 10^{-11} \text{ m}^3/\text{kg s}^2$ the gravitational constant) one can determine the structure of the star. Note that the stability of the star depends on its polytrope.

Finally, we consider a *neutron star*, where the reaction $p + e^- \rightarrow n + \nu$ has transformed nearly all protons and electrons into neutrons. Gravity and pressure are now generated by the same fermionic particle. Typical mass densities of around $\rho \sim 10^{15} \text{ g/cm}^3$ generate fermion densities $n \sim 10^{39} \text{ cm}^{-3}$ (non-relativistic case; $m_N \approx 1.7 \cdot 10^{-24} \text{ g}$) corresponding to degeneracy temperatures

$$T_F \sim 10^{12} \text{ K.} \quad (23.38)$$

The polytropes are $p = K\rho^{5/3}$ and $p = K\rho$ in the non-relativistic and relativistic limits, respectively. A typical star life goes through the sequence of a burning star to a white dwarf when the thermonuclear fuel runs out. If the mass of the core is too large a collapse takes the object to a neutron-star. Very complicated variants with supernova-explosions are possible.

23.2.3 Metallic sound

A similar division of tasks between mass density ρ and pressure p as in the white dwarf appears in the sound generation in metals. There, the electronic Fermi-sea generates the pressure p while the ions provide the mass density ρ . From the dispersion $\omega = (\partial_\rho p|_{s/n})^{1/2} q = c q$ (s/n is the entropy per particle in adiabatic sound) and using the compressibility

$$\kappa^{-1} = -V \partial_V p = n \partial_n p = \frac{5}{3} p = \frac{2}{3} n \varepsilon_F$$

in a degenerate Fermi gas, it follows that⁵

$$c^2 = \frac{1}{\rho \kappa} = \frac{2Zn\varepsilon_F}{nM3} = \frac{1}{3} \frac{mZ}{M} \frac{p_F^2}{m^2} = \frac{Z}{3} \frac{m}{M} v_F^2, \quad (23.39)$$

$$\rightarrow c = \sqrt{\frac{Zm}{3M}} v_F, \quad (23.40)$$

⁵In ^3He -fluids and in neutron stars $c = v_F/\sqrt{3}$.

where we have used that $\rho = nM/Z$ with n the electron density, M the mass of the ions, and Z is the ion's valency. Using

$$v_F \sim \frac{p_F}{m} \sim \frac{\hbar(3\pi^2 n)^{1/3}}{m} \sim 10^8 \text{ cm/s}, \quad (23.41)$$

we obtain a typical sound (or phonon) velocity $c_{\text{metal}} \sim 10^6 \text{ cm/s}$ in a metal. Note that in these collective excitations both the electrons and the ions are oscillating. What happens if only the electrons oscillate?

Chapter 24

Bosons

Material for this chapter can be found in the book by Kerson Huang. We study two types of bosonic systems: Photons are the quantum particles constituting the classical waves of electromagnetic radiation—when discussing photons in a general context, their particle number is not conserved and their chemical potential is zero, $\mu = 0$, meaning the cost of adding a photon is always its energy. Similar is the case for the phonons, the quantum particle associated with sound waves in a material. Quantum engineered photonic systems involving cavities and waveguides are an area of current research, in particular, the non-equilibrium aspect that is associated with the omnipresent drive and dissipation in this type of systems—steady states then replace the equilibrium states. The second class of bosons are massive particles, in particular atoms, that are experimentally studied in the form of trapped cold gases; below, we will discuss the properties of an ideal Bose-gas and its Bose-Einstein condensation.

24.1 Photons

The quantum mechanical description of the electromagnetic field is given by the *photon number states*

$$|n_{\vec{k}_1, \lambda_1}, n_{\vec{k}_2, \lambda_2}, \dots, n_{\vec{k}_i, \lambda_i}, \dots\rangle \quad (24.1)$$

in Fock space, where the \vec{k}, λ -mode is occupied $n_{\vec{k}, \lambda} \in \mathbb{N}_0$ times. The parameters \vec{k} and λ describe the *wavevector* and the *polarization* of the mode.

In a quantization volume L^3 , the wavevectors are given by

$$\vec{k} = \frac{2\pi}{L} \vec{n}, \quad \vec{n} \in \mathbb{Z}^3. \quad (24.2)$$

For every \vec{k} there exist two polarizations (either linear or circular), specified by the polarization vectors

$$\vec{\varepsilon}_x, \vec{\varepsilon}_y; \quad \vec{\varepsilon}_+, \vec{\varepsilon}_-; \quad |\vec{\varepsilon}_\lambda| = 1; \quad \vec{\varepsilon}_\lambda \cdot \vec{k}_\lambda = 0. \quad (24.3)$$

The corresponding field is *transversal*¹. The *gauge degree of freedom* is a consequence of the *masslessness* of photons and reduces the (spin-)polarizations to the values $\pm\hbar$ (Photon = massless spin-1 boson with $S_z \neq 0$). A further consequence of the lack of mass is the *linear dispersion*

$$\begin{aligned} E &= \sqrt{p^2 c^2 + m^2 c^4}|_{m=0} = pc, \\ E &= \hbar\omega, \quad \omega_k = ck, \quad p = \hbar k. \end{aligned} \quad (24.4)$$

Energy, momentum and angular momentum of the photon field are given by

$$\begin{aligned} E[\{n_{\vec{k},\lambda}\}] &= \sum_{\vec{k},\lambda} \hbar\omega_k n_{\vec{k},\lambda}, \\ \vec{P}[\{n_{\vec{k},\lambda}\}] &= \sum_{\vec{k},\lambda} \hbar\vec{k} n_{\vec{k},\lambda}, \\ \vec{S}[\{n_{\vec{k},\lambda}\}] &= \sum_{\vec{k}} \hbar\hat{k} (n_{\vec{k},+} - n_{\vec{k},-}). \end{aligned} \quad (24.5)$$

The corresponding operators are obtained by substituting the bosonic creation and annihilation operators

$$n_{\vec{k},\lambda} = a_{\vec{k},\lambda}^\dagger a_{\vec{k},\lambda}, \quad [a_{\vec{k},\lambda}, a_{\vec{k}',\lambda'}^\dagger] = \delta_{\vec{k}\vec{k}'} \delta_{\lambda\lambda'}. \quad (24.6)$$

We consider a box (cavity), whose walls are at a fixed temperature T and bring the radiation to equilibrium (alternatively, use a black body which captures all radiation and radiates a (different) spectral distribution determined by its temperature). We investigate the spectral distribution of the

¹ $\vec{\nabla} \cdot \vec{A} = 0$ in the Coulomb gauge; the physical fields are transverse, with Coulomb's law $\vec{\nabla} \cdot \vec{E} = 0$ and Gauss' law $\vec{\nabla} \cdot \vec{B} = 0$.

radiation in the box by calculating the grandcanonical partition function²
 \mathcal{Z}

$$\mathcal{Z} = \prod_{\vec{k}, \lambda} \frac{1}{1 - e^{-\beta \hbar \omega_k}}; \quad \log \mathcal{Z} = -2 \sum_{\vec{k}} \log(1 - e^{-\beta \hbar \omega_k}). \quad (24.7)$$

The mean occupation numbers are easily found,

$$\langle n_{\vec{k}, \lambda} \rangle = -\frac{1}{\beta} \frac{\partial \log \mathcal{Z}}{\partial \hbar \omega_k} = \frac{1}{e^{\beta \hbar \omega_k} - 1}, \quad (24.8)$$

and the internal energy is to be calculated from

$$U = -\frac{\partial}{\partial \beta} \log \mathcal{Z} = \sum_{\vec{k}, \lambda} \hbar \omega_k \langle n_{\vec{k}, \lambda} \rangle \quad (24.9)$$

with the equation of state following immediately from U (use that $\mathcal{Z} = Z$ and $p = -\partial_V F$, $F = -k_B T \ln Z$)

$$p = \frac{1}{\beta} \frac{\partial}{\partial V} \log \mathcal{Z} = -\frac{1}{3V} \frac{\partial}{\partial \beta} \log \mathcal{Z} = \frac{U}{3V}. \quad (24.10)$$

Here, we used that \mathcal{Z} depends on V through $x = \beta \omega_k$: With $\omega_k = 2\pi c |\vec{k}| V^{-1/3}$, we find $\partial_V = (\partial_V x) \partial_x$ and $\partial_V x = -x/3V$, $\partial_x = \omega_k^{-1} \partial_\beta$, thus $\partial_V = -(\beta/3V) \partial_\beta$. Note that (24.10) (for a mass $m = 0$) is different from (22.19) (for massive particles with $m > 0$, we have $\varepsilon_p \propto V^{-2/3}$ and $\partial_V = -(2\beta/3V) \partial_\beta$). To calculate U one substitutes

$$\sum_{\vec{k}, \lambda} \xrightarrow{V \rightarrow \infty} 2V \int \frac{d^3 k}{(2\pi)^3} = \frac{V}{\pi^2} \int_0^\infty dk k^2 = \frac{V}{\pi^2 c^3} \int_0^\infty d\omega \omega^2 \quad (24.11)$$

and finds the **Stefan-Boltzmann** law,

$$U = \frac{V \hbar}{\pi^2 c^3} \int_0^\infty d\omega \frac{\omega^3}{e^{\beta \hbar \omega} - 1} = \frac{V \hbar}{\pi^2 c^3 \hbar^4 \beta^4} \underbrace{\int_0^\infty dx \frac{x^3}{e^x - 1}}_{\pi^4/15}$$

²Alternatively consider a (classical) harmonic oscillator and calculate its (canonical) partition function by a gaussian integral over phase space (p, q) . Then go over to the quantum mechanical oscillator with energies $E_n = \hbar \omega n$ and calculate the corresponding result $Z = \text{Tr}[\exp(-\beta H)] = 1/(1 - \exp[-\beta \hbar \omega])$ (without quanta/particle interpretation); for many independent oscillators with frequencies ω_k , we find the canonical partition function $Z = \prod_k Z_k$, which is exactly the result quoted in (24.7). The equality $\mathcal{Z} = Z$ (or $F = \Omega$) follows from (21.20) with $z = 1$, $\mu = 0$. Thus, we can interpret the quanta of the harmonic oscillator as bosonic particles with chemical potential $\mu = 0$. The linear spectrum in n of the harmonic oscillator is an essential part of this argumentation.

$$= \frac{\pi^2 (k_B T)^4}{15 (\hbar c)^3} V = \frac{\pi^2 V}{15 \lambda^3} k_B T, \quad (24.12)$$

where we have defined the thermal wavelength λ through $2\pi k_B T = \hbar c k = \hbar c / \lambda$ in the last equation. Comparing this last expression $U \sim (V/\lambda^3) k_B T$ with the ideal gas result $U \sim N k_B T$, we see that the particle number N corresponds to the ratio V/λ^3 , i.e., the ‘particles are created by the temperature’ rather than being there ‘from the beginning’. The specific heat per unit volume continuously grows with temperature,

$$c_V = \frac{4\pi^2 k_B^4 T^3}{15 (\hbar c)^3} \xrightarrow{T \rightarrow \infty} \infty \quad (24.13)$$

as steadily additional modes get activated when temperature increases.

Of particular significance is Wien’s displacement law. We write the radiation law in the form (u is the spectral energy density; in the classical case it is given by the density of states times the thermal energy $k_B T$ per mode from the equipartition theorem)

$$u(\omega, T) = \frac{\omega^2}{\pi^2 c^3} k_B T f(\beta \hbar \omega), \quad (24.14)$$

separating the classical Rayleigh-Jeans law $u \propto \omega^2$ from the Planck factor $f = \beta \hbar \omega / (e^{\beta \hbar \omega} - 1)$ (the modification due to the quantization of energy, cf. (24.12)); the latter cuts off of the spectral energy density at large frequencies, see Fig. 24.1. Wien’s displacement law, the shift of the maximum in $u(\omega)$

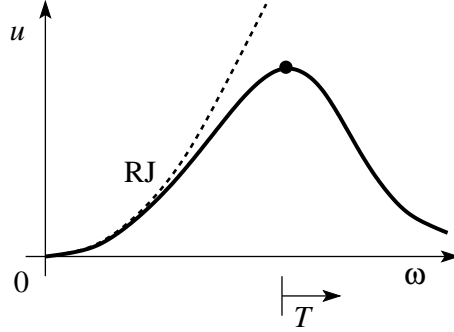


Fig. 24.1: The Rayleigh-Jeans law valid at low frequencies is cutoff by the Planck factor $f(\beta \hbar \omega)$. Their competition leads to Wien’s displacement law, telling that the maximum in $u(\omega)$ shifts linearly in T .

with the temperature T follows from the requirement $\partial_\omega u = 0$,

$$\rightarrow [2f + x f']|_{x_m} = 0, \quad \beta \hbar \omega_m = x_m \rightarrow \omega_m \propto T \quad (24.15)$$

since x_m is merely a number determined by the shape of the function $f(x)$.

The Stephan-Boltzmann law (24.12) can be derived classically: With the Maxwellian radiation pressure³ $p = u/3$ and the internal energy $U = Vu(T)$, we find the entropy differential

$$\begin{aligned} dS &= \frac{C_V}{T} dT + \frac{1}{T} \left(\frac{\partial U}{\partial V} + p \right) dV \\ \rightarrow dS &= \frac{V}{T} \frac{\partial u}{\partial T} dT + \frac{4}{3} \frac{u}{T} dV, \end{aligned} \quad (24.16)$$

and using the condition for integrability $\partial^2 S / \partial T \partial V = \partial^2 S / \partial V \partial T$, we arrive at the differential equation

$$\begin{aligned} \frac{1}{T} \frac{\partial u}{\partial T} &= \frac{4}{3} \frac{\partial}{\partial T} \left(\frac{u}{T} \right) \\ \rightarrow \frac{du}{u} &= 4 \frac{dT}{T} \quad \rightarrow \quad \ln u = 4 \ln T + \text{const.} \end{aligned} \quad (24.17)$$

$$\rightarrow u = aT^4. \quad (24.18)$$

Planck's derivation of the radiation law

$$u(\omega, T) = \frac{\hbar \omega^3}{\pi^2 c^3} \frac{1}{\exp(\hbar \omega / k_B T) - 1} \quad (24.19)$$

makes use of an ingenious trick:⁴ He considered a charged harmonic oscillator of frequency ω coupled to a radiation field of temperature T and found the relation

$$U = U_{\text{HO}} = \frac{\pi c^3}{2\omega^2} u(\omega, T), \quad (24.20)$$

between the spectral weight $u(\omega, T)$ of the radiation field and the time averaged energy U of the oscillator. Next, he considered the entropy at constant volume

$$dS = \frac{dU}{T} \quad \rightarrow \quad \frac{dS}{dU} = \frac{1}{T}. \quad (24.21)$$

He makes use of a more precise Wien displacement law that was based on improved measurements of Paschen,

$$u(\omega, T) \sim \omega^3 e^{-\alpha \omega / T}, \quad (24.22)$$

³Average the pressure $(2S/c) \cos^2 \theta$, \mathbf{S} the Pointing vector, over a half-sphere in order to describe the reflection (at an angle θ) from a boundary.

⁴One may find more about this sketch in arXiv:quant-ph/0010008, "...ich dachte mir nicht viel dabei...", Planck's uneven path to the radiation formula.

and using (24.20) for large frequencies ω , he found

$$\frac{dS}{dU} = \frac{1}{T} = -\frac{1}{\alpha\omega} \ln U, \quad \frac{d^2S}{dU^2} = -\frac{1}{\alpha\omega U}. \quad (24.23)$$

For *small* ω , he makes use of the equipartition theorem $U = k_B T$ for the harmonic oscillator

$$\rightarrow \frac{dS}{dU} = \frac{1}{T} = \frac{k_B}{U}, \quad \frac{d^2S}{dU^2} = -\frac{k_B}{U^2}. \quad (24.24)$$

Interpolating, one obtains

$$\frac{d^2S}{dU^2} = -\frac{1}{\alpha\omega U + U^2/k_B} \quad (24.25)$$

Using that $\partial_U^2 S = -T^{-2} \partial_U T$, one finds a differential equation for $U(T)$, which is solved by

$$U = \frac{\alpha\omega k_B}{e^{\alpha\omega/T} - 1}. \quad (24.26)$$

Making use of (24.20) one finally arrives at Planck's law.

24.2 Phonons

Quite often, the vacuum state of a condensed matter system can be viewed as an elastic deformable medium. The excitations of this medium, the phonons, then play the same role as the photons in the electromagnetic vacuum. The important difference is that condensed matter systems are composed of atoms: The *continuum* nature of the vacuum in electromagnetism is replaced by the *discrete* atomic lattice of condensed matter. Instead of infinitely many modes the condensed matter system has $3N$ modes corresponding to the $3N$ degrees of freedom of the N atoms. The spectrum in momentum space is reduced to the Brillouin zone (BZ) due to the fact that large wave numbers $k > K_n$ (with K_n the reciprocal lattice vectors) do not make sense because such modes resolve the (inexistent) structure *between* the lattice sites. Furthermore, the spectrum at the BZ edge is deformed as the phonons “scatter on the lattice” and are thus reflected, implying that $\partial\omega/\partial k|_{\text{BZ}} = 0$. Thirdly, there are now “3 polarisations”, two transversal and one longitudinal, for every \vec{k} -vektor, transversal shear modes and longitudinal compression modes. The latter correspond to the hydrodynamic sound in gas (longitudinal compression modes)—the shear modes are a characteristic of the hard condensed matter system (they appear neither in liquids

nor in gases). Note that the classification of longitudinal (\parallel) and transversal (\perp) (to \vec{k}) directions only make sense for special high-symmetry directions of \vec{k} in the BZ.

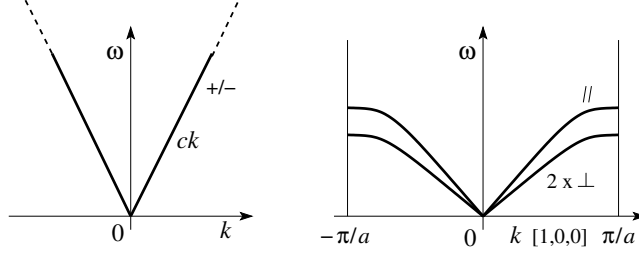


Fig. 24.2: Summary of differences between photons und phonons. Extended k -space versus finite Brillouin zone, overall linear double-degenerate spectrum versus spectral deformation at the BZ edges and three modes, double degenerate transverse plus longitudinal, velocity of light versus velocity of sound at small k , about 10^4 to 10^5 times smaller.

For a simple isotropic elastic medium, we can *model* the phonon dispersion by

$$\omega_\lambda = c_\lambda |\vec{k}| \quad (24.27)$$

and the *density of states* has the form (with $d^3k \sim 4\pi k^2 dk$)

$$\begin{aligned} \rho(\omega) d\omega &\simeq \sum_\lambda 4\pi k^2 \left(\frac{L}{2\pi}\right)^3 dk \\ &= \sum_\lambda \frac{V}{2\pi^2} \omega^2 d\omega \left(\frac{1}{c_\lambda^3}\right) \\ &= \frac{3V\omega^2}{2\pi^2 \bar{c}^3} d\omega, \\ \frac{3}{\bar{c}^3} &= \frac{2}{c_\perp^3} + \frac{1}{c_\parallel^3}. \end{aligned} \quad (24.28)$$

We find the size of the isotropic *Brillouin zone* through the constraint

$$\begin{aligned} \int_0^{\omega_{\max}} \rho(\omega) d\omega &= \frac{V}{2\pi^2} \frac{\omega_{\max}^3}{\bar{c}^3} = \frac{V}{2\pi^2} k_{\text{BZ}}^3 = 3N \\ \rightarrow k_{\text{BZ}}^3 &= 6\pi^2 n, \quad \text{cf. Eq. (23.28)} \\ \omega_{\max} &= \bar{c} k_{\text{BZ}}. \end{aligned} \quad (24.29)$$

The internal energy follows from Eqs. (24.12) and (24.29),

$$\begin{aligned} U &= \frac{3V}{2\pi^2\bar{c}^3} \int_0^{\omega_{\max}} d\omega \omega^2 \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} \\ \rightarrow \frac{U}{N} &= \underbrace{\frac{3\hbar}{2\pi^2 n \bar{c}^3} \frac{1}{(\hbar\beta)^4}}_{9(k_B T)^4 / (\hbar\omega_{\max})^3} \int_0^{\beta\hbar\omega_{\max}} dt \frac{t^3}{e^t - 1}. \end{aligned} \quad (24.30)$$

The *Debye function* $\mathcal{D}(x)$ behaves as

$$\mathcal{D}(x) \equiv \frac{3}{x^3} \int_0^x dt \frac{t^3}{e^t - 1} = \begin{cases} 1 - 3x/8 + x^2/20 - \dots, & x \ll 1, \\ \pi^4/5x^3 + \mathcal{O}(e^{-x}), & x \gg 1, \end{cases} \quad (24.31)$$

for small and large arguments x and we obtain the result

$$\frac{U}{N} \approx 3k_B T \begin{cases} 1 - \frac{3}{8} \frac{\Theta}{T} + \dots, & T \gg \Theta, \\ \frac{\pi^4}{5} \left(\frac{T}{\Theta}\right)^3, & T \ll \Theta. \end{cases} \quad (24.32)$$

Here, we have introduced the Debye temperature

$$k_B \Theta = \hbar\omega_{\max}. \quad (24.33)$$

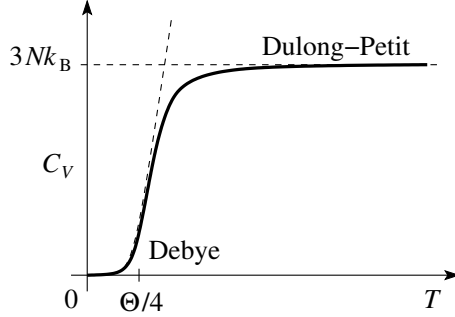


Fig. 24.3: Specific heat of phonons. As the spectrum is cutoff at the Debye energy $k_B \Theta$ (at the BZ), the specific heat saturates at the Dulong-Petit value.

Above Θ all modes are excited and we find the classical result of Dulong-Petit for the specific heat of a solid

$$C_V = 3Nk_B, \quad T > \Theta, \quad (24.34)$$

(6 degrees of freedom per *bound* atom). For $T \ll \Theta$, excitations are added continuously starting from zero and we obtain the (SB)-Debye result

$$C_V = \frac{12\pi^4}{5} Nk_B \left(\frac{T}{\Theta}\right)^3, \quad T < \Theta. \quad (24.35)$$

Here is a collection of Debye temperatures

Ag	215 K
Al	428 K
Au	170 K
Cu	343 K
He	25 K
C (diamond)	2250 K
C (graphite)	413 K
Si	645 K
MgO	750 K
SiO ₂ (quartz)	470 K
NaCl	310 K

Small values of Θ describe “soft”, while large values refer to “hard” materials.

In general, every excitation/degree of freedom contributes to C_V ; it is their spectral density as it follows from the dispersion $\omega \sim k^n$ and the dimensionality of the system ($\rightarrow \rho \sim k^{d-1}$) that determines the temperature dependence of C_V . Simple dimensional considerations, valid for small T such that the cutoff in the integral is on T and not on ρ , then provide the scaling behavior

$$\begin{aligned}
 U &\sim \int d^d k \frac{\hbar \omega}{e^{\beta \hbar \omega} - 1} \sim \frac{1}{\beta} \int^{\omega \sim T} dk \rho, \\
 C_V &\sim \int^{\omega \sim T} dk \rho \sim \int dk k^{d-1} \sim k^d \sim \omega^{d/n} |^{\omega \sim T} \sim T^{d/n}.
 \end{aligned}$$

When all modes are exhausted (with the integral cutoff⁵ by ρ), we have $U \propto T$ and C_V is constant. For phonons in 3D, we have $n = 1$, $d = 3$,

$$C_V \sim T^3. \quad (24.36)$$

Massive Bosons in 3D have $n = 2$ and hence we expect $C_V \sim T^{3/2}$, see Eq. (24.59) below. Similar estimates can be made for Fermions when properly accounting for the Pauli blockade.

24.3 Ideal Bose-gas: Bose-Einstein-condensation

We recapitulate the results from Sec. 22.0.2:

Partition function:

$$\mathcal{Z}(V, T, z) = \prod_{\vec{p}} (1 - z e^{-\beta \varepsilon_p})^{-1}. \quad (24.37)$$

⁵Note that the density of states vanishes outside the BZ, i.e., for $k > k_{\text{BZ}}$.

Equation of state:

$$p = \frac{k_B T}{V} \log \mathcal{Z} = \frac{k_B T}{\lambda^3} g_{5/2}(z) - \frac{k_B T}{V} \log(1 - z). \quad (24.38)$$

Particle density:

$$n = \frac{z}{V} \frac{\partial}{\partial z} \log \mathcal{Z} = \frac{1}{\lambda^3} g_{3/2}(z) + \frac{1}{V} \frac{z}{1 - z}. \quad (24.39)$$

Occupation number:

$$\langle n_{\vec{p}} \rangle = \frac{1}{e^{\beta(\varepsilon_p - \mu)} - 1}. \quad (24.40)$$

$$\begin{aligned} g_{5/2}(z) &= -\frac{4}{\sqrt{\pi}} \int_0^\infty dx x^2 \log(1 - z e^{-x^2}) = \sum_{\ell=1}^\infty \frac{z^\ell}{\ell^{5/2}}, \\ g_{3/2}(z) &= \frac{4}{\sqrt{\pi}} \int_0^\infty dx \frac{x^2}{e^{x^2}/z - 1} = \sum_{\ell=1}^\infty \frac{z^\ell}{\ell^{3/2}}. \end{aligned} \quad (24.41)$$

Again, we first solve the fugacity relation

$$n\lambda^3 = g_{3/2}(z) + \frac{\lambda^3}{V} \frac{z}{1 - z} \quad (24.42)$$

for z . The function $g_{3/2}(z)$, see Fig. 24.4, can be expanded in the range $0 \leq z \leq 1$ using the series expansion $\sum_\ell z^\ell / \ell^{3/2}$,

$$g_{3/2}(z) = \begin{cases} z + z^2/2\sqrt{2} + z^3/3\sqrt{3} + \dots, & z \ll 1, \\ \zeta(3/2) = 2.612, & z = 1. \end{cases} \quad (24.43)$$

The slope of $g_{3/2}(z)$ diverges in $z = 1$. We see that for $n\lambda^3 < 2.612$, equation (24.42) can be satisfied with a value $z < 1$ deriving from $n\lambda^3 = g_{3/2}(z)$ with the second term playing no role in the limit $V \rightarrow \infty$. In contrast when $n\lambda^3 > 2.612$, we find a solution to (24.42) requires that $z = 1 - \mathcal{O}(1/V)$ and the second term has to compensate the term $n\lambda^3 - 2.612$. In the limit $V \rightarrow \infty$, we find $z \rightarrow 1$ and using (24.42), the ground state occupation $\langle n_0 \rangle = \langle n_{\vec{p}=0} \rangle$

$$\langle n_0 \rangle = \frac{z}{1 - z} = \frac{V}{\lambda^3} (n\lambda^3 - \underbrace{n_{\text{BE}}\lambda^3}_{=2.612}) \propto V \quad (24.44)$$

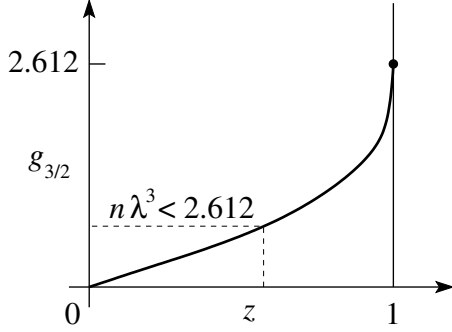


Fig. 24.4: Sketch for $g_{3/2}(z)$ in the interval $[0, 1]$, linear at small z and saturating at $z = 1$ with infinite slope. Small densities $n\lambda^3 < 2.612$ can be brought to the normal phase without condensation as $z < 1$.

assumes a *macroscopic* value $\langle n_0 \rangle \propto N$: This is the phenomenon of **Bose-Einstein-condensation** (BEC). One often talks about condensation in momentum space: Instead of condensing in a liquid drop (real space), the condensation occurs in a quantum state with definite \vec{p} . We will come back to the peculiarities of the condensate later on. The constraint

$$n\lambda^3 = n \left(\frac{2\pi\hbar^2}{mk_B T} \right)^{3/2} = 2.612 \quad (24.45)$$

defines the **critical temperature** for the condensation

$$k_B T_{\text{BE}} = \underbrace{\frac{\hbar^2 n^{2/3}}{2m}}_{\text{Energie}} \cdot \underbrace{\frac{4\pi}{2.612^{2/3}}}_{6.63} \quad (24.46)$$

at constant density n , or, alternatively, the **critical density**

$$n_{\text{BE}} = \frac{2.612}{\lambda^3} \quad (24.47)$$

at constant temperature T ; both criteria signal the onset of condensation. These strong quantum effects appear when ~ 2.6 particles can be found in a de Broglie volume λ^3 . For $T \rightarrow 0$ the condensation is complete and $\langle n_0 \rangle / N = 1$. At finite temperatures, the fraction of particles in the condensate is found from (24.44), see Fig. 24.5 for a sketch,

$$\frac{\langle n_0 \rangle}{N} = \begin{cases} (1 - n_{\text{BE}}/n) = [1 - (T/T_{\text{BE}})^{3/2}], & T < T_{\text{BE}}, \\ 0, & T > T_{\text{BE}}. \end{cases} \quad (24.48)$$

One may ask why only the $\vec{p} = 0$ state condenses, as the momentum $p_1 = 2\pi\hbar/L$ also goes to 0 for $V = L^3 \rightarrow \infty$. Thus, we would have

$$\frac{N}{V} = \frac{1}{\lambda^3} g_{3/2}(z) + \frac{\langle n_0 \rangle}{V} + \frac{\langle n_1 \rangle}{V} \quad (24.49)$$

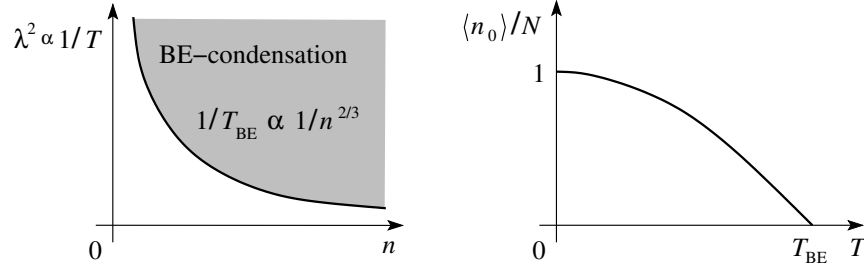


Fig. 24.5: Left: particles Bose condense when sufficiently many particles occupy the de Broglie volume λ^3 . Right: the fraction of particles in the condensate grows continuously with lowering temperature T .

with the last term the next in the sum $\sum_{\vec{p}} \langle n_{\vec{p}} \rangle$. We fix T below T_{BE} and

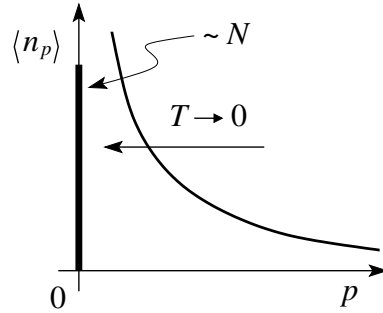


Fig. 24.6: The condensation with macroscopic occupation is confined to only the ground state $p = 0$; the occupations $\langle n_{p>0} \rangle$ for finite momenta remains finite, i.e., do not scale with N .

determine $\langle n_1 \rangle$ from Eq. (24.40) for a large volume and $\beta \varepsilon_1 \ll 1$ (we use $\langle n_0 \rangle/V = 1/\alpha > 0$),

$$\frac{\langle n_1 \rangle}{V} = \frac{1}{V} \frac{1}{e^{\beta \varepsilon_1}/z - 1} \stackrel{z=e^{\beta \varepsilon_0}/(1+\alpha/V)}{\approx} \frac{1/V}{\alpha/V + \beta(\varepsilon_1 - \varepsilon_0)} \propto \frac{T}{V^{1/3}} \xrightarrow{V \rightarrow \infty} 0, \quad (24.50)$$

i.e., in the thermodynamic limit $V \rightarrow \infty$ only the ground state condenses; while $\vec{p} = 0$ is occupied macroscopically ($\langle n_0 \rangle \propto N$) all $\langle n_{p>0} \rangle$ remain finite, see Fig. 24.6.

Knowing the fugacity z as a function of density n and temperature T ,

$$\frac{g_{3/2}(z)}{g_{3/2}(1)} = \frac{n}{n_{BE}} = \left(\frac{T_{BE}}{T} \right)^{3/2}, \quad n\lambda^3 < 2.612, \quad T > T_{BE}, \quad n < n_{BE}, \quad (24.51)$$

$$z = 1, \quad n\lambda^3 > 2.612, \quad T < T_{BE}, \quad n > n_{BE},$$

we can evaluate the equation of state (see 24.38) of the ideal Bose-gas

$$p = \frac{k_B T}{\lambda^3} g_{5/2}(z) \rightarrow k_B T n \left(1 - \frac{n \lambda^3}{4\sqrt{2}} + \dots \right), \quad \begin{cases} T > T_{\text{BE}}, \\ n < n_{\text{BE}}, \end{cases} \quad (24.52)$$

$$p = \frac{k_B T}{\lambda^3} g_{5/2}(1) = 1.342 \frac{k_B T}{\lambda^3} \propto T^{5/2} \xrightarrow{T \rightarrow 0} 0, \quad \begin{cases} T < T_{\text{BE}}, \\ n > n_{\text{BE}}. \end{cases} \quad (24.53)$$

This result defines the transition line $p_{\text{BE}}(T)$ in the p - T diagram, see Fig. 24.7,

$$p_{\text{BE}}(T) = g_{5/2}(1) \left(\frac{m}{2\pi\hbar^2} \right)^{3/2} (k_B T)^{5/2}. \quad (24.54)$$

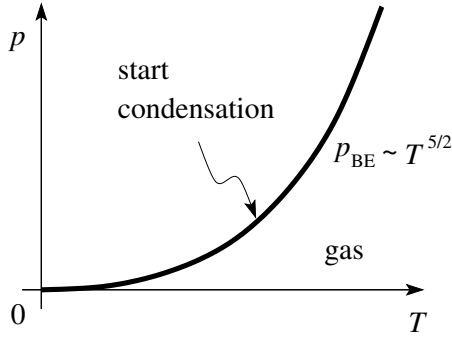


Fig. 24.7: Transition line for the Bose Einstein condensation in the p - T diagram.

In a further step, we determine the p - v diagram for the ideal Bose gas. For a volume $v < v_{\text{BE}}(T)$ ($v = 1/n$ is the volume per particle, $n > n_{\text{BE}}(T)$), we have $p = p_{\text{BE}} = \text{const.}$ This is the vapour pressure above the condensate with the gas-phase assuming a fraction $n_{\text{BE}}(T)/n$ of the density (the remaining particles are in the condensate and generate no pressure). At the transition, $n\lambda^3 = \lambda^3/v = g_{3/2}(1)$ and hence $T_{\text{BE}}(v) \propto v^{-2/3}$. Using the expression (24.53) for the pressure at the transition, we find the transition line in the p - v diagram,

$$p_{\text{BE}}(v) = \frac{k_B T_{\text{BE}}(v)}{\lambda^3(T_{\text{BE}}(v))} g_{5/2}(1) = \frac{2\pi\hbar^2}{m v^{5/3}} \frac{g_{5/2}(1)}{[g_{3/2}(1)]^{5/3}} \rightarrow p_{\text{BE}} v^{5/3} = \text{const.} \quad (24.55)$$

The phase diagrams in the Figs. 24.7 and 24.8 resemble those for the real gas with the liquid-gas transition. We can make an analogy between Bose gases and real classical gases with the Bose condensed (C) phase corresponding to the liquid and the Bose gas phase (G) with $p > 0$ corresponding to

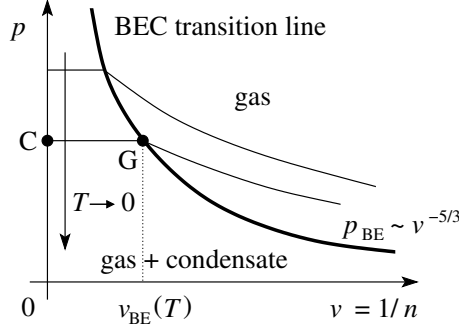


Fig. 24.8: p - v diagram for the Bose System with a gas phase and a Bose-Einstein condensate. Curves shown are isothermes with fixed temperature T .

the usual gas phase (G). The condensate replacing the liquid, however, has different, atypical properties. In particular, it does not sustain a pressure, and thus cannot oppose any compression. One then may observe the following similarities and differences: First, the transition between G and C in Fig. 24.8 obeys the Clausius-Clapeyron relation with the volume jump from $v_G = v_{BE}(T)$ to $v_C = 0$, $\Delta v = v_{BE}(T)$. Correspondingly, one obtains the latent heat (per particle joining the condensate)

$$\ell = \underbrace{\frac{dp_{BE}}{dT}}_{(5/2)g_{5/2}(1)k_B/\lambda^3} T \Delta v = \frac{5}{2} \frac{g_{5/2}(1)}{g_{3/2}(1)} k_B T, \quad (24.56)$$

where we have used $v_{BE}/\lambda^3 = 1/g_{3/2}(1)$ for the volume jump. On the other hand, we cannot reach the $p > p_{BE}(T)$ region of the p - T diagram. If we want to cool further below T_{BE} , there is no state that can resist a pressure and thus the system follows the transition line $p_{BE}(T)$ until it reaches $p = 0$, $T = 0$ (up to the impossibility to reach $T = 0$ exactly), see Fig. 24.9. The full condensation can then be realized by i) ‘pressing’ the gas into the condensate at p and T (reduce the volume), or ii) reducing the temperature at constant volume (stretching the de Broglie wavelength beyond the system volume); naturally, variations of these two methods are possible.

Further, we show (see below) that $C_V(T)$ only shows a kink. The interpretation of the BEC as a phase transition of first or second order according to the Ehrenfest classification is not really applicable, but note that L&L classify the transition as being second order.⁶

⁶There is no finite free energy surface that defines a finite slope $\partial_p g$ or $\partial_T g$ below the $p_{BE}(T)$ transition line.

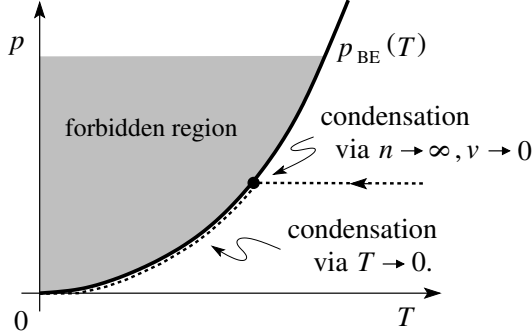


Fig. 24.9: The Bose gas condenses completely by compressing the gas into the condensate at T_{BE} and fixed pressure or by lowering the temperature $T \rightarrow 0$ at fixed volume and following the transition line $p_{\text{BE}}(T)$.

For the internal energy U , the entropy S , and the specific heat C_V , one finds

$$\frac{U}{N} = \frac{3}{2}pv = \begin{cases} \frac{3}{2} \frac{k_B T}{n\lambda^3} g_{5/2}(z), \\ \frac{3}{2} \frac{k_B T}{n\lambda^3} g_{5/2}(1), \end{cases} \quad (24.57)$$

$$\frac{S}{Nk_B} = \begin{cases} \frac{5}{2} \frac{g_{5/2}(z)}{n\lambda^3} - \log z, \\ \frac{5}{2} \frac{g_{5/2}(1)}{n\lambda^3}, \end{cases} \quad (24.58)$$

$$\frac{C_V}{Nk_B} = \begin{cases} \frac{15}{4} \frac{g_{5/2}(z)}{n\lambda^3} - \frac{9}{4} \frac{g_{3/2}(z)}{g_{1/2}(z)}, \\ \frac{15}{4} \frac{g_{5/2}(1)}{n\lambda^3}. \end{cases} \quad (24.59)$$

The internal energy appears from the pressure and $U = (3/2)pV$. The entropy S follows from

$$\begin{aligned} G &= N\mu = Nk_B T \log z, \\ S &= -\frac{\partial G}{\partial T}, \end{aligned} \quad (24.60)$$

$$\frac{\partial z}{\partial T} = -\frac{5z}{2T} \frac{g_{5/2}(z)}{g_{3/2}(z)}, \quad (24.61)$$

where the last equation is derived from $\partial_T [g_{5/2}(z) = p\lambda^3/k_B T]$ (using the relation $\partial_z g_n = g_{n-1}/z$). The specific heat is found via

$$\begin{aligned} C_V &= \frac{\partial U}{\partial T}, \\ \frac{\partial z}{\partial T} &= -\frac{3z}{2T} \frac{g_{3/2}}{g_{1/2}}, \end{aligned} \quad (24.62)$$

and making use of $\partial_T [g_{3/2}(z) = n\lambda^3]$. The dependence $C_V(T)$, see Fig. 24.10, shows a kink at the transition.

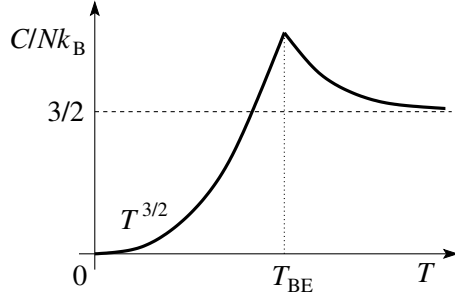


Fig. 24.10: Specific heat of a BEC with a cusp at T_{BE} .

In the condensate phase the entropy is (we make use of Eq. (24.48))

$$\frac{S}{N} = \left(\frac{T}{T_{\text{BE}}} \right)^{3/2} s = \frac{n_{\text{BE}}}{n} s = [1 - \langle n_0 \rangle / N] s \quad (24.63)$$

with the entropy per particle in the gas phase

$$s = \frac{g_{5/2}(1)}{g_{3/2}(1)} \frac{5}{2} k_B, \quad (24.64)$$

i.e., the entropy S (and the pressure p) are generated purely by the gas phase. For $T \rightarrow 0$ the gas phase vanishes along with the entropy and the specific heat, in accordance with the third law of thermodynamics. The latent heat $\ell = T\Delta s = Ts$ originating from (24.64) agrees with (24.56). Note that the latent heat and volume jump are realized when hitting the transition point $p_{\text{BE}}(T)$ and either condensing the gas by further decrease of p and T along the transition line or by squeezing the gas into the condensate at fixed p and T . Thereafter, there is no new thermodynamic phase which can be further cooled with a new thermodynamic Gibbs function, i.e., we cannot calculate an entropy/volume jump as a jump in derivatives of two free energy surfaces across the transition line as is the case for a usual first-order transition.

Chapter 25

Relativistic Particles: Klein-Gordon equation*

This is the first of the last three chapters in which we investigate (massive) relativistic particles and their quantum mechanical properties. In the present chapter, we study spin 0 bosonic particles, e.g., π mesons or K mesons.¹ In the next chapter, we will introduce the Dirac equation which describes fermionic spin 1/2 particles (such as the electron).

The relativistic description adds a new phenomenon to quantum mechanics: that of anti-particles. For example, we find the π^- as the anti-particle of π^+ , the \bar{K}^0 as the anti-particle of K^0 , the positron e^+ as the counterpart to the electron e^- . The appearance of anti-particles is a consequence of the structure of the theory: In the context of a relativistic description, we require that our equations are covariant.² The resulting equations show solutions with negative energy and negative conserved norm, which we will interpret as anti-particles.

The Klein-Gordon and Dirac theories are *one-particle* theories in first-quantized form. They are a priori unsuitable for describing particle production in a consistent manner; the latter is possible within the framework of a field theory where both the particles and the mediators of the interac-

¹ π^\pm and K^\pm mesons couple with the electric charge $\pm e$ to the electromagnetic field, while K^0 mesons are neutral. Mesons are made from two bound quarks.

²The principle of covariance tells that the laws of nature have an identical form in systems linked by Lorentz- (in SRT) or by general coordinate- (in GRT) transformations. Accordingly, they can be written as tensor equations.

tion, i.e., the gauge bosons³ are described in a second-quantized form. Such a complete and consistent description of relativistic many-body theories is the topic of quantum field theory (QFT) in the next semester. Before carrying out this last step, however, we need the one-particle description in a covariant form.

ZZ

25.1 Klein-Gordon equation

According to the principle of correspondence, the following procedure seems appropriate to arrive at a Schrödinger equation for relativistic particles: The relativistic dispersion connecting energy and momentum of a particle with mass m is

$$E = c\sqrt{p^2 + m^2c^2} = mc^2\sqrt{1 + p^2/m^2c^2}, \quad (25.1)$$

$$\left[\approx mc^2 + p^2/2m + (p^2/2m)(p/2mc)^2 \right]$$

and with

$$E \leftrightarrow i\hbar\partial_t, \quad \vec{p} \leftrightarrow -i\hbar\vec{\nabla}, \quad (25.2)$$

the first-order equation describing the time evolution of the scalar wave function $\Psi(\vec{r}, t)$ for a spin 0 particle takes the form

$$i\hbar\partial_t\Psi(\vec{r}, t) = c\sqrt{m^2c^2 - \hbar^2\nabla^2} \Psi(\vec{r}, t). \quad (25.3)$$

The relation (25.3) is not manifestly covariant, however: time and space coordinates do not appear on an equal footing.⁴ We obtain relativistic covariance if we square (25.1)

$$E^2 - c^2p^2 - m^2c^4 = 0, \quad (25.4)$$

and use the correspondence principle to construct the *Klein-Gordon equation*

$$\left[\frac{1}{c^2} \partial_t^2 - \nabla^2 + \frac{m^2c^2}{\hbar^2} \right] \Psi(\vec{r}, t) = 0,$$

$$\left[\frac{1}{c^2} \partial_t^2 - \nabla^2 + \frac{1}{\lambda^2} \right] \Psi(\vec{r}, t) = 0. \quad (25.5)$$

³Gauge bosons are γ with spin 0 (photons, electromagnetic interaction), W^\pm , Z^0 weak (or vector) bosons with spin 1 (weak interaction, e.g., in the process $n \rightarrow p + (W^- \rightarrow e^- + \bar{\nu}_e)$) or the spin 1 gluons g (strong interaction, e.g., in and between atomic nuclei).

⁴Furthermore, (25.3) is non-local on the scale of the Compton wavelength and thus not relativistic causal.

For an explicit covariant form, we use $x^\mu = (t, \vec{r})$, $\partial_\mu = d/dx^\mu = (\partial_t, \vec{\nabla})$, and $\partial^\mu = d/dx_\mu = (\partial_t, -\vec{\nabla})$ to write

$$[\partial_\mu \partial^\mu + m^2] \Psi(x) = 0, \quad (25.6)$$

where we have set $\hbar = 1 = c$. The Compton wavelength

$$\lambda = h/mc \quad (25.7)$$

determines the length scale in (25.5); e.g., for the π^\pm mesons, it is of order femtometer,

$$\lambda = \lambda/2\pi = \hbar/mc = 1.4 \cdot 10^{-15} \text{ m} = 1.4 \text{ fm}. \quad (25.8)$$

Comparing $E = c\sqrt{p^2 + m^2c^2}$ and $E^2 = c^2p^2 + m^2c^4$, we seem to have ‘only squared’ an expression. The difference between (25.3) and (25.5) is, however, more dramatic: equation (25.5) is of second order in time and its solution $\Psi(\vec{r}, t)$ is fixed by two conditions $\Psi(\vec{r}, t_0)$ and the derivative $\partial_t \Psi(\vec{r}, t_0)$. Furthermore, taking the square provides us with free solutions involving both positive and negative energies,

$$\begin{aligned} \Psi(\vec{r}, t) &= e^{i(\vec{p} \cdot \vec{r} - Et)/\hbar}, \quad \text{with} \\ E &= \pm c\sqrt{p^2 + m^2c^2}. \end{aligned} \quad (25.9)$$

ZZ

The negative energy E solutions show a new type of behavior where the energy decreases with increasing momentum, see Fig. ?? . A dispersion of this type also occurs in solid state physics at band electrons on the upper band edge, where it gives rise to the definition of ‘holes’ — accordingly we will find anti-particles in the relativistic context.

If the particle is charged (charge q , including sign), so we have to identify the derivatives through their calibration-invariant extensions

$$\begin{aligned} i(\partial_t, -\vec{\nabla}) &= i\partial^\mu = p^\mu \rightarrow p^\mu - qA^\mu, \\ i\partial_t &\rightarrow i\partial_t - 2\pi c\phi/\Phi_0, \\ -i\vec{\nabla} &\rightarrow -i\vec{\nabla} - 2\pi\vec{A}/\Phi_0, \end{aligned} \quad (25.10)$$

replacing with ⁵

$$\begin{aligned} A^\mu &= (\phi, \vec{A}), \\ \Phi_0 &= \frac{hc}{q} = \text{Flussquant}, \end{aligned} \quad (25.11)$$

⁵ $\Phi_0 > 0$ for positively charged particles.

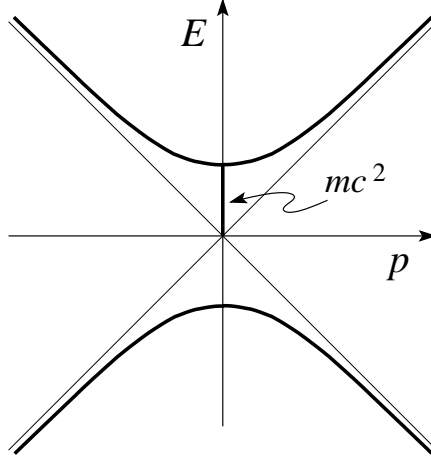


Fig. 25.1: Dispersion of massive relativistic particles in the Klein-Gordon equation. The positive branch has a conventional dispersion with $\vec{v} = \partial_{\vec{p}} E$ parallel to \vec{p} , the negative branch is abnormal with \vec{v} and \vec{p} anti-parallel, the characteristic of a hole or anti-particle.

and get the equations (in covariant and in conventional form)

$$\begin{aligned} & [-(i\partial_\mu - qA_\mu)(i\partial^\mu - qA^\mu) + m^2] \Psi(x) = 0, \\ & \left[-\frac{1}{c^2} \left(i\partial_t - \frac{2\pi c}{\Phi_0} \phi \right)^2 + \left(i\vec{\nabla} + \frac{2\pi}{\Phi_0} \vec{A} \right)^2 + \frac{m^2 c^2}{\hbar^2} \right] \Psi(\vec{r}, t) = 0. \end{aligned} \quad (25.12)$$

25.1.1 Preserved norm

The Klein-Gordon equation (25.12) has maintained 4th current density $\partial_\mu j^\mu = 0$ ($\partial^\mu = d/dx_\mu = (\partial_t, -\vec{\nabla})$) (we need two wave functions Ψ_a and Ψ_b to properly define the effect of the operators)

$$j^\mu = \frac{1}{2m} \left(\Psi_a^* i \overset{\leftrightarrow}{\partial}^\mu \Psi_b - 2qA^\mu \Psi_a^* \Psi_b \right), \quad (25.13)$$

where the arrow indicates the direction in which the operator is acting,

$$\begin{aligned} \Psi_a^* i \overset{\leftrightarrow}{\partial}^\mu \Psi_b &\equiv \Psi_a^* i \overset{\rightarrow}{\partial}^\mu \Psi_b - \Psi_a^* i \overset{\leftarrow}{\partial}^\mu \Psi_b \\ &= \Psi_a^* (i\partial^\mu \Psi_b) - (i\partial^\mu \Psi_a^*) \Psi_b. \end{aligned} \quad (25.14)$$

Equivalent applies

$$\partial_t \rho(\vec{r}, t) + \vec{\nabla} \cdot \vec{j} = 0, \quad (25.15)$$

with the density

$$\rho(\vec{r}, t) = \frac{\hbar}{2mc^2} \left[\Psi^* \left(i\partial_t - \frac{2\pi c}{\Phi_0} \phi \right) \Psi - \Psi \left(i\partial_t + \frac{2\pi c}{\Phi_0} \phi \right) \Psi^* \right] \quad (25.16)$$

and the current density

$$\vec{j}(\vec{r}, t) = \frac{\hbar}{2m} \left[\Psi^* \left(\frac{\vec{\nabla}}{i} - \frac{2\pi}{\Phi_0} \vec{A} \right) \Psi - \Psi \left(\frac{\vec{\nabla}}{i} + \frac{2\pi}{\Phi_0} \vec{A} \right) \Psi^* \right]. \quad (25.17)$$

While \vec{j} , as defined in (25.17), matches the nonrelativistic expression, the density is no longer given by $\Psi^* \Psi$ but by the expected value of the calibration invariant time derivative. The size $\int d^3r \Psi^* \Psi$ or more generally $\int d^3r \Psi_a^* \Psi_b = \langle \Psi_a | \Psi_b \rangle$ is no longer preserved ($\langle \cdot | \cdot \rangle$ does not define a conserved norm, but a conserved norm is determined by the size

$$\int d^3r \left(\Psi_a^* \frac{i \overleftrightarrow{\partial}_t}{2m} \Psi_b - \frac{qA^0}{m} \Psi_a^* \Psi_b \right) \quad (25.18)$$

defined because with $0 = \partial_\mu j^\mu = \partial_t \rho + \vec{\nabla} \cdot \vec{j}$ we find through integration

$$\begin{aligned} 0 &= \partial_t \int_V d^3r j^0 + \int_V d^3r \vec{\nabla} \cdot \vec{j}, \\ &= \partial_t \int_V d^3r j^0 + \underbrace{\int_{\partial V} d^3r \vec{n} \cdot \vec{j}}_{=0}. \end{aligned} \quad (25.19)$$

We are therefore introducing a new scalar product and a new norm (here including \hbar and c defined),

$$\langle \Psi_a | \Psi_b \rangle_{\text{KG}} = \int d^3r \left(\Psi_a^* \frac{i \hbar \overleftrightarrow{\partial}_t}{2mc^2} \Psi_b - \frac{q\phi}{mc^2} \Psi_a^* \Psi_b \right), \quad (25.20)$$

$$\|\Psi\|_{\text{KG}}^2 = \langle \Psi | \Psi \rangle_{\text{KG}}. \quad (25.21)$$

With regard to this dot product, a completely orthonormal basis can be obtained from the free solutions; With

$$\begin{aligned} \Psi_{\vec{k}}^{(\pm)}(\vec{r}, t) &= N_{\vec{k}} e^{i(\vec{k} \cdot \vec{r} \mp E_k t)/\hbar}, \\ E_k &= +\sqrt{p^2 c^2 + m^2 c^4}, \\ \vec{k} &= 2\pi \vec{n}/L, \\ \vec{p} &= \hbar \vec{k}, \end{aligned} \quad (25.22)$$

we get the normalization

$$\begin{aligned} \langle \Psi_{\vec{k}}^{(\pm)} | \Psi_{\vec{k}}^{(\pm)} \rangle_{\text{KG}} &= N_{\vec{k}}^2 \int d^3r e^{-i(\vec{k} \cdot \vec{r} \mp E_k t)/\hbar} \frac{i \hbar \overleftrightarrow{\partial}_t}{2mc^2} e^{i(\vec{k} \cdot \vec{r} \mp E_k t)/\hbar} \\ &= N_{\vec{k}}^2 V \frac{(\pm E_k)}{mc^2} = \pm 1 \end{aligned} \quad (25.23)$$

and thus the normalized wave function

$$\Psi_{\vec{k}}^{(\pm)}(\vec{r}, t) = \sqrt{\frac{mc^2}{VE_k}} e^{i(\vec{p} \cdot \vec{r} \mp E_k t)/\hbar}. \quad (25.24)$$

Note that $\langle \Psi_{\vec{k}}^{(-)} | \Psi_{\vec{k}}^{(-)} \rangle = -1$ has a negative ‘norm’. All of these states are also orthogonal,

$$\begin{aligned} \langle \Psi_{\vec{k}}^{(\pm)} | \Psi_{\vec{k}'}^{(\pm)} \rangle_{\text{KG}} &= \frac{mc^2}{V\sqrt{E_k E_{k'}}} \int d^3r e^{-i(\vec{k} \cdot \vec{r} \mp \omega_k t)} \frac{\pm E_k \pm E_{k'}}{2mc^2} e^{i(\vec{k}' \cdot \vec{r} \mp \omega_{k'} t)} \\ &= \pm \delta_{\vec{k}, \vec{k}'}, \end{aligned} \quad (25.25)$$

$$\begin{aligned} \langle \Psi_{\vec{k}}^{(+)} | \Psi_{\vec{k}'}^{(-)} \rangle_{\text{KG}} &= \frac{mc^2}{V\sqrt{E_k E_{k'}}} \int d^3r e^{-i(\vec{k} \cdot \vec{r} - \omega_k t)} \frac{E_k - E_{k'}}{2mc^2} e^{i(\vec{k}' \cdot \vec{r} + \omega_{k'} t)} \\ &= 0, \end{aligned} \quad (25.26)$$

$$\langle \Psi_{\vec{k}}^{(-)} | \Psi_{\vec{k}'}^{(+)} \rangle_{\text{KG}} = 0. \quad (25.27)$$

The wave functions (25.24) result in ‘twice as many base vectors’ as before, since two energies $\pm E_k$ are now allowed for each wave vector \vec{k} . We therefore have to determine ‘twice as many’ coefficients if we want to represent any wave function Ψ as the superposition of the base vectors $\Psi_{\vec{k}}^{(\pm)}$. This is consistent with the fact that Ψ ‘twice as many’ initial conditions $\Psi(t=0)$ and $\partial_t \Psi(t=0)$ are required to clearly define them.

With the Klein-Gordon scale product $\langle \cdot | \cdot \rangle_{\text{KG}}$ and the complete orthonormal system $\{\Psi_{\vec{k}}^{(\pm)}\}$ we could almost build a Hilbert space on which the Klein-Gordon equation (25.5) defines the dynamics. For the norm, however, we expect the property that it is positive (semi-) definite, which is obviously not the case here. We can also no longer interpret the obtained density $\rho(\vec{r}, t)$ as probability density, since $\rho(\vec{r}, t)$ can take negative values. Historically, these defects in theory have been reason enough to dismiss them. Surprisingly, it was shown (again) that the equations are ‘smarter than the physicists’: the physical origin of the above formal problems lies in the existence of the anti-particles, which ‘the equations knew’, but was unknown to physicists at the time. One can interpret the Klein-Gordon equations with their conserved norm and their positive and negative energy solutions consistently so that the states with negative energy and norm describe precisely the anti-particles. What is then conserved is no longer the particle density but the particle charge density, where the charge can be an electromagnetic charge

or another quantum number, for example, strangeness or the hyper charge in the K^0 , \bar{K}^0 - Mesons.⁶ A particle of the charge then creates a pair with charge ± 1 so the total charge $+1$ is retained. In a field theory, anti-matter can then be described consistently, in the sense that the states with negative energy and norm are reinterpreted as anti-particles with positive energy and norm, so that again a semi-definite scalar product $\langle \cdot | \cdot \rangle \geq 0$ and energies $E > 0$ result for all states.

Although a consistent interpretation is only possible in second quantized form, it is still appropriate to understand the first quantized form of Klein-Gordon theory. Dirac theory is less problematic, where we find anti-particles in a consistent one-particle theory where the norm of the wave functions is always positive. We next examine the production of anti-particles.

25.2 Pair generation at potential barriers

We consider a particle (eg a Pion π^+) with charge $q = e > 0$, which is currently t_0 by a wave packet from states $\Psi_k^{(+)}$. In homogeneous space this state evolves in the sector of positive energy and norm and we can interpret $\rho(\vec{r}, t)$ as particle density. If the particle encounters an inhomogeneity, for example a potential barrier $V(x) = V\Theta(x)$, see Fig. 25.2, components also arise via the Klein-Gordon dynamics $\Psi_k^{(-)}$, that is, anti-particles are created.

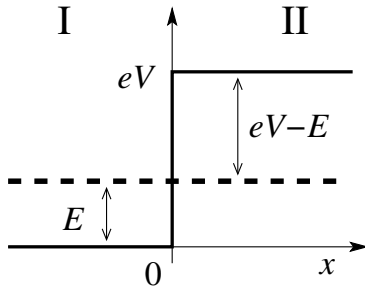


Fig. 25.2: An incident particle (region I) with energy E scatters at a potential level of high eV , which in the case of enough high energy $eV > E + mc^2$ creates a particle - anti-particle pair.

Here we write $V > 0$ as the electromagnetic potential, then $eV > 0$ is the potential energy. According to (25.12) the dynamics through the Klein-

⁶There is no charge quantum number for the π^0 meson and π^0 is its own antiparticle; with Ψ also Ψ^* is a solution, so that Ψ can be chosen real and a density $\rho = 0$ results.

Gordon equation is in the form

$$(\hbar^2 \partial_t^2 - c^2 \hbar^2 \partial_x^2 + 2eV(x) i\hbar \partial_t - e^2 V^2(x) + m^2 c^4) \Psi(x) = 0 \quad (25.28)$$

given and we can take the solution

$$\Psi_I(x, t) = Ae^{i(px-Et)/\hbar} + Be^{-i(px+Et)/\hbar}, \quad (25.29)$$

$$\Psi_{II}(x, t) = Ce^{i(Px-Et)/\hbar} + De^{-i(Px+Et)/\hbar}, \quad (25.30)$$

with

$$\begin{aligned} p &= \sqrt{E^2 - m^2 c^4}/c, \\ P &= \sqrt{(eV - E)^2 - m^2 c^4}/c, \end{aligned} \quad (25.31)$$

with $E > 0$ for regions I and II. (??) to (??) is obviously a solution to (??) in regions I and II, see Fig. 25.28. A total solution in I and II requires the calculation of the coefficients A to D , such that $\Psi_I(0) = \Psi_{II}(0)$ and $\partial_x \Psi_I(0) = \partial_x \Psi_{II}(0)$ ist ⁷. The waves $\propto A, B$ are right- and left-running waves in I; to understand the physical meaning of the solution we consider an incident wave packet with E around E_0 , then ⁸

$$\begin{aligned} & \int_{E_0-\Delta E}^{E_0+\Delta E} dE e^{\pm i(\sqrt{E^2 - m^2 c^4} x/c \mp Et)/\hbar} \\ & \approx e^{\pm i(p_0 x \mp E_0 t)/\hbar} \int_{E_0-\Delta E}^{E_0+\Delta E} dE e^{\pm iE(x/v_0 \mp t)/\hbar} \\ & = e^{\pm i(p_0 x \mp E_0 t)/\hbar} f(x/v_0 \mp t) \end{aligned} \quad (25.32)$$

with $f(t) = 2 \sin(\Delta E t/\hbar)/(\Delta E/\hbar)$. In (25.32) the upper (lower) sign stands for type A (type B) waves. Thus in area I it applies that the phase and group speeds run in parallel and type A , (B) waves run to the right (left).

Next we consider the type C and D waves in region II. The momentum $P = \sqrt{(eV - E)^2 - m^2 c^4}/c$ is for $eV - E < mc^2$ or $E < eV < E + mc^2$

⁷The initial conditions are not relevant for a stationary solution. If we construct wave packets from the stationary solutions, the initial conditions become relevant.

⁸we develop around E_0 ,

$$\begin{aligned} \sqrt{E^2 - m^2 c^4} x/c \mp Et & \approx \sqrt{(E_0 + \delta E)^2 - m^2 c^4} x/c \mp (E_0 + \delta E)t \\ & = p_0 x \mp E_0 t + c^{-1} \partial_{\delta E} \sqrt{\cdots}|_{E_0} \delta E x \mp \delta E t \\ & = p_0 x \mp E_0 t + (\delta E/v_0) x \mp \delta E t, \end{aligned}$$

with $c^{-1} \partial_{\delta E} \sqrt{\cdots}|_{E_0} = (1/2c\sqrt{\cdots})2E_0 = E_0/p_0 c^2 = 1/v_0$, since according to special relativity theory applies $p = mv/\sqrt{1 - v^2/c^2} \leftrightarrow v = p/\sqrt{m^2 + p^2/c^2} = pc^2/E$.

purely imaginary and we get the well-known exponentially damped waves⁹: For $E < eV < E + mc^2$ the π^+ meson is reflected by the potential barrier, but what if the barrier is very high, $eV > E + mc^2$? Then $P = \sqrt{[(eV - E)^2 > mc^2] - mc^2}/c > 0$ real and we get propagating waves again, but these waves have special properties: their norm is negative,

$$\begin{aligned} \int dx e^{\mp i(Px \mp Et)/\hbar} \left(i\hbar \frac{\overleftrightarrow{\partial}}{\partial t} / 2 - eV \right) e^{\pm i(Px \mp Et)/\hbar} \\ = \left(-(-\frac{\overleftarrow{E}}{2}) + \frac{\overrightarrow{E}}{2} - eV \right) L \\ = (E - eV)L < -mc^2L < 0, \end{aligned} \quad (25.33)$$

and the relationship between phase and group speed is just inverse, that is anti-parallel,

$$\begin{aligned} \int_{E_0 - \Delta E}^{E_0 + \Delta E} dE e^{\pm i(\sqrt{(eV - E)^2 - m^2 c^4} x / c \mp Et)/\hbar} \\ \approx e^{\pm i(P_0 x \mp E_0 t)/\hbar} \int_{E_0 - \Delta E}^{E_0 + \Delta E} dE e^{\mp iE \left(\frac{eV - E_0}{P_0 c^2} x \pm t \right) / \hbar} \\ = e^{\pm i(P_0 x \mp E_0 t)/\hbar} f(x/u_0 \pm t), \end{aligned} \quad (25.34)$$

with $u_0 = P_0 c^2 / (eV - E_0) > 0$ (the upper (lower) sign applies to C type (D type) waves). Furthermore it follows from (??) $u_0 = P_0 c^2 / E_{\text{kin}}$ that the particle has a positive kinetic energy even though it is in a forbidden zone.

We are now looking for a solution in which a π^+ meson from region I is currently hitting and scattering $t = 0$. The signal in II should run to the right and is therefore of type D . We'll find a consistent solution if

$$\begin{aligned} A + B &= D, & A &= 1 \\ p(A - B) &= -PD, \end{aligned} \quad (25.35)$$

and thus

$$\begin{aligned} \Psi_{\text{I}}(x, t) &= e^{i(px - Et)/\hbar} - \underbrace{\frac{P + p}{P - p} e^{-i(px + Et)/\hbar}}_{\text{reflected wave}}, \\ \Psi_{\text{II}}(x, t) &= - \underbrace{\frac{2p}{P - p} e^{-i(Px + Et)/\hbar}}_{\text{transmitted wave}} \end{aligned} \quad (25.36)$$

⁹tunnels under the potential barrier.

is. We get the solution superposed to wave packets

$$\begin{aligned}\Psi_{\text{I}}(x, t) &= e^{i(p_0 x - E_0 t)/\hbar} f(x/v_0 - t) - \frac{P_0 + p_0}{P_0 - p_0} e^{-i(p_0 x + E_0 t)/\hbar} f(x/v_0 + t), \\ \Psi_{\text{II}}(x, t) &= -\frac{2p_0}{P_0 - p_0} e^{-i(P_0 x + E_0 t)/\hbar} f(x/u_0 - t).\end{aligned}\quad (25.37)$$

The transmission and reflection coefficients (to be calculated from the integrated densities / KG standard) are

$$\begin{aligned}T &= -\left(\frac{2p_0}{P_0 - p_0}\right)^2 \frac{eV - E_0}{E_0} \frac{u_0}{v_0} \\ &= -\frac{4p_0 P_0}{(P_0 - p_0)^2} < 0,\end{aligned}\quad (25.38)$$

$$R = \left(\frac{P_0 + p_0}{P_0 - p_0}\right)^2 > 1,\quad (25.39)$$

$$R + T = 1.\quad (25.40)$$

The factor $(eV - E_0)/E_0$ comes from $(i\hbar\partial_t/2 - eV)$ in the KG scalar product $\langle \cdot | \cdot \rangle_{\text{KG}}$, the Factor u_0/v_0 of integrals of the type

$$\int dx g(x/v - t) = v \int dx' g(x' - t).\quad (25.41)$$

We find an interesting result. The π^+ meson hits the barrier and is likely to be reflected $R > 1$: More π^+ mesons are coming back than we sent. Added to that is that in the region II for $t > 0$ (causal) particles appear that run to the right. They have negative norm, positive kinetic energy and, in terms of amount, result in the same number of particles as additional π^+ mesons for $t > 0$ to fly back to the left ($R + T = 1$ applies). The only possible consistent interpretation of this result is that these particles are negative loaded objects. Because for negatively charged particles the potential has the form ¹⁰ according to Fig. 25.3.

The total energy of the particle is negative

$$E_{\text{total}} = E_{\text{kin}} + E_{\text{pot}} = (eV - E) - eV = -E.\quad (25.42)$$

The negatively charged particle ‘sees’ an attractive potential and the positive kinetic energy is no longer a problem. The particles appearing in region II

¹⁰The potential for positively charged particles is sketched in Fig. 25.2. In addition, the shape for negative particles is as shown in Fig. 25.3

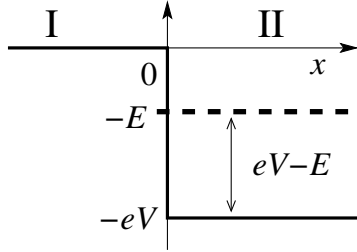


Fig. 25.3: Potential for negatively charged particles as they arise in the scattering of high-energy positively charged particles at a potential level of the height eV .

are π^- mesons, the anti-particles to the π^+ mesons. $\pi^- - \pi^+$ pairs are thus generated at the barrier and the barrier separates the particles by moving the π^+ to the left and the π^- to the right into the respective attractive ones Regions. Note that an induced particle - anti-particle production is possible for bosons since double occupation of the reflected state is not a fundamental problem.¹¹ For fermions (see Dirac equation in the next capital) the induced pair generation is prohibited because of the Pauli principle. As a counterpart we find the spontaneous electron-positron there ($e^- - e^+$) pair generation in a strong electric field, similar to the spontaneous particle-hole pair production via tunnel effect on a semiconductor pn -" transition in the breakdown area, see Fig. ?? ,

In Fig. ?? we illustrate the pair generation by shooting a π^+ - meson at a potential level and thereby opening $\pi^+ - \pi^-$ - meson pairs on the computer. With the values $E = (13/12)mc^2$, $eV = (221/60)mc^2$, we create two from one particle at $t < 0$ additional particles, so that the reflected and transmitted wave packets show the weights $1 + 1.02$ and -1.02 . The process, interpreted as scattering and pair generation, receives charge and energy,

$$\int e \rho(x, t) dx = e, \quad (\text{Ladung}),$$

$$E + 0 = 2.02E + 1.02[(eV - E) - eV]. \quad (25.43)$$

The next thing we do is work on the formalism: instead of a 2th order dynamic for a scalar field, our goal is a 1st order dynamic for a 'two component spinor'. We get a formalism that is closer to the known formalism of nonrelativistic quantum mechanics.

¹¹In analogy to the spontaneous + induced emission of photons, $\pi^+ - \pi^-$ meson pairs can also be generated spontaneously.

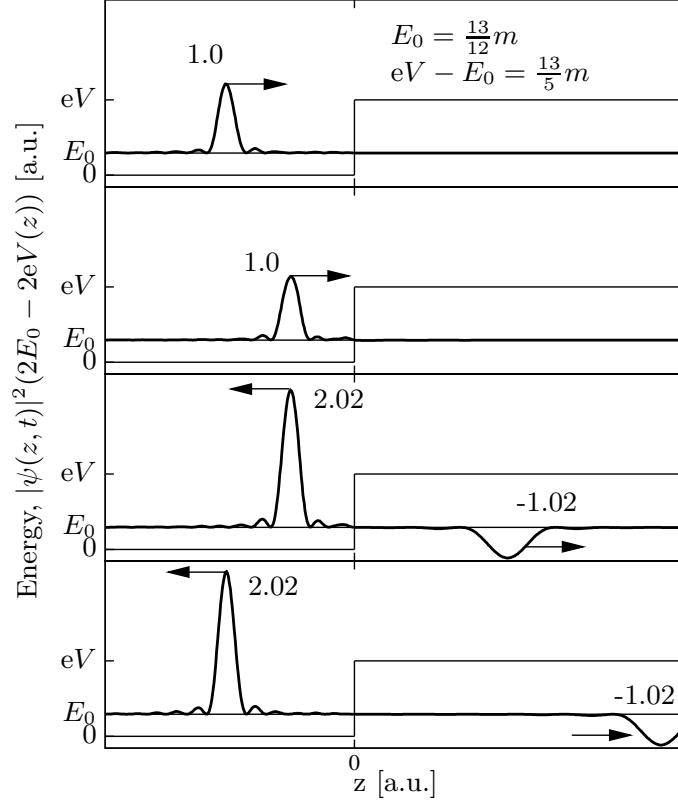


Fig. 25.4: Klein-Gordon evolution of a wave packet ($c = 1$). The wave packet is centered around the energy $E_0 = (13/12)m$ and the step has the height $eV = (13/5)m + E_0$. The charge density $\|\Psi(z, t)\|_{\text{KG}}$ is drawn for $t = -2, -1, 1, 2$. The integrated charge density is given for the individual packets; an induced pair generation is observed at the barrier.

25.3 Two-component representation

Every second-order differential equation can be divided into two first differential equations according to general principles. Order to be converted. So the goal is to create a dynamic in the form

$$i\hbar\partial_t\Phi = H\Phi \quad (25.44)$$

to get with

$$\Phi = \begin{pmatrix} \Phi_+ \\ \Phi_- \end{pmatrix} \quad (25.45)$$

a two vector and H a 2×2 matrix¹². With (25.44) we have our first order dynamics back and we can use the usual rules of nonrelativistic QM, especially the perturbation theory described in Chapter 8. Also a dynamic with a matrix structure is a good preparation for the Dirac equation, where we have 4th Vectors and 4×4 matrices must be treated (2 components for particles and anti-particles, 2 components for spin $\pm 1/2$). A disadvantage of this two-component representation is that (25.44) is no longer manifestly covariant; The old covariant form (25.6) is used for field quantization.

Instead of trivial

$$\Phi = \begin{pmatrix} \Psi \\ \partial_t \Psi \end{pmatrix}, \quad (25.46)$$

we choose the symmetrized combinations

$$\begin{aligned} \Phi_+ &= \frac{1}{2} \left(1 + \frac{i\hbar}{mc^2} \partial_t - \frac{q\phi}{mc^2} \right) \Psi, \\ \Phi_- &= \frac{1}{2} \left(1 - \frac{i\hbar}{mc^2} \partial_t + \frac{q\phi}{mc^2} \right) \Psi. \end{aligned} \quad (25.47)$$

With

$$H = \begin{pmatrix} \frac{(\vec{p}-q\vec{A}/c)^2}{2m} + q\phi + mc^2 & \frac{(\vec{p}-q\vec{A}/c)^2}{2m} \\ -\frac{(\vec{p}-q\vec{A}/c)^2}{2m} & -\frac{(\vec{p}-q\vec{A}/c)^2}{2m} + q\phi - mc^2 \end{pmatrix} \quad (25.48)$$

the vector Φ obeys the first-order differential equation

$$i\hbar\partial_t\Phi = H\Phi. \quad (25.49)$$

¹²the indexes \pm in Φ are not identical to those in $\Psi_k^{(\pm)}$.

With the help of the Pauli matrices

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (25.50)$$

l can be written compactly according to H (25.48)

$$H = \left(\frac{(\vec{p} - q\vec{A}/c)^2}{2m} + mc^2 \right) \tau_3 + \frac{(\vec{p} - q\vec{A}/c)^2}{2m} i\tau_2 + q\phi \mathbb{1}. \quad (25.51)$$

The two-component representations (25.47) and (25.51) of the Klein-Gordon theory is equivalent to the scalar form (25.12). In the 2- Component representation, the Klein-Gordon scalar product can be elegantly written; use of

$$\begin{aligned} \Psi &= \Phi_+ + \Phi_-, \\ \frac{i\hbar\partial_t\Psi}{mc^2} &= (\Phi_+ - \Phi_-) + \frac{q\phi}{mc^2} (\Phi_+ + \Phi_-) \end{aligned} \quad (25.52)$$

into the dot product (25.20)

$$\begin{aligned} \langle \Psi_a | \Psi_b \rangle_{\text{KG}} &= \int d^3r \left(\Psi_a^* i \frac{\overleftrightarrow{\partial}_t}{2mc^2} \Psi_b - \frac{q\phi}{mc^2} \Psi_a^* \Psi_b \right) \\ &= \int d^3r (\Phi_{a+}^* \Phi_{b+} - \Phi_{a-}^* \Phi_{b-}) \\ &= \int d^3r \underbrace{(\Phi_{a+}^*, \Phi_{a-}^*)}_{\Phi_a^\dagger \tau_3 \equiv \bar{\Phi}_a} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \underbrace{\begin{pmatrix} \Phi_{b+} \\ \Phi_{b-} \end{pmatrix}}_{\Phi_b} \\ &= \int d^3r \bar{\Phi}_a \Phi_b = \langle \Phi_a | \Phi_b \rangle_{2-\text{KG}}. \end{aligned} \quad (25.53)$$

So let's redefine the Klein-Gordon adjunct of the state Φ by

$$\begin{aligned} \bar{\Phi} &= \Phi^\dagger \tau_3, \\ &= (\Phi_+^*, -\Phi_-^*), \end{aligned} \quad (25.54)$$

with

$$\Phi = \begin{pmatrix} \Phi_+ \\ \Phi_- \end{pmatrix} \quad \text{und} \quad \Phi^\dagger = (\Phi_+^*, \Phi_-^*), \quad (25.55)$$

so we can write the Klein-Gordon scalar product in the usual form

$$\langle \Psi_a | \Psi_b \rangle_{\text{KG}} = \int d^3r \bar{\Phi}_a \Phi_b = \langle \Phi_a \Phi_b \rangle_{2-\text{KG}}. \quad (25.56)$$

We can also use the operator's Klein-Gordon adjoint and define A as

$$\bar{A} = \tau_3 (A^T)^* \tau_3 = \tau_3 A^\dagger \tau_3, \quad (25.57)$$

with the usual adjoint A^\dagger . With this we get

$$\langle \Phi_a | \bar{A} \Phi_b \rangle = \langle \Phi_b | A \Phi_a \rangle^*. \quad (25.58)$$

Note that $H \neq H^\dagger$, so H is not Hermitian in the usual sense, but H is Klein-Gordon-Hermitian,¹³

$$H = \bar{H} = \tau_3 H^\dagger \tau_3. \quad (25.59)$$

25.3.1 Solutions of the Klein-Gordon equation

The solutions of the free Klein-Gordon equation take in the 2-component representation in the following form: with

$$i\hbar \partial_t \Phi = \begin{pmatrix} \frac{p^2}{2m} + mc^2 & \frac{p^2}{2m} \\ -\frac{p^2}{2m} & -\frac{p^2}{2m} - mc^2 \end{pmatrix} \Phi, \quad (25.60)$$

and the approach

$$\Phi_{\vec{k}}^{(\pm)} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} e^{i(\vec{p} \cdot \vec{r} \mp Et)/\hbar}, \quad (25.61)$$

where $\vec{p} = \hbar \vec{k}$ and $\vec{k} = 2\pi \vec{n}/L$, we need to solve the system.

$$\begin{pmatrix} \pm E - mc^2 - \frac{p^2}{2m} & -\frac{p^2}{2m} \\ \frac{p^2}{2m} & \pm E + mc^2 + \frac{p^2}{2m} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0 \quad (25.62)$$

The spectrum follows from the acute equation

$$E^2 = m^2 c^4 + p^2 c^2 \quad (25.63)$$

¹³The term $\propto i\tau_2$ transformed according to $\tau_3(i\tau_2)^\dagger \tau_3 = i\tau_2$.

and the eigenfunctions result in inclusive standardization (note the consistency with ?? taking into account the signs of the \mathbf{k} 's)

$$\begin{aligned}\Phi_{\vec{k}}^{(+)} &= \frac{E + mc^2}{2\sqrt{mc^2EV}} \left(-\left(\frac{pc}{E+mc^2}\right)^2 \right) e^{i(\vec{p}\cdot\vec{r}-Et)/\hbar}, \\ \Phi_{-\vec{k}}^{(-)} &= \frac{E + mc^2}{2\sqrt{mc^2EV}} \left(-\left(\frac{pc}{E+mc^2}\right)^2 \right) e^{-i(\vec{p}\cdot\vec{r}-Et)/\hbar}.\end{aligned}\quad (25.64)$$

This results in standardization

$$\begin{aligned}\langle \Phi_{\vec{k}}^{(+)} | \Phi_{\vec{k}}^{(+)} \rangle &= 1, \\ \langle \Phi_{\vec{k}}^{(-)} | \Phi_{\vec{k}}^{(-)} \rangle &= -1\end{aligned}\quad (25.65)$$

where we needed that

$$\begin{aligned}(E + mc^2)^2 &= 2mc^2(E + mc^2) + p^2c^2, \quad \text{und} \\ (E + mc^2)^4 - p^4c^4 &= 4mc^2E(E + mc^2)^2\end{aligned}\quad (25.66)$$

is. The solutions (25.64) of the free ones are very similar to the solution of the Dirac equation, see later. We are now investigating how the anti-particles appear in 2-component form. Before that, as an intermezzo, the non-relativistic limit.

25.4 Non-relativistic limit

Let $p^2c^2 \ll m^2c^4$ and we introduce the reduced energy ε ,

$$\varepsilon = E - mc^2 \stackrel{\text{freie T.}}{=} \sqrt{m^2c^4 + p^2c^2} - m^2c^2 \approx \frac{p^2}{2m}, \quad (25.67)$$

The reduced energy ε no longer contains any rest mass. With the approach

$$\Phi = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} e^{-iEt/\hbar} \quad (25.68)$$

we get using (25.48)

$$\begin{aligned}\varepsilon\alpha &= \left(\frac{(\vec{p} - q\vec{A}/c)^2}{2m} + q\phi \right) \alpha + \frac{(\vec{p} - q\vec{A}/c)^2}{2m} \beta, \\ (2mc^2 + \varepsilon)\beta &= -\frac{(\vec{p} - q\vec{A}/c)^2}{2m} \alpha - \left(\frac{(\vec{p} - q\vec{A}/c)^2}{2m} - q\phi \right) \beta.\end{aligned}\quad (25.69)$$

Divide the second equation by $2mc^2$ there

$$\beta = -\frac{(\vec{p} - \frac{q}{c}\vec{A})^2}{4m^2c^2}\alpha + \mathcal{O}\left[\left(\frac{\varepsilon}{mc^2}\right)^2\right] \quad (25.70)$$

and inserting it into the first equation of (25.69) results

$$\varepsilon\alpha \approx \left(\underbrace{\frac{(\vec{p} - q\vec{A}/c)^2}{2m} + q\phi}_{\text{nichtrel. SG}} - \underbrace{\frac{1}{8m^3c^2}(\vec{p} - q\vec{A}/c)^4}_{\text{rel. Korrekturen}} \right)\alpha. \quad (25.71)$$

The relativistic corrections in (25.71) are of the order $\mathcal{O}[(v/c)^2]$. For a particle in the constant magnetic field $\vec{B} = \vec{\nabla} \wedge \vec{A}$, for example $\vec{A} = (\vec{B} \wedge \vec{r})/2$ where $\vec{\nabla} \cdot \vec{A} = 0$, we get

$$i\hbar\partial_t\alpha \approx \left(\frac{p^2}{2m} - \frac{p^4}{8m^3c^2} - \frac{q\vec{B} \cdot \vec{L}}{2mc} \left(1 - \frac{p^2}{2m^2c^2}\right) + q\phi \right)\alpha. \quad (25.72)$$

Since we have no spin, we only find an orbital Zeeman Effect, which also by higher order in $(v/c)^2$ is reduced ¹⁴.

25.5 Charge conjugation

The Klein-Gordon equation, as well as the Dirac equation, both manifestly covariant, have an additional symmetry, which converts the states of positive and negative energies into one another. This symmetry is called charge conjugation and contains the following statement:

Let $\Phi = \Phi^{(-)}$ be a solution of the Klein-Gordon equation to negative energy and a corresponding negative norm. Then $\Phi^c \equiv \tau_1 \Phi^*$ is a solution of the charge-inverted Klein-Gordon equation with positive norm. The charge-inverted Klein-Gordon equation is obtained by replacing the charge q with $-q$ in the equation (25.48). In particular, $\Phi_{-\vec{k}}^{(-)}$ under charge conjugation in $\Phi_{\vec{k}}^{(+)}$ so that a state of charge q , with momentum $-\vec{k}$, energy $-E < 0$, and negative norm, which propagates in time backwards, an anti-particle of charge $-q$, with impulse \vec{k} , energy $E > 0$ and positive norm, which in advancing the time corresponds to.

¹⁴Let's replace the mass m with the relativistic mass $m_r = m/\sqrt{1-v^2/c^2}$ so the development gives the correction $m_r \approx m(1 + p^2/2m^2c^2)$.

We first show that $\Phi^c = \tau_1 \Phi^*$ solves the Klein-Gordon equation with inverted charge

$$i\hbar\partial_t\Phi = H\Phi \quad (25.73)$$

one also finds the equation by conjugation ($\vec{p}^* \rightarrow -\vec{p}$)

$$-i\hbar\partial_t\Phi^* = \left(\left[\frac{(\vec{p} + q\vec{A}/c)^2}{2m} + mc^2 \right] \tau_3 + \frac{(\vec{p} + q\vec{A}/c)^2}{2m} i\tau_2 + q\phi \right) \Phi^*. \quad (25.74)$$

The multiplication with $-\tau_1$ and application of $-\tau_1\tau_2 = \tau_2\tau_1$ und $-\tau_1\tau_3 = \tau_3\tau_1$ results (with $\tau_1\Phi^* = \Phi^c$)

$$i\hbar\partial_t\Phi^c = \left(\left[\frac{(\vec{p} + q\vec{A}/c)^2}{2m} + mc^2 \right] \tau_3 + \frac{(\vec{p} + q\vec{A}/c)^2}{2m} i\tau_2 - q\phi \right) \Phi^c, \quad (25.75)$$

and thus with Φ is a solution of

$$i\hbar\partial_t\Phi = H(q)\Phi \quad (25.76)$$

also $\Phi^c = \tau_1\Phi^*$ a solution from

$$i\hbar\partial_t\Phi^c = H(-q)\Phi^c. \quad (25.77)$$

Φ^c has an inverted norm; e.g. for $\Phi = \Phi^{(-)}$ and use of $-\tau_3 = \tau_1\tau_3\tau_1$

$$\begin{aligned} 0 &< - \int d^3r \, \overline{\Phi}^{(-)} \Phi^{(-)} = - \int d^3r \, \underbrace{\Phi^{(-)cT} \tau_1}_{\overline{\Phi}^{(-)}} \underbrace{\tau_3 \tau_1 \Phi^{(-)c*}}_{\Phi^{(-)}} \\ &= - \int d^3r \, \Phi^{(-)cT} (-\tau_3) \Phi^{(-)c*} = \int d^3r \, \overline{\Phi}^{(-)c} \Phi^{(-)c}. \end{aligned} \quad (25.78)$$

$\Phi_{-\vec{k}}^{(-)}$ goes in under charge conjugation $\Phi_{\vec{k}}^{(+)}$ about,

$$\begin{aligned} \Phi_{-\vec{k}}^{(-)c} &= \tau_1 \Phi_{-\vec{k}}^{(-)*} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} N \begin{pmatrix} -p^2 c^2 \\ (E+mc^2)^2 \\ 1 \end{pmatrix} e^{i(\vec{p}\cdot\vec{r}-Et)/\hbar} \\ &= N \begin{pmatrix} 1 \\ -p^2 c^2 \\ (E+mc^2)^2 \end{pmatrix} e^{i(\vec{p}\cdot\vec{r}-Et)/\hbar} = \Phi_{\vec{k}}^{(+)}. \end{aligned} \quad (25.79)$$

Instead of a theory with states $\Phi_{\vec{k}}^{(-)}$ and $\Phi_{\vec{k}}^{(+)}$ which solve $i\hbar\partial_t\Phi = H(q)\Phi$, we can formulate a theory of particles and anti-particles, in which the waves function

$$\begin{aligned}\Phi_{\vec{k}}^{(t)} &= \Phi_{\vec{k}}^{(+)} \quad \text{für Teilchen, und} \\ \Phi_{\vec{k}}^{(a)} &= \Phi_{-\vec{k}}^{(-)c} \quad \text{für Anti-Teilchen}\end{aligned}\tag{25.80}$$

stand and the equations

$$\begin{aligned}i\hbar\partial_t\Phi_{\vec{k}}^{(t)} &= H(q)\Phi_{\vec{k}}^{(t)}, \\ i\hbar\partial_t\Phi_{\vec{k}}^{(a)} &= H(-q)\Phi_{\vec{k}}^{(a)}\end{aligned}\tag{25.81}$$

A state $\Phi_{\vec{k}}^{(+)}$ describes either a particle with charge $q > 0$, momentum $\vec{k} > 0$ and energy $E > 0$, or an anti-particle with charge $-q$, momentum \vec{k} and energy $E > 0$, both advancing in time. With this we have again formulated a conventional quantum mechanics with exclusively positive energies and norms, but have been forced to introduce two types of particles. A consistent description of this system is only possible in the second quantized form, the quantum field theory ,

25.6 Coulomb scattering and anti-particles

In this section we examine the Coulomb scattering problem, the scatter at the potential $eV(r) = e^2/r$. We calculate the time evolution $\langle f|\Phi(t)\rangle$ in 2^{ter} order perturbation theory with the following goals:

- To complete an exercise in Klein-Gordon dynamics.
- We show that for $|i\rangle = |\Phi_{\vec{k}_i}^{(+)}\rangle$, and $|f\rangle = |\Phi_{\vec{k}_f}^{(+)}\rangle$ in 2nd Order always virtual states $\Phi_{\vec{k}}^{(-)}$ occur and are therefore relevant for the correct calculation of the cross section.
- We will interpret the appearance of these states with negative norm through virtual pair production. We will interpret backward propagating states as forward anti-particles and construct a simple Feynman graph.

In the two-component representation of the Klein-Gordon theory, we can directly apply the results of the perturbation theory from QM I. The amplitude for the process $|i\rangle = |\Phi_{\vec{k}_i}^{(+)}\rangle \rightarrow |f\rangle = |\Phi_{\vec{k}_f}^{(+)}\rangle$ is in 2nd Order perturbation theory given by

$$\begin{aligned} \langle f|\Phi(t)\rangle &= \frac{1}{i\hbar} \int_{t_0}^t dt_1 e^{-iE_f(t-t_1)/\hbar} \langle f|H'(t_1)|i\rangle e^{-iE_i(t_1-t_0)/\hbar} \\ &\quad - \frac{1}{\hbar^2} \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \sum_v \left[e^{-iE_f(t-t_1)/\hbar} \langle f|H'(t_1)|v\rangle \right. \\ &\quad \left. e^{-iE_v(t_1-t_2)/\hbar} \langle v|H'(t_2)|i\rangle e^{-iE_i(t_2-t_0)/\hbar} \right], \end{aligned} \quad (25.82)$$

with $H'(t) = e^2/r$ independent of time and $\sum_v |v\rangle\langle v| = \mathbb{1}$. The completeness of the base $\Phi_{\vec{k}}^{(+)}, \Phi_{\vec{k}}^{(-)}$ States that

$$\sum_{\vec{k}} \left\{ |\Phi_{\vec{k}}^{(+)}\rangle \langle \Phi_{\vec{k}}^{(+)}| - |\Phi_{\vec{k}}^{(-)}\rangle \langle \Phi_{\vec{k}}^{(-)}| \right\} = \mathbb{1} \quad (25.83)$$

or in location representation ¹⁵

$$\sum_{\vec{k}} \left\{ \Phi_{\vec{k}}^{(+)}(\vec{r}) \overline{\Phi_{\vec{k}}^{(+)}(\vec{r}')} - \Phi_{\vec{k}}^{(-)}(\vec{r}) \overline{\Phi_{\vec{k}}^{(-)}(\vec{r}')} \right\} = \mathbb{1} \delta(\vec{r} - \vec{r}'). \quad (25.84)$$

We need the matrix elements $\langle a|H'(t)|b\rangle$,

$$\langle \vec{k}_f^+ | H' | \vec{k}^+ \rangle = f_{\vec{k}_f, \vec{k}}^+ = \int d^3r \left(\Phi_{\vec{k}_f}^{(+)}(\vec{r}) \right)^\dagger \tau_3 \frac{e^2}{r} \Phi_{\vec{k}}^{(+)}(\vec{r})$$

¹⁵To calculate, use the expressions for the wave functions,

$$\begin{aligned} &\sum_{\vec{k}} \left\{ \Phi_{\vec{k}}^{(+)}(\vec{r}) \overline{\Phi_{\vec{k}}^{(+)}(\vec{r}')} - \Phi_{\vec{k}}^{(-)}(\vec{r}) \overline{\Phi_{\vec{k}}^{(-)}(\vec{r}')} \right\} \\ &= \sum_{\vec{k}} \exp[-\vec{k} \cdot (\vec{r} - \vec{r}')] \frac{(E + mc^2)^2}{4mc^2 E V} \\ &\quad \times \left\{ \left(\frac{1}{\frac{-p^2 c^2}{(E + mc^2)^2}} \right) \left(1, \frac{p^2 c^2}{(E + mc^2)^2} \right) - \left(\frac{\frac{-p^2 c^2}{(E + mc^2)^2}}{1} \right) \left(\frac{-p^2 c^2}{(E + mc^2)^2}, -1 \right) \right\} \end{aligned}$$

The expression in the curly bracket is to be understood as the difference between two direct products and results in the difference between two 2×2 matrices

$$\begin{pmatrix} 1 & \frac{p^2 c^2}{(\dots)^2} \\ \frac{-p^2 c^2}{(\dots)^2} & \frac{-p^4 c^4}{(\dots)^4} \end{pmatrix} - \begin{pmatrix} \frac{p^4 c^4}{(\dots)^4} & \frac{p^2 c^2}{(\dots)^2} \\ \frac{-p^2 c^2}{(\dots)^2} & -1 \end{pmatrix} = \frac{4mc^2 E}{(E + mc^2)^2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

$$\begin{aligned}
&= \int \frac{d^3r}{V} \frac{(E_f + mc^2)(E + mc^2)}{4mc^2 \sqrt{E_f E}} \frac{e^2}{r} e^{i(\vec{k} - \vec{k}_f) \cdot \vec{r}} \\
&\quad \times \left(1, \frac{-p_f^2 c^2}{(E_f + mc^2)^2} \right) \left(\frac{1}{(E + mc^2)^2} \right) \\
&= \frac{1}{V} \frac{4\pi e^2}{|\vec{k}_f - \vec{k}|^2} \frac{E_f + E}{2\sqrt{E_f E}}, \tag{25.85}
\end{aligned}$$

where we used is that $\int d^3r \exp(-\mu r) \exp(i\vec{q} \cdot \vec{r})/r = 4\pi/(q^2 + \mu^2)$ and $(E_f + mc^2)^2(E + mc^2)^2 - p_f^2 p^2 c^4 = 2mc^2(E_f + mc^2)(E + mc^2)(E_f - E)$. You can also find that

$$\langle \vec{k}_f^+ | H' | \vec{k}^- \rangle = f_{\vec{k}_f, \vec{k}}^- = \frac{1}{V} \frac{4\pi e^2}{|\vec{k}_f - \vec{k}|^2} \frac{E_f - E}{2\sqrt{E_f E}}, \tag{25.86}$$

is, using that

$$p_f^2(E + mc^2)^2 - p^2 c^2(E_f + mc^2)^2 = 2mc^2(E_f + mc^2)(E + mc^2)(E_f - E).$$

Note that is $\delta(E_f - E) \langle \vec{k}_f^+ | H' | \vec{k}^- \rangle = 0$.

1. Order: In the first order we get

$$\langle f | \Phi(t) \rangle = \frac{1}{i\hbar} e^{-iE_f t/\hbar} \int_{t_0}^t dt_1 e^{i(E_f - E_i)t_1/\hbar} f_{\vec{k}_f, \vec{k}_i}^+ \tag{25.87}$$

and with $\partial_t |\int dt_1 \exp(iEt_1/\hbar)|^2 = 2\pi\hbar\delta(E)$ we get the propagator

$$P_{i \rightarrow f} = \partial_t |\langle f | \Phi \rangle|^2 = \frac{2\pi}{\hbar} \delta(E_f - E_i) |f_{\vec{k}_f, \vec{k}_i}^+|^2. \tag{25.88}$$

The scattering cross section is ($j = v/V$ is the incident particle current density)

$$\begin{aligned}
d\sigma_{fi} &= \frac{V}{v} \int_{\Omega_{\vec{k}}} \underbrace{\frac{d\Omega}{4\pi} \frac{k_f^2 dk_f}{(2\pi)^2} L^3}_{\text{aus } \Sigma_{k_f \rightarrow f}} P_{i \rightarrow f} \\
&\downarrow v = \frac{pc^2}{E}, \quad \frac{dE}{dp} = \frac{pc^2}{E}, \quad k_f^2 dk_f = \frac{pE_f}{c^2 \hbar^3} dE_f, \\
&= d\Omega \frac{4e^4 E^2}{c^4 \hbar^4 q^4} \Big|_{q=|\vec{k}_f - \vec{k}_i|}, \\
\frac{d\sigma}{d\Omega} &= \frac{4e^4 E^2}{c^4 \hbar^4 |\vec{k}_f - \vec{k}_i|^4}. \tag{25.89}
\end{aligned}$$

2nd Order: In the second order we have to calculate the following amplitude:

$$\begin{aligned} \langle f | \Phi(t) \rangle &\stackrel{(2)}{\approx} \frac{-1}{\hbar^2} \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \underbrace{\Theta(t_1 - t_2)}_{t_2 < t_1} \sum_k e^{-iE_f(t-t_1)/\hbar} \\ &\quad \times \left\{ f_{\vec{k}_f, \vec{k}}^+ f_{\vec{k}_i, \vec{k}}^{+*} e^{-iE_k(t_1-t_2)/\hbar} - f_{\vec{k}_f, \vec{k}}^- f_{\vec{k}_i, \vec{k}}^{-*} e^{iE_k(t_1-t_2)/\hbar} \right\} e^{iE_i(t_2-t_0)/\hbar}. \end{aligned}$$

Note: The $(-)$ states occur unavoidably and they propagate in time with the phase factor $\exp(iE_k t/\hbar)$, with $t = t_1 - t_2 \geq 0$. We let $t_0 \rightarrow -\infty$ and $t \rightarrow \infty$ go and ignore the relevant irrelevant phases. With $T = \frac{1}{2}(t_1 + t_2)$ and $t = t_1 - t_2$ we get

$$\begin{aligned} \langle f | \Phi(t) \rangle &\stackrel{(2)}{\approx} \frac{-1}{\hbar^2} \int_{-\infty}^{\infty} dT \int_{-\infty}^{\infty} dt \Theta(t) \sum_{\vec{k}} \left\{ f_{\vec{k}_f, \vec{k}}^+ f_{\vec{k}_i, \vec{k}}^{+*} e^{-i(E_k - (E_i + E_f)/2)t/\hbar} \right. \\ &\quad \left. - f_{\vec{k}_f, \vec{k}}^- f_{\vec{k}_i, \vec{k}}^{-*} e^{i(E_k + (E_i + E_f)/2)t/\hbar} \right\} e^{i(E_f - E_i)T/\hbar} \\ &= -\frac{2\pi}{\hbar^2} \delta(E_i - E_f) \int_{-\infty}^{\infty} dt \Theta(t) \sum_k \left\{ f_{\vec{k}_f, \vec{k}}^+ f_{\vec{k}_i, \vec{k}}^{+*} e^{-i(E_k - E_i)t/\hbar} \right. \\ &\quad \left. - f_{\vec{k}_f, \vec{k}}^- f_{\vec{k}_i, \vec{k}}^{-*} e^{i(E_k + E_i)t/\hbar} \right\} \end{aligned} \quad (25.90)$$

We can convert the forward propagation in the second term. Convert a trick into a backward arts propagation where the time variable $t < 0$ grows in the negative direction: the expression

$$\int_{-\infty}^{\infty} dt \sum_{\vec{k}} f_{\vec{k}_f, \vec{k}}^- f_{\vec{k}_i, \vec{k}}^{-*} e^{i(E_k + E_i)t/\hbar} \quad (25.91)$$

disappears because $\int dt$ creates the δ function $\delta(E_k + E_i)$ which always disappears because $E_k, E_i > 0$. With this we can rewrite the second term from (25.90)

$$\begin{aligned} \langle f | \Phi(t) \rangle &\stackrel{(2)}{\approx} -\frac{2\pi}{\hbar^2} \delta(E_i - E_f) \int_{-\infty}^{\infty} dt \sum_k \left\{ f_{\vec{k}_f, \vec{k}}^+ f_{\vec{k}_i, \vec{k}}^{+*} \Theta(t) e^{-i(E_k - E_i)t/\hbar} \right. \\ &\quad \left. + f_{\vec{k}_f, \vec{k}}^- f_{\vec{k}_i, \vec{k}}^{-*} \Theta(-t) e^{i(E_k + E_i)t/\hbar} \right\}. \end{aligned} \quad (25.92)$$

This result allows a consistent interpretation based on pair generation and forward propagation of anti-particles: The first term in (25.92) has the time

evolution (with $t = t_1 - t_2 > 0$ and $E_i t = E_f t_1 - E_i t_2$)

$$e^{-iE_f(t-t_1)/\hbar} e^{-iE_k(t_1-t_2)/\hbar} e^{-iE_i(t_2-t_0)/\hbar}, \quad (25.93)$$

which are represented by the following graph, see Fig. 25.6

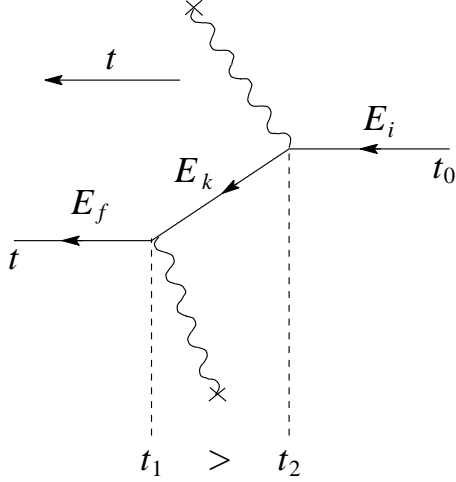


Fig. 25.5: Graph for the first term in the expression in (25.92) with transitions $i \rightarrow k \rightarrow f$ at times $t_2 < t_1$.

The second term in (25.92) can be written as (with $t = t_1 - t_2 < 0$) as

$$e^{-iE_f(t-t_1)/\hbar} e^{iE_k(t_1-t_2)/\hbar} e^{-iE_i(t_2-t_0)/\hbar}, \quad (25.94)$$

which corresponds to the graph in Fig. 25.6.

If we interpret the backward propagated state as an anti particle, which propagates in the early days, we get the graph in Fig. 25.6.

With this interpretation, the charge is preserved at $t = t_1$, whereas the energy for this virtual process is not preserved (according to Heisenberg, the energy theorem can be canceled for a correspondingly short period of time) The question of conservation of momentum is irrelevant since the translation invariance is broken by H' . Finally we simplify the result (25.92) by executing the time integration. With

$$\frac{1}{2\pi i} \int_{-\infty}^{\infty} d\omega \frac{\omega^n \exp(-i\omega t)}{E - \omega - i\varepsilon} \stackrel{\text{Cauchy}}{=} \Theta(t) E^n e^{-iEt} \quad (25.95)$$

we receive

$$\langle f | \Phi(t) \rangle \stackrel{(2)}{\approx} -\frac{2\pi}{\hbar^2} \delta(E_i - E_f) \int_{-\infty}^{\infty} dt \int \frac{d\omega}{2\pi i} e^{-i\omega t} \quad (25.96)$$

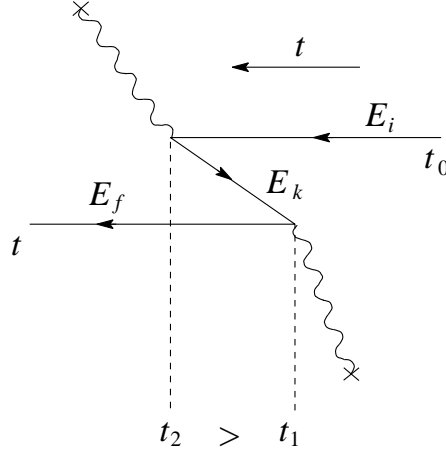


Fig. 25.6: Graph for the second term in the expression in (25.92) with transitions $i \rightarrow k \rightarrow f$ at times $t_2 > t_1$.

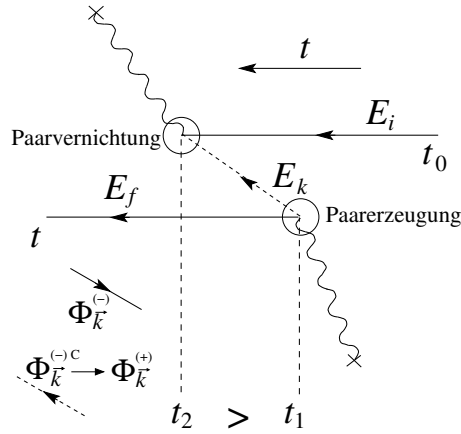


Fig. 25.7: Graph for the second term in the expression in (25.92) after replacing the backward-propagating particle as a forward-propagating anti-particle. The "transition at time t_1 is interpreted as pair generation, that at $t_2 (> t_1)$ as pair destruction.

$$\begin{aligned}
 & \sum_k \left\{ \frac{f_{\vec{k}_f, \vec{k}}^+ f_{\vec{k}_i, \vec{k}}^{+*}}{E_k - E_i - \omega - i\varepsilon} + \frac{f_{\vec{k}_f, \vec{k}}^- f_{\vec{k}_i, \vec{k}}^{-*}}{E_k + E_i + \omega - i\varepsilon} \right\} \\
 &= \frac{2\pi i}{\hbar^2} \delta(E_i - E_f) \sum_k \left\{ \frac{f_{\vec{k}_f, \vec{k}}^+ f_{\vec{k}_i, \vec{k}}^{+*}}{E_k - E_i - i\varepsilon} + \frac{f_{\vec{k}_f, \vec{k}}^- f_{\vec{k}_i, \vec{k}}^{-*}}{E_k + E_i - i\varepsilon} \right\}.
 \end{aligned}$$

Remarks:

- The energy denominator of the second term is just the energy of the virtual state $-E_i$ ("equivalently this also applies trivially for the first

term)

$$E_k + E_i = \underbrace{E_i + E_k + E_f - E_i}_{E_i + \text{Particle-antiparticle pair}} = E_k + E_f \text{ und } E_i = E_f.$$

– $t \rightarrow -t$ in the second term of (25.92)

$$\Theta(-t) e^{iE_k t/\hbar} \rightarrow \Theta(t) e^{-iE_k t/\hbar}, \quad (25.97)$$

and you find the usual forward propagation of the anti-particle.

– Pair generation: One particle becomes three particles in a one-particle theory. For a consistent description with interactions one needs the stronger formalism of the quantum field theory.

Chapter 26

Relativistic Particles: Dirac Equation*

The simplest relativistic equation we have constructed so far, the Klein-Gordon equation, has the following properties: it is

- scalar,
- covariant (in the scalar form),
- 2nd order in time (in the scalar form),
- has a *not positive definite* preserved norm and a Hamiltonian H (in the 2-component representation (25.49)) which is not Hermitian (but KG-Hermitian, (25.59)).

In the following we raise our ambitions and try to construct a theory based on a Hamiltonian with the following properties: the Hamiltonian

- may be tensorial,¹
- should be covariant,
- should be 1st order in time,
- should lead to a positively definite norm so that we get a hermitian Hamiltonian H .

¹The limitation to a scalar theory produces no result.

26.1 Dirac equation

The most general ansatz for a Hamiltonian with these properties (linear in ∂_t and covariant implies linearity in space, i.e. in $\vec{\nabla}$) has the form

$$i\hbar\partial_t\Psi = H\Psi = (c\vec{\alpha}\cdot\vec{p} + \beta mc^2)\Psi, \quad (26.1)$$

with Ψ a (Dirac) Spinor with n components, $\vec{\alpha}$, β are $n \times n$ matrices, hermitian und dimensionless; the momentum is defined as usual, $\vec{p} = -i\hbar\vec{\nabla}$. Our task is to find n , $\vec{\alpha}$ and β in such a way, that (26.1) reproduces the usual relativistic dispersion

$$E^2 = p^2c^2 + m^2c^4 \quad (26.2)$$

To do this, we square the equation according to

$$\begin{aligned} E^2\Psi \rightarrow -\hbar^2\partial_t^2\Psi &= i\hbar\partial_t(c\vec{\alpha}\cdot\vec{p} + \beta mc^2)\Psi \\ &= (c\vec{\alpha}\cdot\vec{p} + \beta mc^2)i\hbar\partial_t\Psi \\ &= (c\vec{\alpha}\cdot\vec{p} + \beta mc^2)^2\Psi \\ &\stackrel{!}{=} (p^2c^2 + m^2c^4)\Psi. \end{aligned} \quad (26.3)$$

Therefore we must have

$$\begin{aligned} (c\alpha_i p^i + \beta mc^2)(c\alpha_j p^j + \beta mc^2) &= \underbrace{c^2\alpha_i^2(p^i)^2}_{=c^2p^2} \\ &+ \underbrace{\beta^2 m^2 c^4}_{=m^2c^4} + \underbrace{(\alpha_i\beta + \beta\alpha_i)p^i mc^3}_{=0} + \underbrace{\frac{1}{2}(\alpha_i\alpha_j + \alpha_j\alpha_i)_{i\neq j} p^i p^j c^2}_{=0}. \end{aligned} \quad (26.4)$$

So we get the conditions

$$\alpha_i^2 = \mathbb{1}, \quad \beta^2 = \mathbb{1}, \quad (26.5)$$

$$\{\alpha_i, \beta\} = 0, \quad \{\alpha_i, \alpha_j\}_{i\neq j} = 0, \quad (26.6)$$

where $\{\cdot, \cdot\}$ designates the anti-commutator as usual. So we need four hermitian, anti-commuting matrices whose squares result in $\mathbb{1}$. It turns out that these conditions can only be fulfilled if $n \geq 4$. As proof, note that

1. The anti-comuutator $\{\alpha_i, \beta\} = 0$ implies, that $\beta\alpha_i\beta = -\alpha_i$ and therefore $\text{Tr}\alpha_i = -\text{Tr}[\beta\alpha_i\beta] = -\text{Tr}[\beta^2\alpha_i] = -\text{Tr}\alpha_i$ and thus $\text{Tr}\alpha_i = 0$. Likewise from $\alpha_i\beta\alpha_i = -\beta$, it follows that $\text{Tr}\beta = 0$, so all four matrices α_i, β are traceless.

2. The four matrices α_i and β have eigenvalues ± 1 , because α_i, β should be hermitian, i.e. also diagonalizable; since $\alpha_i^2 = \beta^2 = \mathbb{1}$ the squares of the eigenvalues must be equal to 1.
3. n must be even, $n = 2m$, with $m \in \mathbb{N}$, because with $\text{Tr}\alpha_i = 0$, $\text{Tr}\beta = 0$ the eigenvalues 1 and -1 must occur in pairs.
4. Finally it must be $n > 2$, because a hermitian matrix has a real main diagonal and complex conjugate elements symmetric to the main diagonal, so we find the following number of degrees of freedom

$$\underbrace{n}_{\text{realHD}} + \underbrace{2}_{\text{complex}} \cdot \underbrace{(n^2 - n)}_{\text{non-diag.}} \underbrace{/ 2}_{\text{symm.}} = n^2. \quad (26.7)$$

In addition, $\text{Tr} = 0$ and $n^2 - 1$ degrees of freedom remain. For $n = 2$ we can find three independent traceless hermitian matrices, e.g., the Pauli matrices σ_i ,

$$\left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right] = [\sigma_1, \sigma_2, \sigma_3] = \vec{\sigma}. \quad (26.8)$$

with $\sigma_i^2 = \mathbb{1}$ and the exchange rules $\sigma_i \sigma_j + \sigma_j \sigma_i = 0$.

For $n = 4$ we find 15 independent matrices with trace $\text{Tr} = 0$; from this we choose the matrices,

$$\begin{aligned} \beta &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \\ \alpha_1 &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{pmatrix}, \\ \alpha_2 &= \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \sigma_2 \\ \sigma_2 & 0 \end{pmatrix}, \\ \alpha_3 &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \sigma_3 \\ \sigma_3 & 0 \end{pmatrix}, \end{aligned} \quad (26.9)$$

with the required properties $\beta^2 = \alpha_i^2 = \mathbb{1}$ and $\{\alpha_i, \alpha_j\} = 0$, $\{\alpha_i, \beta\} = 0$.

This gives us the Dirac equation in the form

$$i\hbar\partial_t\Psi = (-i\hbar c\alpha_i\partial_{x_i} + \beta mc^2)\Psi, \quad (26.10)$$

with the Dirac 4th Spinor

$$\Psi = \begin{pmatrix} \Psi_1 \\ \vdots \\ \Psi_4 \end{pmatrix} = \begin{pmatrix} \chi \\ \eta \end{pmatrix}, \quad (26.11)$$

and the 2nd Spinors χ and η .

We get the covariant form of Dirac-equation by multiplying (26.10) with β and defining the Dirac matrices

$$\begin{aligned} \gamma^\mu &= (\beta, \beta\alpha^i), \\ \gamma^0 &= \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \\ \gamma^i &= \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}; \end{aligned} \quad (26.12)$$

the Dirac matrices are characterized by the following anti-commutators ²,

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}, \quad (26.13)$$

with $g^{\mu\nu}$ the metric tensor ($= \text{diag}\{1, -1, -1, -1\}$, cf. (26.83)). (26.10) then becomes (we again set $\hbar = c = 1$ for covariant formulas; it is $\partial_\mu = \partial_{x^\mu}$)

$$(i\gamma^\mu\partial_\mu - m)\Psi(x) = 0. \quad (26.14)$$

We use the Einstein sum convention and introduce the notation of the ‘slash’ operator, $\not{a} = \gamma^\mu a_\mu$; this simplifies (26.14) again to

$$(i\not{\partial}_x - m)\Psi(x) = 0. \quad (26.15)$$

(26.15) is covariant when Ψ transforms under Lorentz transformations in such a way that the new system has the identical equation. This condition defines the behavior of $\Psi \in \mathcal{H}$ (the Hilbert space of the states) under

²A set γ^ν of Dirac matrices is consistent if (26.10) is fulfilled.

Lorentz transformations and thus defines a 4-dimensional representation of the Lorentz group in \mathcal{H} : For $\Lambda \in \mathcal{L}$ a Lorentz transformation there is a representation matrix $S(\Lambda)$, such that

$$\Psi'(x') = S(\Lambda)\Psi(x), \quad (26.16)$$

$$x' = \Lambda x, \quad (26.17)$$

the wave function is in the new (with prime) coordinate system.³ The covariance then requires that

$$(i\gamma^\mu \partial'_\mu - m) \Psi'(x') = 0 \quad (26.18)$$

is satisfied. (26.18) determines the representation $S(\Lambda)$ of the Lorentz group: with

$$\begin{aligned} x'^\mu &= \Lambda^\mu{}_\nu x^\nu \\ \partial'_\mu &= (\Lambda^{-1})^\nu{}_\mu \partial_\nu \end{aligned} \quad (26.19)$$

and

$$S^{-1}(\Lambda) [i\gamma^\mu (\Lambda^{-1})^\nu{}_\mu \partial_\nu - m] S(\Lambda) \Psi(x) \stackrel{!}{=} (i\gamma^\mu \partial_\mu - m) \Psi(x) = 0$$

$S(\Lambda)$ must meet the following conditions,

$$\begin{aligned} (\Lambda^{-1})^\nu{}_\mu S^{-1}(\Lambda) \gamma^\mu S(\Lambda) &= \gamma^\nu, \\ \Rightarrow S^{-1}(\Lambda) \gamma^\mu S(\Lambda) &= \Lambda^\mu{}_\nu \gamma^\nu. \end{aligned} \quad (26.20)$$

With help of (26.20) we can determine the representation $S(\Lambda)$, see later.

Finally, we want to take into account the electromagnetic interaction, which we achieve through the substitution,

$$-i\hbar \partial^\mu = p^\mu \rightarrow p^\mu - qA^\mu, \quad (26.21)$$

$$i\hbar \partial_t \Psi = H\Psi = [c\alpha_i (-i\hbar \partial_{x_i} - qA_i/c) + q\phi + \beta mc^2] \Psi, \quad (26.22)$$

or in covariant notation with $\hbar = c = 1$,

$$[\gamma^\mu (i\partial_\mu - qA_\mu) - m] \Psi(x) = 0. \quad (26.23)$$

³Compare with (??): We write $U_{\vec{\omega}} \vec{\Psi}(\vec{r}) = \mathbf{R}_{\vec{\omega}} \vec{\Psi}(\mathbf{R}_{-\vec{\omega}} \vec{r})$ with $\vec{\Psi}' = U_{\vec{\omega}} \vec{\Psi}$ and $\vec{r}' = \mathbf{R}_{-\vec{\omega}} \vec{r}'$ and $\vec{\Psi}'(\vec{r}') = \mathbf{R}_{\vec{\omega}} \vec{\Psi}(\vec{r})$, which gives us the same form as (26.16).

26.1.1 Preserved norm

Let Ψ_a and Ψ_b be solutions of the Dirac equation, which we multiply by the corresponding spinor as follows,

$$\begin{aligned}\Psi_a^\dagger \cdot [i\hbar\partial_t\Psi_b &= H\Psi_b], \\ [-i\hbar\partial_t\Psi_a^\dagger &= (H\Psi_a)^\dagger] \cdot \Psi_b.\end{aligned}\quad (26.24)$$

The subtraction of the two equations gives

$$\begin{aligned}i\hbar\partial_t(\Psi_a^\dagger\Psi_b) &= \Psi_a^\dagger H\Psi_b - (H\Psi_a)^\dagger\Psi_b \\ &= \frac{\hbar}{i}\alpha_i\Psi_a^\dagger(\vec{\partial}_{x_i}^{\rightarrow} + \vec{\partial}_{x_i}^{\leftarrow})\Psi_b \\ &= \frac{\hbar}{i}\partial_{x_i}(\Psi_a^\dagger\alpha_i\Psi_b)\end{aligned}\quad (26.25)$$

and we get

$$\partial_t(\Psi_a^\dagger\Psi_b) + \partial_{x_i}(\Psi_a^\dagger\alpha_i\Psi_b) = 0, \quad (26.26)$$

or, after inserting ones, taking into account $\mathbb{1} = \beta^2 = (\gamma^0)^2$ and $\alpha_i = \beta\beta\alpha_i = \gamma^0\gamma^i$,

$$\partial_\mu(\bar{\Psi}_a\gamma^\mu\Psi_b) = 0. \quad (26.27)$$

We have the Dirac adjoint defined as

$$\bar{\Psi}_a \equiv \Psi_a^\dagger\gamma^0 \quad (26.28)$$

The integration over the spatial coordinates that takes into account the boundary conditions ($\int d^3r \vec{\nabla} \cdot \vec{j} = \int d\sigma \vec{j} \cdot \vec{n} = 0$) gives us the preserved size

$$\begin{aligned}\text{const.} &= \int d^3r \bar{\Psi}_a\gamma^0\Psi_b \\ &= \int d^3r \Psi_a^\dagger\Psi_b = \langle\Psi_a, \Psi_b\rangle.\end{aligned}\quad (26.29)$$

(26.29) defines a norm, $\|\Psi\|^2 = \langle\Psi, \Psi\rangle \geq 0$.

Notice: The size $\Psi^\dagger\Psi = \bar{\Psi}\gamma^0\Psi$ is not a scalar under Lorentz transformation, but is the 0th component of 4^{er} vector

$$\rho^\mu = \bar{\Psi} \gamma^\mu \Psi. \quad (26.30)$$

The density ρ^μ transformed like a 4th vector:

$$\begin{aligned} \rho'^\mu(x') &= \bar{\Psi}'(x') \gamma^\mu \Psi'(x') \\ &= \bar{\Psi}(x) \underbrace{S^{-1}(\Lambda) \gamma^\mu S(\Lambda)}_{\substack{(26.20) \rightarrow \Lambda_\nu^\mu \gamma^\nu}} \Psi(x) \\ &= \Lambda_\nu^\mu \rho^\nu(x), \end{aligned} \quad (26.31)$$

where we have used that

$$\begin{aligned} \bar{\Psi}'(x') &= \Psi'^\dagger(x') \gamma^0 = \Psi^\dagger(x) S^\dagger \gamma^0 \\ &= \underbrace{\Psi^\dagger(x) \gamma^0}_{\bar{\Psi}(x)} \underbrace{\gamma^0 S^\dagger \gamma^0}_{S^{-1}} \\ &= \bar{\Psi}(x) S^{-1}. \end{aligned} \quad (26.32)$$

The final step, $\gamma^0 S^\dagger \gamma^0 = S^{-1}$, will be verified later in this chapter.

26.1.2 Free particles

We take the (block) ansatz⁴,

$$\Psi_{\vec{p}}^{(\pm)}(\vec{r}, t) = N_p e^{i(\vec{p} \cdot \vec{r} \mp E_p t)/\hbar} \begin{pmatrix} \chi \\ \eta \end{pmatrix}, \quad (26.33)$$

with χ, η two 2^{er} spinors, the momentum $\vec{p} = \hbar \vec{k}$, $\vec{k} = 2\pi \vec{n}/L$ and $E_p > 0$. Inserting it into the Dirac equation (26.10) gives

$$\pm E_p \begin{pmatrix} \chi \\ \eta \end{pmatrix} = \begin{pmatrix} mc^2 & c\vec{\sigma} \cdot \vec{p} \\ c\vec{\sigma} \cdot \vec{p} & -mc^2 \end{pmatrix} \begin{pmatrix} \chi \\ \eta \end{pmatrix}. \quad (26.34)$$

The secular determinant is

$$(mc^2 \mp E_p)(-mc^2 \mp E_p) - c^2(\vec{\sigma} \cdot \vec{p})^2, \quad (26.35)$$

⁴Despite the positive semidefinite norm $\|\cdot\| \geq 0$ we will find solutions with negative energies.

and we get solutions if the dispersion

$$E_p^2 = c^2 p^2 + m^2 c^4 \quad (26.36)$$

is fulfilled, whereby we have used that $(\vec{\sigma} \cdot \vec{p})^2 = \sigma_i^2 p_i^2 + \{\sigma_i, \sigma_j\} p_i p_j / 2 = \sigma_i^2 p_i^2 = \mathbb{1}_i p_i^2 = p^2$. We get for the eigenvectors

$$\begin{aligned} + : \eta &= \left(\frac{c\vec{\sigma} \cdot \vec{p}}{E_p + mc^2} \right) \chi, \\ - : \chi &= - \left(\frac{c\vec{\sigma} \cdot \vec{p}}{E_p + mc^2} \right) \eta, \end{aligned} \quad (26.37)$$

and with $\chi^\dagger \chi \equiv 1$, and $\eta^\dagger \eta \equiv 1$ we find for the normalization⁵

$$\begin{aligned} \langle \Psi_{\vec{p}}^{(+)} \Psi_{\vec{p}}^{(+)} \rangle &= 1 = N_p^2 V \chi^\dagger \left(1, \frac{c\vec{\sigma} \cdot \vec{p}}{E_p + mc^2} \right) \begin{pmatrix} 1 \\ \frac{c\vec{\sigma} \cdot \vec{p}}{E_p + mc^2} \end{pmatrix} \chi \\ &= N_p^2 V \left(1 + \frac{c^2 p^2}{(E_p + mc^2)^2} \right) \\ &= N_p^2 V \left(1 + \frac{E_p^2 - m^2 c^4}{(E_p + mc^2)^2} \right) \\ &= N_p^2 V \frac{2E_p}{E_p + mc^2} \end{aligned} \quad (26.38)$$

and the same result applies to $\Psi_{\vec{p}}^{(-)}$.

To simplify this complex writing style, we turn to the Dirac spinor.

26.1.3 Dirac spinor

The free solutions of the Dirac theory can be defined by the Dirac spinors

$$\begin{aligned} u(\vec{p}, s) &\equiv \sqrt{E_p + mc^2} \begin{pmatrix} 1 \\ \frac{c\vec{\sigma} \cdot \vec{p}}{E_p + mc^2} \end{pmatrix} \chi^{(s)}, \\ v(\vec{p}, s) &\equiv \sqrt{E_p + mc^2} \begin{pmatrix} \frac{c\vec{\sigma} \cdot \vec{p}}{E_p + mc^2} \\ 1 \end{pmatrix} (-i\sigma_2 \chi^{(s)}), \end{aligned} \quad (26.39)$$

⁵the result is not Lorentz-covariant, because $\|\cdot\|$ is not a scalar but the 0-component of a 4-vector.

just express it through

$$\begin{aligned}\Psi_{\vec{p},s}^{(+)}(\vec{r},t) &= \frac{1}{\sqrt{2E_p V}} u(\vec{p},s) e^{i(\vec{p}\cdot\vec{r}-E_p t)/\hbar}, \\ \Psi_{-\vec{p},-s}^{(-)}(\vec{r},t) &= \frac{1}{\sqrt{2E_p V}} v(\vec{p},s) e^{-i(\vec{p}\cdot\vec{r}-E_p t)/\hbar}.\end{aligned}\quad (26.40)$$

The 2nd Spinor $\chi^{(s)}$ describes the spin of a spin-1/2-particle,

$$\begin{aligned}\chi^{1/2} &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ \chi^{-1/2} &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}.\end{aligned}\quad (26.41)$$

Likewise, $-i\sigma_2\chi^{(s)}$ describes a spin, but is

$$\begin{aligned}-i\sigma_2\chi^{(1/2)} &= \eta^{(1/2)} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \\ -i\sigma_2\chi^{(-1/2)} &= \eta^{(-1/2)} = \begin{pmatrix} -1 \\ 0 \end{pmatrix},\end{aligned}\quad (26.42)$$

and $\eta^{(s)}$ describes a down (up) spin for $s = 1/2$ ($= -1/2$), just the opposite way to $\chi^{(s)}$,

$$\chi^{(s)} \sim \eta^{(-s)}.\quad (26.43)$$

26.1.4 Interpretation

The Dirac equation has solutions again with negative energy, see $\Psi_{-\vec{p},-s}^{(-)}$ in (26.40). In addition to the component structure (χ, η) , which we also found for the Klein-Gordon equation and interpreted as a particle-antiparticle pair, the Dirac equation has a 2^{er} spinor structure in χ and η . We interpret this spinor structure as the spin state of a spin-1/2 particle. Thus $\Psi_{\vec{p},s}^{(+)}$ describes a particle with momentum \vec{p} and spin s , while the state $\Psi_{-\vec{p},-s}^{(-)}$ with impulse $-\vec{p}$ and spin $-s$ describes an anti-particle with impulse \vec{p} and spin s . The states $\Psi_{\vec{p},s}^{(+)}$ and $\Psi_{\vec{p}',s'}^{(-)}$ are orthogonal; Combinations with $\vec{p} \neq \vec{p}'$ disappear due to $\int d^3r \dots = 0$. For $\vec{p} = \vec{p}'$ results in $\int d^3r \dots = V$ and the orthogonality follows from

$$v^\dagger(-\vec{p},s') \cdot u(\vec{p},s) = u^\dagger(\vec{p},s) \cdot v(-\vec{p},s') = 0.\quad (26.44)$$

So that is

$$\begin{aligned}
 \int d^3r \Psi_{\vec{p},-s'}^{(-)\dagger} \Psi_{\vec{p},s}^{(+)} &= \int \frac{d^3r}{2E_p V} e^{-i(\vec{p}\cdot\vec{r}+E_p t)/\hbar} v^\dagger(-\vec{p}, s') \cdot u(\vec{p}, s) e^{i(\vec{p}\cdot\vec{r}-E_p t)/\hbar} \\
 &= \frac{1}{2E_p} e^{-2iE_p t/\hbar} v^\dagger(-\vec{p}, s') \cdot u(\vec{p}, s) \\
 &= \frac{E_p + mc^2}{2E_p} e^{-2iE_p t/\hbar} \eta^{(s')\dagger} \underbrace{\left(\frac{-c\vec{\sigma} \cdot \vec{p}}{E_p + mc^2}, 1 \right)}_0 \left(\frac{1}{\frac{c\vec{\sigma} \cdot \vec{p}}{E_p + mc^2}} \right) \chi^{(s)}. \\
 &= 0.
 \end{aligned} \tag{26.45}$$

Finally, the free solutions form a complete system,⁶

$$\sum_{\vec{p},s} \left[\Psi_{\vec{p},s}^{(+)}(\vec{r}) \Psi_{\vec{p},s}^{(+)\dagger}(\vec{r}') + \Psi_{\vec{p},s}^{(-)}(\vec{r}) \Psi_{\vec{p},s}^{(-)\dagger}(\vec{r}') \right] = \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \delta^3(\vec{r} - \vec{r}'). \tag{26.46}$$

We again obtain the antiparticles by charge conjugation.

26.2 Charge conjugation

The charge conjugation

$$C : \Psi \rightarrow \Psi^C = C\beta\Psi^* \tag{26.47}$$

with

$$C = -i\alpha_2 = \begin{pmatrix} 0 & -i\sigma_2 \\ -i\sigma_2 & 0 \end{pmatrix} \tag{26.48}$$

transforms states with negative energy into those with positive energy, which obey a dynamic with reversed charge: with

$$i\hbar\partial_t\Psi = [c\alpha_i(-i\hbar\partial_{x_i} - eA_i/c) + e\phi + \beta mc^2]\Psi \tag{26.49}$$

⁶Compare with the completeness relation for the free solutions of the Klein-Gordon equation (25.84) with a minus sign.

holds by complex conjugation and multiplication by $C\beta$:

$$\begin{aligned}
-i\hbar\partial_t\Psi^C &= C\beta [c\alpha_i^* (i\hbar\partial_{x_i} - qA_i/c) + q\phi + \beta mc^2] \Psi^* \\
\downarrow C\beta\alpha_i^*\Psi^* &= \underbrace{C\beta C^{-1}}_{-\beta} \underbrace{C\alpha_i^* C^{-1}}_{-\alpha_i} \underbrace{C\beta C^{-1}}_{-\beta} \underbrace{C\beta\Psi^*}_{\Psi^C} \\
&= -\beta\alpha_i\beta\Psi^C = \alpha_i\Psi^C, \\
\downarrow C\beta^2\Psi^* &= \underbrace{C\beta C^{-1}}_{-\beta} \underbrace{C\beta\Psi^*}_{\Psi^C} = -\beta\Psi^C, \\
&= [c\alpha_i (i\hbar\partial_{x_i} - qA_i/c) + q\phi - \beta mc^2] \Psi^C, \quad (26.50)
\end{aligned}$$

and thus Ψ^C follows the dynamic

$$\begin{aligned}
i\hbar\partial_t\Psi^C &= [c\alpha_i (-i\hbar\partial_{x_i} + qA_i/c) - q\phi + \beta mc^2] \Psi^C, \\
&= H(-q)\Psi^C. \quad (26.51)
\end{aligned}$$

The charge conjugation C also maps the state $\Psi_{-\vec{p},-s}^{(-)}$ to $\Psi_{\vec{p},s}^{(+)}$ because

$$\begin{aligned}
C\beta v^*(\vec{p}, s) &= \sqrt{E_p + mc^2} \begin{pmatrix} 0 & -i\sigma_2 \\ -i\sigma_2 & 0 \end{pmatrix} \begin{pmatrix} \frac{c\vec{\sigma}^* \cdot \vec{p}}{E_p + mc^2} \\ -1 \end{pmatrix} (-i\sigma_2 \chi^{(s)}) \\
&= \sqrt{E_p + mc^2} \begin{pmatrix} 1 \\ \frac{c\vec{\sigma} \cdot \vec{p}}{E_p + mc^2} \end{pmatrix} \chi^{(s)} = u(\vec{p}, s). \quad (26.52)
\end{aligned}$$

We used $\sigma_2 \vec{\sigma}^* \sigma_2 = -\vec{\sigma}$. Analogous to (26.52) is

$$C\beta u^*(\vec{p}, s) = v(\vec{p}, s). \quad (26.53)$$

Accordingly, we find that

$$\Psi_{-\vec{p},-s}^{(-)C} = C\beta\Psi_{-\vec{p},-s}^{(-)*} = \Psi_{\vec{p},s}^{(+)}, \quad (26.54)$$

and we identify the $(+)$ energy state $\Psi_{-\vec{p},-s}^{(-)C}$, which obeys the $-q$ -dynamics as anti-particles with identical mass and spin and with impulse \vec{p} and spin component s :

$\Psi_{\vec{p},s}^{(+)}$ = Particles with $E_p > 0$, momentum \vec{p} and spin s , dynamics of charge q ,

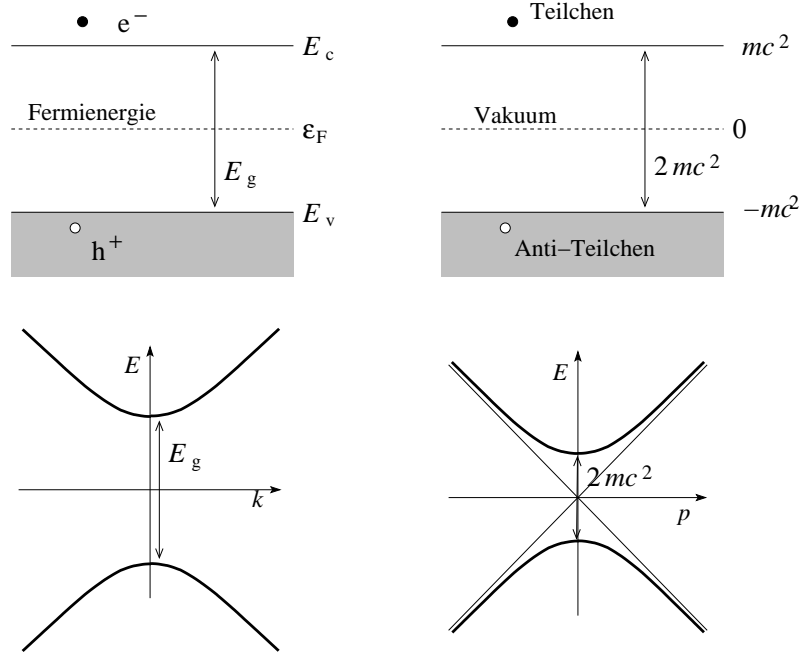


Fig. 26.1: Left: Conduction and valence band (top) and dispersion (bottom) of the electrons in a semiconductor with direct energy gaps (eg, GaAs, ZnO, ...). The excitations (missing electrons) in the valence band are interpreted as holes. Right: Hole theory of the Dirac equation. The states with negative energy are filled and block the decay of states with positive energy. The excitations with negative energy are interpreted as anti-particles.

$\Psi_{-\vec{p},-s}^{(-)}$ State with $E_p < 0$, impulse $-\vec{p}$ and spin $-s$; with associated $\Psi_{-\vec{p},-s}^{(-)C}$ interpreted as anti-particles with $E_p > 0$, Momentum \vec{p} and spin s , dynamics of charge $-q$.

Since these particles are fermions, the Pauli exclusion principle applies and we can give an interesting interpretation that was not possible for the bosonic Klein-Gordon particles:

If we interpret the physical vacuum as filled with negative energy states, the $E_p > 0$ states are stable because of the Pauli blocking. The lowest state therefore has the energy mc^2 and not $-\infty$. The state $\Psi_{-\vec{p},-s}^{(-)}$ describes the absence of the particle with the same mass and the same spin, but because

it is a hole, inverse charge. Its energy is

$$0 - (-E) = \underbrace{0}_{\text{Vacuum energy}} - \underbrace{-}_{\text{Missing particles E of particles}} \underbrace{(-E)}_{\text{particles}} = E > 0, \quad (26.55)$$

its impulse $0 - (-\vec{p}) = \vec{p}$ and its spin $0 - (-s) = s$. This hole theory of the Dirac equation is the relativistic generalization of hole theory in semiconductors, as shown in Fig. 26.1.

It is also interesting to discuss the Kleinian paradox (scattering at a potential level) for fermions (although pair generation only really makes sense in many-body physics). The following results are found: 1) for $eV > 2mc^2$ both bosons and fermions are totally reflected. 2) Due to the Pauli principle, no pair can accompany the reflected particle in fermions. 3) Reflected bosons can be accompanied by further boson pairs (stimulated emission in the channel of the reflected boson). 4) In the case of fermions, particles - anti-particle pairs can be produced spontaneously. As a counterpart to the spontaneous electron-positron ($e^- - e^+$) pair generation in a strong electric field, we mention the particle-hole pair production via tunnel effect on a semiconductor $p - n$ - transition in the breakdown area, see Fig. 26.2.

26.3 Non-relativistic limits

A final ‘Tour de Force’. We want to derive the non-relativistic limits up to the order p^4/m^3 , that is $p^2/2m \cdot (v/c)^2$, i.e. including first, relativistic corrections about the (non-trivial) for leading terms $p^2/2m \sim q\phi \ll mc^2$. We first discuss the result and then show the solution path.

The dynamics of a spin-1/2 particle in the electromagnetic field ϕ, \vec{A} with relativistic corrections up to order $(v/c)^2$ results in (for a centrally symmetric potential $\phi(\vec{r}) = \phi(r)$ we can replace $\vec{\nabla}\phi \wedge \vec{p}$ with $r^{-1}(\partial_r\phi)\vec{L}$)

$$H = \underbrace{\frac{1}{2m} \left(\vec{p} - \frac{q}{c} \vec{A} \right)^2 + q\phi}_{H_0} - \underbrace{\frac{1}{2mc^2} \left(\frac{p^2}{2m} \right)^2 + \frac{q\hbar}{4m^2c^2} \frac{(\partial_r\phi)}{r} \vec{\sigma} \cdot \vec{L} + \frac{q\hbar^2 (\Delta\phi)}{8m^2c^2}}_{\text{Finestructure}} - \underbrace{\frac{q\hbar}{2mc} \vec{\sigma} \cdot \vec{B}}_{\text{Zeeman}}. \quad (26.56)$$

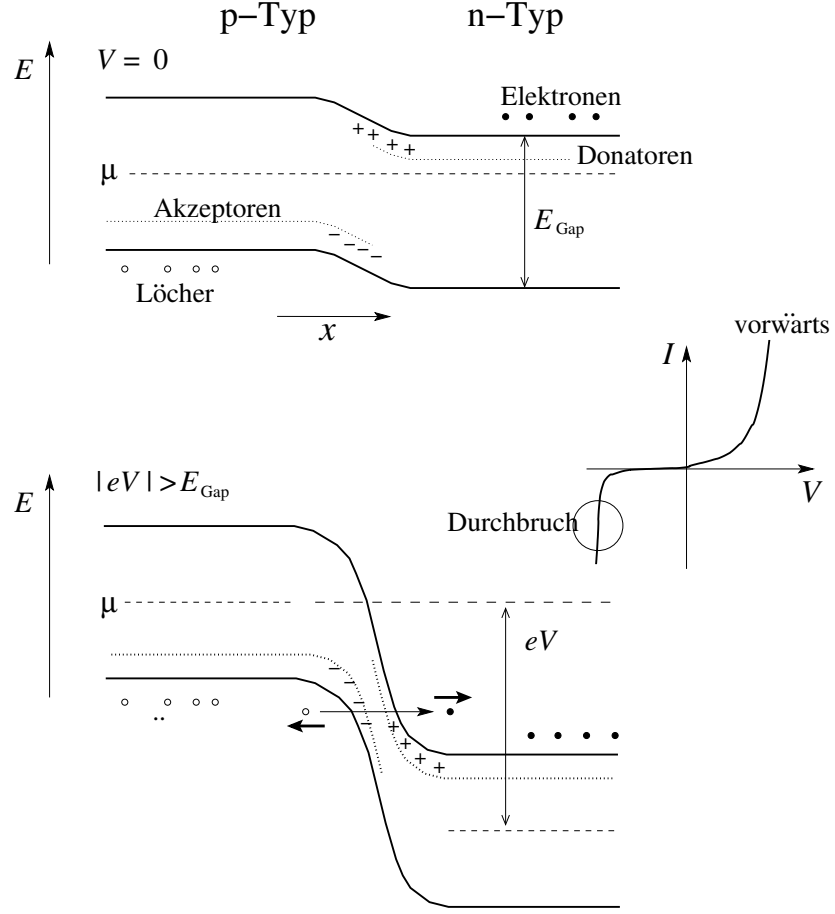


Fig. 26.2: Electron-hole pair generation in the electric field of a p - n semiconductor junction. Above: situation in equilibrium without applied voltage, $V = 0$. Below: situation in the backward polarized p - n junction in the breakdown regime where particle Hole pairs are generated via the tunnel process. See also the current-voltage characteristic of the p - n diode.

If we combine the orbital and spin Zeeman term we get

$$H_{\text{Zeeman}} = -\frac{q}{2mc}\vec{B} \cdot (\vec{L} + 2\vec{S}), \quad (26.57)$$

and from it the gyromagnetic factor of the electron spin ($q = -e$), $g = 2$. We know the different terms (spin-orbit coupling, Zeeman term) from previous discussions; the correction $\propto p^4$ to the kinetics is trivial. The Darwin term

$\propto \Delta\phi$ is new and interesting. It can be argued that this originates from the smearing of the electron over the Compton wavelength ⁷ \hbar/mc : The averaging of potential

$$\begin{aligned}\langle q\phi(\vec{r} + \delta\vec{r}) \rangle &\approx \langle q\phi(\vec{r}) + q\delta\vec{r} \cdot \vec{\nabla}\phi(\vec{r}) + \frac{q}{2}(\delta\vec{r} \cdot \vec{\nabla})^2\phi(\vec{r}) \rangle \\ &\approx q\phi(\vec{r}) + \frac{q}{6} \frac{\hbar^2}{m^2 c^2} (\Delta\phi),\end{aligned}\quad (26.58)$$

gives a term of the corresponding structure and order of magnitude. The associated quantum fluctuations in the position of the electron is called 'Zitterbewegung', see also the discussion at the end of this section. It is unfortunate that this argument does not involve the spin, because the Klein-Gordon equation does not show a corresponding term.

The application to the electron in the Coulomb potential $\phi(r) = eZ/r$, $q = -e$ ($Z = 1$ for the hydrogen atom) results in a fine structure that only the (main and magnetic) Quantum numbers n and j are involved, although the Darwin term only contributes to the orbital angular momentum $l = 0$ and the spin orbit term to $l \neq 0$. The compensation of these contributions then gives the energy ⁸

$$E_{nj} = mc^2 + T = mc^2 \left[1 - \frac{(Z\alpha)^2}{2n^2} - \frac{(Z\alpha)^4}{2n^4} \left(\frac{n}{j + 1/2} - \frac{3}{4} \right) \right] \quad (26.60)$$

with $\alpha = e^2/\hbar c \approx 1/137$ the fine structure constant.

Derivation*:

Unfortunately, the simplest way of derivation does not work: consider a solution

$$\Psi = \begin{pmatrix} \chi \\ \eta \end{pmatrix} e^{-iEt/\hbar} \quad (26.61)$$

with $E > 0$, $\chi = \chi(\vec{r})$ and $\eta = \eta(\vec{r})$. With $E = T + mc^2$ we can write

$$T\chi = q\phi\chi + c\vec{\sigma} \cdot (\vec{p} - q\vec{A}/c) \eta,$$

⁷The Compton wavelength \hbar/mc is the length scale of the relativistic differential equation.

⁸The exact energies of the relativistic electron are in the Coulomb potential

$$E_{nj} = mc^2 \left[1 + \left(\frac{Z\alpha}{n - (j + 1/2) + \sqrt{(j + 1/2)^2 - Z^2\alpha^2}} \right)^2 \right]^{-1/2}. \quad (26.59)$$

There is no stable s state for $j = 1/2$ and $Z > 137$.

$$(2mc^2 + T)\eta = q\phi\eta + c\vec{\sigma} \cdot (\vec{p} - q\vec{A}/c)\chi. \quad (26.62)$$

The second equation gives

$$\eta = \frac{1}{2mc^2 + T - q\phi} c\vec{\sigma} \cdot (\vec{p} - q\vec{A}/c)\chi \quad (26.63)$$

and insert in (26.62) lists

$$T\chi = \left[q\phi + c\vec{\sigma} \cdot (\vec{p} - q\vec{A}/c) \frac{1}{2mc^2 + T - q\phi} c\vec{\sigma} \cdot (\vec{p} - q\vec{A}/c) \right] \chi. \quad (26.64)$$

If we want to go to order p^4/m^3 , we have to develop and maintain the denominator accordingly

$$\begin{aligned} T\chi = & \left[\left(q\phi + \frac{1}{2m} \left[\vec{\sigma} \cdot (\vec{p} - q\vec{A}/c) \right]^2 - \frac{1}{4m^2c^2} \vec{\sigma} \cdot (\vec{p} - q\vec{A}/c) \right. \right. \\ & \left. \left. \times (T - q\phi) \vec{\sigma} \cdot (\vec{p} - q\vec{A}/c) \right] \right] \chi = H(\text{Energie} = E)\chi. \end{aligned} \quad (26.65)$$

We do not like the dependency of $T = E - mc^2$ on the right side because it forces us to use an energy-dependent Hamiltonian. We therefore replace T with the lowest approximation

$$T \approx q\phi + \frac{1}{2m} \left[\vec{\sigma} \cdot (\vec{p} - q\vec{A}/c) \right]^2. \quad (26.66)$$

However, this step is not clear, because the order $(\vec{p} - q\vec{A}/c)(T - q\phi)$ generates a term $\vec{p}q\phi \propto q\vec{E}$, while the expression $T - q\phi \approx [\vec{\sigma} \cdot (\vec{p} - q\vec{A}/c)]^2/2m$ exchanges with $\vec{\sigma} \cdot (\vec{p} - q\vec{A}/c)$. Alternatively, you can first commute the factor $\vec{\sigma} \cdot (\vec{p} - q\vec{A}/c)$ with $(T - q\phi)$ and then use (26.66), however this procedure is also not clear and also does not yet generate Hermitian terms.

A clear non-relativistic limit can be constructed with the Foldy-Wouthuysen transformation, a unitary transformation that makes the non-diagonal terms small. We can then solve the coupled system of equations in the transformed system. We start from the Dirac equation in the form

$$\begin{aligned} T \begin{pmatrix} \chi \\ \eta \end{pmatrix} &= \left[-mc^2 + e\phi + c\vec{\alpha} \cdot (\vec{p} - e\vec{A}/c) + mc^2\beta \right] \begin{pmatrix} \chi \\ \eta \end{pmatrix} \\ &= H \begin{pmatrix} \chi \\ \eta \end{pmatrix} \end{aligned} \quad (26.67)$$

$$\sim mc^2 \left[-1 + \left(\frac{v}{c} \right)^2 + \left(\frac{v}{c} - \left(\frac{v}{c} \right)^2 \right) + 1 \right] \begin{pmatrix} \chi \\ \eta \end{pmatrix}. \quad (26.68)$$

The last equation lists the order of the terms in the relativistic parameter v/c , whereby the term $c\vec{\alpha} \cdot \vec{p} \propto v/c$ is large. In contrast, only terms $\propto 1$, $(v/c)^2$ appeared in the Klein-Gordon equation, which is why their non-relativistic limits were not problematic. It is precisely the term $c\vec{\alpha} \cdot \vec{p} \propto (v/c)mc^2$, which gives us trouble and we will transform it away. We use a unitary transformation of the form

$$U = U^\dagger = \underbrace{\sqrt{1 - \frac{\lambda^2 p^2}{m^2 c^2}}}_{\Lambda} \beta + \frac{\lambda}{mc} \vec{\alpha} \cdot \vec{p}, \quad (26.69)$$

with the appropriate parameter λ ; we will choose λ so that the term $c\vec{\alpha} \cdot \vec{p}$ disappears in the transformed system⁹. We define the transformed spinor

$$\begin{pmatrix} \chi' \\ \eta' \end{pmatrix} = U \begin{pmatrix} \chi \\ \eta \end{pmatrix}, \quad (26.71)$$

multiply (26.67) by U ,

$$UT \begin{pmatrix} \chi \\ \eta \end{pmatrix} = T \begin{pmatrix} \chi' \\ \eta' \end{pmatrix} = U H U^{-1} \begin{pmatrix} \chi' \\ \eta' \end{pmatrix}, \quad (26.72)$$

and get the transformed Hamiltonian

$$H' = U H U^{-1}. \quad (26.73)$$

We transform the terms mc^2 , $q\phi$, $c\vec{\alpha} \cdot (\vec{p} - q\vec{A}/c)$, and $mc^2\beta$ and find

$$U [mc^2] U^{-1} = mc^2, \quad (a)$$

$$U [q\phi] U^{-1} = \Lambda q\phi \Lambda + \beta \frac{\lambda}{mc} (\Lambda q\phi \vec{\alpha} \cdot \vec{p} - \vec{\alpha} \cdot \vec{p} q\phi \Lambda) \quad (b)$$

$$+ \frac{\lambda^2}{m^2 c^2} \vec{\alpha} \cdot \vec{p} q\phi \vec{\alpha} \cdot \vec{p},$$

$$U [c\vec{\alpha} \cdot (\vec{p} - q\vec{A}/c)] U^{-1} = -\Lambda c \vec{\alpha} \cdot (\vec{p} - q\vec{A}/c) \Lambda \quad (c)$$

⁹The ansatz (26.69) actually defines a unitary transformation,

$$\begin{aligned} UU^\dagger &= \left(\Lambda \beta + \frac{\lambda}{mc} \vec{\alpha} \cdot \vec{p} \right)^2 \\ &= \Lambda^2 \beta^2 + \frac{\lambda}{mc} \Lambda \{\beta, \alpha_i\} p_i + \frac{\lambda^2}{m^2 c^2} [\{\alpha_i, \alpha_j\} p_i p_j / 2 + p^2] \\ &= \Lambda^2 + \frac{\lambda^2}{m^2 c^2} p^2 = 1. \end{aligned} \quad (26.70)$$

$$\begin{aligned}
& +\beta \frac{\lambda}{m} [\Lambda \vec{\alpha} \cdot (\vec{p} - q\vec{A}/c) \vec{\alpha} \cdot \vec{p} + \vec{\alpha} \cdot \vec{p} \vec{\alpha} \cdot (\vec{p} - q\vec{A}/c) \Lambda] \quad (26.74) \\
& + \frac{\lambda^2}{m^2 c} \vec{\alpha} \cdot \vec{p} \vec{\alpha} \cdot (\vec{p} - q\vec{A}/c) \vec{\alpha} \cdot \vec{p}, \\
U [mc^2 \beta] U^{-1} & = mc^2 \beta \Lambda^2 + \underline{2\lambda\Lambda c \vec{\alpha} \cdot \vec{p}} - \beta \lambda^2 p^2 / m. \quad (d)
\end{aligned}$$

The (underlined) non-diagonal terms of odd order in $\vec{\alpha}$ and of relative order v/c are critical. With

$$\Lambda \approx 1 - \frac{\lambda^2 p^2}{2m^2 c^2} = 1 - \mathcal{O}[(v/c)^2] \quad (26.75)$$

we get (the index 'nd' stands for 'non-diagonal')

$$\begin{aligned}
H'_{\text{nd}, v/c} & = -c\vec{\alpha} \cdot (\vec{p} - q\vec{A}/c) + 2\lambda c\vec{\alpha} \cdot \vec{p} + \mathcal{O}[(v/c)^3] \\
& \stackrel{\lambda=1/2}{=} q\vec{\alpha} \cdot \vec{A} \sim \mathcal{O}[(v/c)^2]. \quad (26.76)
\end{aligned}$$

The equation (26.72) then takes the form

$$T \begin{pmatrix} \chi' \\ \eta' \end{pmatrix} = \begin{pmatrix} H'_{11} & q\vec{\alpha} \cdot \vec{A} \\ q\vec{\alpha} \cdot \vec{A} & -2mc^2 \end{pmatrix} \begin{pmatrix} \chi' \\ \eta' \end{pmatrix}, \quad (26.77)$$

with (we refer to (26.74); the terms $\propto mc^2$ in (a) and (d) combine to $(-11 + \beta)mc^2$ and do not contribute to H'_{11} but contribute $-2mc^2$ to H'_{22})

$$\begin{aligned}
H'_{11} & \approx \underbrace{q\phi - \frac{p^2}{8m^2 c^2} q\phi - q\phi \frac{p^2}{8m^2 c^2}}_{(b): \Lambda q\phi \Lambda} + \underbrace{\frac{\vec{\sigma} \cdot \vec{p} q\phi \vec{\sigma} \cdot \vec{p}}{4m^2 c^2}}_{(b)} \quad (26.78) \\
& + \underbrace{\frac{1}{2m} [\vec{\sigma} \cdot (\vec{p} - q\vec{A}/c) \vec{\sigma} \cdot \vec{p} + \vec{\sigma} \cdot \vec{p} \vec{\sigma} \cdot (\vec{p} - q\vec{A}/c)]}_{(c)} - \frac{p^4}{8m^3 c^2} - \underbrace{\frac{p^2}{2m}}_{(d)}.
\end{aligned}$$

The rest are rewriting; with $\sigma_i \sigma_j = \delta_{ij} + i\varepsilon_{ijk} \sigma_k$ you get for the (c) term

$$\begin{aligned}
& \vec{\sigma} \cdot (\vec{p} - q\vec{A}/c) \vec{\sigma} \cdot \vec{p} + \vec{\sigma} \cdot \vec{p} \vec{\sigma} \cdot (\vec{p} - q\vec{A}/c) \quad (26.79) \\
& = 2p^2 - \vec{\sigma} \cdot \vec{p} \vec{\sigma} \cdot (q\vec{A}/c) - \vec{\sigma} \cdot (q\vec{A}/c) \vec{\sigma} \cdot \vec{p} \\
& = 2p^2 - (q/c)(\vec{p} \cdot \vec{A} + \vec{A} \cdot \vec{p}) - i(q/c) \vec{\sigma} \cdot [(\vec{p} \wedge \vec{A}) - (\vec{A} \wedge \vec{p})] \\
& = (\vec{p} - q\vec{A}/c)^2 + p^2 - (q\vec{A}/c)^2 - (q\hbar/c) \vec{\sigma} \cdot (\vec{\nabla} \wedge \vec{A}) \\
& = (\vec{p} - q\vec{A}/c)^2 + p^2 - (q/c)^2 A^2 - (q\hbar/c) \vec{\sigma} \cdot \vec{B}.
\end{aligned}$$

The (b) terms can also be simplified according to

$$\begin{aligned} p^2\phi + \phi p^2 &= (p^2\phi) + 2(\vec{p}\phi) \cdot \vec{p} + 2\phi p^2, \\ \vec{\sigma} \cdot \vec{p} \phi \vec{\sigma} \cdot \vec{p} &= \vec{\sigma} \cdot (\vec{p}\phi) \vec{\sigma} \cdot \vec{p} + \phi p^2 \\ &= (\vec{p}\phi) \cdot \vec{p} + i\vec{\sigma} \cdot [(\vec{p}\phi) \wedge \vec{p}] + \phi p^2 \end{aligned}$$

and we get

$$\begin{aligned} &-(q/8m^2c^2)\{p^2\phi + \phi p^2 - 2\vec{\sigma} \cdot \vec{p} \phi \vec{\sigma} \cdot \vec{p}\} \\ &= -(q/8m^2c^2)\{(p^2\phi) - 2i\vec{\sigma} \cdot [(\vec{p}\phi) \wedge \vec{p}]\}. \end{aligned}$$

Inserting in (26.78) gives the Hamiltonian (the diamagnetic term $\propto A^2$ is obtained from $\eta' = (q/2mc^2)\vec{\alpha} \cdot \vec{A} \chi' + \mathcal{O}(v^4/c^4)$)

$$\begin{aligned} H'_{11} + \frac{e^2}{2mc^2}A^2 &= \frac{1}{2m} \left(\vec{p} - \frac{q}{c}\vec{A} \right)^2 + q\phi - \frac{q\hbar}{2mc}\vec{\sigma} \cdot \vec{B} - \frac{p^4}{8m^3c^2} \\ &+ \frac{q\hbar^2(\Delta\phi)}{8m^2c^2} + \frac{q\hbar}{4m^2c^2}\vec{\sigma} \cdot [(\vec{\nabla}\phi) \wedge \vec{p}]. \end{aligned} \quad (26.80)$$

The dynamics of a spin-1/2 particle in the electromagnetic field ϕ, \vec{A} with relativistic corrections up to order $(v/c)^2$ finally results in (cf. (26.56)), we consider a centrally symmetric potential $\phi(r)$

$$\begin{aligned} H &= \underbrace{\frac{1}{2m} \left(\vec{p} - \frac{q}{c}\vec{A} \right)^2 + q\phi}_{H_0} - \underbrace{\frac{1}{2mc^2} \left(\frac{p^2}{2m} \right)^2 + \frac{q\hbar}{4m^2c^2} \frac{(\partial_r\phi)}{r} \vec{\sigma} \cdot \vec{L} + \frac{q\hbar^2(\Delta\phi)}{8m^2c^2}}_{\text{Feinstruktur}} \\ &\quad - \underbrace{\frac{q\hbar}{2mc} \vec{\sigma} \cdot \vec{B}}_{\text{Zeeman}}. \end{aligned} \quad (26.81)$$

The different terms are known from earlier discussions, newly added is the relativistic correction $\propto p^4$ to the kinetic energy and the Darwin term which generates the 'Zitterbewegung' of the electron. A more in-depth analysis of the Darwin term shows that this Zitterbewegung comes from the interference of positive and negative energy states: a state with purely positive (negative) energy shows no Zitterbewegung. The scatter, e.g. of an electron at a potential always adds $E > 0$ state with $E < 0$ states. Within the hole theory the Zitterbewegung is interpreted as follows: The potential adds a $E > 0$ state component with $E < 0$. The interference effects that lead to the Zitterbewegung correspond to virtual e^+e^- -pairs. The scattered e^- annihilates the e^+ and the e^- of the e^+e^- -pair takes over its role. This

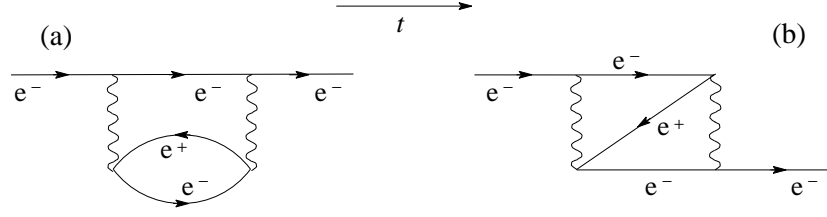


Fig. 26.3: Feynman diagrams for virtual electron-hole pair generation: (a) Vacuum polarization (direct term), similar to the diagram 20.6 of the Lamb shift, but with the photon line replaced by a shielded photon line with a polarization loop. (B) Zitterbewegung (exchange term of vacuum polarization). The electron disappears (via pair annihilation) to reappear at another location (via pair generation).

means that the electron jumps. The process as sketched in Fig. 26.3 is an exchange process.

The extent of the Zitterbewegung can be found with the following estimate: The lifespan Δt of the virtual e^+e^- -pair is given by the uncertainty relation¹⁰: index Heisenberg! Application HUP $\Delta t \sim \hbar/2mc^2$, from which the distance $\Delta r \sim c\Delta t \sim \hbar/mc$ results.

26.4 Covariance of the Dirac equation

26.4.1 Lorentz group

We define the Lorentz group \mathcal{L} as a set of 4×4 matrices Λ with the property

$$\begin{aligned}\Lambda^T G \Lambda &= G, \\ \Lambda^\mu{}_\alpha g_{\mu\nu} \Lambda^\nu{}_\beta &= g_{\alpha\beta},\end{aligned}\tag{26.82}$$

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.\tag{26.83}$$

The Lorentz transformations Λ leave the Minkovski space metric G invariant. This set, together with the operation 'o', the matrix multiplication, de-

¹⁰According to Heisenberg uncertainty principle is $\Delta E \Delta t \geq \hbar$, where ΔE is given by the rest energy mc^2 of the electron.

defines a group¹¹, The group of Lorentz transformations \mathcal{L} . Physical Lorentz transformations are real and decay into four classes listed in the table 26.1.

	det	Λ_{00}	Type
\mathcal{L}_+^\uparrow	1	> 1	proper, orthochronous
\mathcal{L}_+^\downarrow	1	< -1	proper, non-orthochronous
\mathcal{L}_-^\uparrow	-1	> 1	improper, orthochronous
\mathcal{L}_-^\downarrow	-1	< -1	improper, non-orthochronous

Table 26.1: Classes of the Lorentz Group

The two conditions are subject to this classification

det Λ :

$$\begin{aligned}\det(\Lambda^T G \Lambda) &= (\det \Lambda)^2 \det G = \det G \\ \Rightarrow (\det \Lambda)^2 &= 1,\end{aligned}\tag{26.84}$$

Λ_{00} :

$$\begin{aligned}(\Lambda^T G \Lambda)_{00} &= \Lambda_{00}^2 - \sum_j (\Lambda_{j0})^2 = G_{00} = 1, \\ \Rightarrow \Lambda_{00}^2 &= 1 + \sum_j (\Lambda_{j0})^2 \geq 1.\end{aligned}\tag{26.85}$$

Because the sizes

$$\begin{aligned}\det \Lambda &= 1 \quad \& \quad \det \Lambda = -1, \\ \Lambda_{00} &> 1 \quad \& \quad \Lambda_{00} < -1,\end{aligned}\tag{26.86}$$

cannot be continuously converted to each other, these four classes are not related. Only \mathcal{L}_+^\uparrow is continuously related to 1l and forms a restricted group. The other classes are obtained from \mathcal{L}_+^\uparrow through the transformations P = parity, T = time reversal and PT their combination (where $T, P \in \mathcal{L}$),

$$\begin{aligned}P\mathcal{L}_+^\uparrow &= \mathcal{L}_-^\uparrow, \\ T\mathcal{L}_+^\uparrow &= \mathcal{L}_-^\downarrow, \\ PT\mathcal{L}_+^\uparrow &= \mathcal{L}_+^\downarrow.\end{aligned}\tag{26.87}$$

¹¹multiplication in a group: with $\Lambda_1, \Lambda_2 \in \mathcal{L}$ is $\Lambda_1 \circ \Lambda_2 \in \mathcal{L}$, whereby the associativity is fulfilled and there exists the identity and the inverse elements.

The parity and time reversal operation can be represented by the following matrices

$$\begin{aligned} P &= \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{array} \right), \\ T &= \left(\begin{array}{c|ccc} -1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right); \end{aligned} \quad (26.88)$$

it is therefore sufficient to study \mathcal{L}_+^\uparrow .

26.4.2 The sector \mathcal{L}_+^\uparrow

The sector \mathcal{L}_+^\uparrow of \mathcal{L} can be constructed from infinitesimal generators λ ,

$$\Lambda = e^{\theta\lambda}, \quad (26.89)$$

which is characterized by the condition

$$\begin{aligned} \Lambda^T G \Lambda &\stackrel{\theta=\varepsilon \text{ klein}}{\approx} (\mathbb{1} + \varepsilon \lambda^T) G (\mathbb{1} + \varepsilon \lambda) = G \\ &\Rightarrow \lambda^T G + G \lambda = 0 \end{aligned} \quad (26.90)$$

Thus it holds for the generating matrices λ

$$\begin{pmatrix} \lambda_{00} & \lambda_{10} & \lambda_{20} & \lambda_{30} \\ \lambda_{01} & \lambda_{11} & \lambda_{21} & \lambda_{31} \\ \lambda_{02} & \lambda_{12} & \lambda_{22} & \lambda_{32} \\ \lambda_{03} & \lambda_{13} & \lambda_{23} & \lambda_{33} \end{pmatrix} = \begin{pmatrix} -\lambda_{00} & \lambda_{01} & \lambda_{02} & \lambda_{03} \\ \lambda_{10} & -\lambda_{11} & -\lambda_{12} & -\lambda_{13} \\ \lambda_{20} & -\lambda_{21} & -\lambda_{22} & -\lambda_{23} \\ \lambda_{30} & -\lambda_{31} & -\lambda_{32} & -\lambda_{33} \end{pmatrix}, \quad (26.91)$$

that means $\lambda_{\mu\mu} = 0$ and the remaining six degrees of freedom define three symmetri and three anti-symmetric generators.

3 symmetric generators

$$\omega_{10} = \omega_{01} = \left(\begin{array}{c|ccc} 0 & 1 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right),$$

$$\begin{aligned}
\omega_{20} = \omega_{02} &= \left(\begin{array}{c|ccc} 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right), \\
\omega_{30} = \omega_{03} &= \left(\begin{array}{c|ccc} 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array} \right).
\end{aligned} \tag{26.92}$$

3 anti-symmetric generators

$$\begin{aligned}
\omega_{12} = -\omega_{21} &= \left(\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right), \\
\omega_{23} = -\omega_{32} &= \left(\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{array} \right), \\
\omega_{31} = -\omega_{13} &= \left(\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array} \right).
\end{aligned} \tag{26.93}$$

We can write compactly

$$(\omega_{\mu\nu})^\alpha{}_\beta = -\frac{1}{2}\varepsilon_{\mu\nu\lambda\sigma}\varepsilon^{\lambda\sigma\alpha}{}_\beta \tag{26.94}$$

with the totally anti-symmetric tensor $\varepsilon_{\mu\nu\lambda\sigma}$ and $\varepsilon_{0123} = 1$. The most general infinitesimal Lorentz transformation then has the form (cf. (26.90)) with $\varepsilon = (v_i, \theta_j)$, $\lambda = (\omega_{jk})$

$$\Lambda = \mathbb{1} + v_i \omega_{i0} + \frac{1}{2} \theta_i \varepsilon_{ijk} \omega_{jk}, \tag{26.95}$$

and exponentiation gives finite transformations

$$\Lambda = e^{v_i \omega_{i0} + \frac{1}{2} \theta_i \varepsilon_{ijk} \omega_{jk}}. \tag{26.96}$$

The following examples clarify the facts of known transformations.

Rotation

Rotations around the $z = 3$ axis with $\vec{v} = 0$ and $\vec{\theta} = (0, 0, \theta)$,

$$\begin{aligned}
 R_z(\theta) &= e^{\frac{1}{2}\theta_3(\omega_{12}-\omega_{21})} \\
 &= e^{\theta_3\omega_{12}} \\
 &= 1 + \theta\omega_{12} + \frac{1}{2!}\theta^2\omega_{12}^2 + \frac{1}{3!}\theta^3\omega_{12}^3 + \dots \\
 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{26.97}
 \end{aligned}$$

or as a second example the

Boost

along the x -axis with $\vec{v} = (v, 0, 0)$ and $\vec{\theta} = (0, 0, 0)$,

$$\begin{aligned}
 L_x(v) &= e^{v\omega_{10}} \\
 &= 1 + v\omega_{10} + \frac{1}{2!}v^2\omega_{10}^2 + \frac{1}{3!}v^3\omega_{10}^3 + \dots \\
 &= \begin{pmatrix} \cosh v & \sinh v & 0 & 0 \\ \sinh v & \cosh v & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{26.98}
 \end{aligned}$$

with $\tanh v = v/c$.

26.4.3 Covariant Dirac equation

The covariance of the Dirac equation requires that the 4^{er} spinor Ψ under Lorentz transformation according to

$$\begin{aligned}
 \Psi &\mapsto \Psi'(x') = S(\Lambda)\Psi(x) \\
 x' &= \Lambda x, \tag{26.99}
 \end{aligned}$$

transforms, where $S(\Lambda)$ fulfills the condition (see (26.20))

$$S^{-1}\gamma^\mu S = \Lambda^\mu{}_\nu\gamma^\nu \tag{26.100}$$

(26.100) tells us how to construct $S(\Lambda)$. Since $\Lambda \in \mathcal{L}_+^\uparrow$ is determined by the 6 parameters θ_i and v_i and for every $\Lambda \in \mathcal{L}_+^\uparrow$ there is a representation matrix $S(\Lambda)$ belongs to it in the Dirac space, we can for $S(\Lambda)$ make the following ansatz:

$$S(\Lambda) = e^{(v_i B_i + \frac{1}{2} \theta_i \varepsilon_{ijk} R_{jk})}, \quad (26.101)$$

where \vec{v} and $\vec{\theta}$ just define Λ ,

$$\Lambda = e^{(v_i \omega_{i0} + \frac{1}{2} \theta_i \varepsilon_{ijk} \omega_{jk})}, \quad (26.102)$$

and the B_i and R_{jk} are the generators of the representation of \mathcal{L} in the Dirac space. For small parameters \vec{v} , $\vec{\theta}$ is

$$\begin{aligned} S(\Lambda) &= \mathbb{1} + v_i B_i + \frac{1}{2} \theta_i \varepsilon_{ijk} R_{jk}, \\ \Lambda &= \mathbb{1} + v_i \omega_{i0} + \frac{1}{2} \theta_i \varepsilon_{ijk} \omega_{jk}, \end{aligned} \quad (26.103)$$

and with

$$\begin{aligned} S^{-1} \gamma^\mu S &= (1 - v_i B_i - \frac{1}{2} \theta_i \varepsilon_{ijk} R_{jk}) \gamma^\mu (1 + v_i B_i + \frac{1}{2} \theta_i \varepsilon_{ijk} R_{jk}) \\ &\approx \gamma^\mu - v_i [B_i, \gamma^\mu] - \frac{1}{2} \theta_i \varepsilon_{ijk} [R_{ik}, \gamma^\mu] \\ &= \delta^\mu_\nu \gamma^\nu + v_i (\omega_{i0})^\mu_\nu \gamma^\nu + \frac{1}{2} \theta_i \varepsilon_{ijk} (\omega_{ik})^\mu_\nu \gamma^\nu, \end{aligned} \quad (26.104)$$

for all choices of \vec{v} and $\vec{\theta}$, holds by comparison of coefficients

$$\begin{aligned} -[B_i, \gamma^\mu] &= (\omega_{i0})^\mu_\nu \gamma^\nu, \\ -[R_{jk}, \gamma^\mu] &= (\omega_{jk})^\mu_\nu \gamma^\nu. \end{aligned} \quad (26.105)$$

Inserting (26.92) for the boosts results

$$\begin{aligned} [\gamma^0, B_i] &= \gamma^i, & [\gamma^j, B_i] &= \delta_{ji} \gamma^0, \\ \Rightarrow B_i &= \frac{1}{2} \gamma^0 \gamma^i, \end{aligned} \quad (26.106)$$

and from (26.93) for the rotation,

$$\begin{aligned} [\gamma^0, R_{jk}] &= 0, & [\gamma^j, R_{jk}] &= -\delta_{ij} \gamma^k + \delta_{ik} \gamma^j, \\ \Rightarrow R_{jk} &= \frac{1}{2} \gamma^j \gamma^k. \end{aligned} \quad (26.107)$$

The generating matrices can be expressed using the Dirac matrices

$$\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} \quad (26.108)$$

With $\gamma^0 = \beta$ and $\gamma^i = \beta\alpha^i$, as well as $\beta^2 = 1$ follows immediately

$$B_i = \frac{1}{2}\alpha_i, \quad (26.109)$$

and with the definition

$$\gamma^5 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \quad (26.110)$$

we get

$$\begin{aligned} R_{jk} &= \frac{1}{2}\gamma^j\gamma^k \\ &= \frac{1}{2}\begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix} \\ &= -\frac{1}{2}\begin{pmatrix} \sigma_j\sigma_k & 0 \\ 0 & \sigma_j\sigma_k \end{pmatrix} = -\frac{i}{2}\varepsilon_{jkl}\begin{pmatrix} \sigma_l & 0 \\ 0 & \sigma_l \end{pmatrix} \\ &= -\frac{i}{2}\varepsilon_{jkl}\begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma_l \\ \sigma_l & 0 \end{pmatrix} \\ &= -\frac{i}{2}\varepsilon_{jkl}\gamma^5\alpha_l. \end{aligned} \quad (26.111)$$

Thus the representation belonging to $\Lambda(\vec{v}, \vec{\theta}) \in \mathcal{L}_+^\uparrow$ in the Dirac space is given by

$$S(\Lambda) = e^{\frac{1}{2}\vec{v}\cdot\vec{\alpha} - \frac{i}{2}\gamma^5\vec{\theta}\cdot\vec{\alpha}}. \quad (26.112)$$

Examples are the rotations around $\vec{\theta}$, in $\vec{\theta}$ direction with the angle θ (we use that $[\gamma^5, \vec{\alpha}] = 0$, $\gamma^5\gamma^5 = \mathbb{1}$, $(i\gamma^5\vec{\theta}\cdot\vec{\alpha})^2 = -\theta^2$),

$$S(R_{\vec{\theta}}) = \cos\frac{\theta}{2} - i\gamma^5\hat{\theta}\cdot\vec{\alpha}\sin\frac{\theta}{2}, \quad (26.113)$$

and the boosts along \vec{v} ,

$$S(L_{\vec{v}}) = \cosh\frac{v}{2} + \hat{v}\cdot\vec{\alpha}\sinh\frac{v}{2}. \quad (26.114)$$

Note that

$$S(L_{\vec{v}})\Psi_{0,s}(x) = \sqrt{\frac{E}{m}}\Psi_{\vec{p},s}(x'), \quad (26.115)$$

with the boosted sizes

$$\begin{aligned} t' &= t \cosh v + x \sinh v, \\ x' &= t \sinh v + x \cosh v, \\ E &= m \cosh v, \\ p &= m \sinh v, \\ v &= \tanh v, \end{aligned} \quad (26.116)$$

that means, S does not boost $\Psi_{0,s}$ not to $\Psi_{\vec{p},s}$ but to $\sqrt{E/m}\Psi_{\vec{p},s}$: the normalization is not covariant. Also note that $S(R)^\dagger = S(R)^{-1}$ is unitary, while $S(L)^\dagger = S(L)$ applies to the boosts. The density $\Psi^\dagger\Psi$ is thus not Lorentz invariant, whereas the 'metric density' $\bar{\Psi}\Psi = \Psi^\dagger\gamma^0\Psi$ is Lorentz invariant. Obviously it must then hold that

$$S^\dagger\gamma^0S = \gamma^0 \quad (26.117)$$

and thus also

$$\gamma^0S^\dagger\gamma^0 = S^{-1} \quad (26.118)$$

is fulfilled, which with the help of the relation

$$\begin{aligned} [\gamma^5\alpha_i, \gamma^0] &= 0, \\ \{\alpha_i, \gamma^0\} &= 0, \end{aligned} \quad (26.119)$$

is easy to see¹². One last comment about covariance: we usually consider a contraction

$$a^\mu b_\mu \quad (26.120)$$

as a manifestly covariant object. In this sense we could 'read' the Dirac equation as

$$(i\gamma^\mu\partial_\mu - m)\Psi = 0, \quad (26.121)$$

¹²Write down $S^\dagger\gamma^0S = \gamma^0$ for little ones \vec{v} and $\vec{\theta}$.

with the transformation behavior

$$\begin{aligned}\gamma'^\mu &= \Lambda^\mu_\nu \gamma^\nu \leftarrow \text{Vector}, \\ \partial'_\mu &= (\Lambda^{-1})^\nu_\mu \partial_\nu \leftarrow \text{Vector}, \\ \Psi'(x') &= \Psi(x) \leftarrow \text{Scalar}.\end{aligned}\tag{26.122}$$

This would be consistent with the transformation behavior as discussed on page 293, and also the condition

$$\{\gamma'^\mu, \gamma'^\nu\} = 2g^{\mu\nu}\tag{26.123}$$

would be fulfilled. But, then every observer would have to choose different γ matrices, and of course Ψ is not a scalar field either. We really want to understand (26.15) as *form invariant*, we want to set the same matrices for all observers in advance

$$\begin{aligned}\gamma^0 &= \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \\ \gamma^i &= \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}\end{aligned}\tag{26.124}$$

Then the corresponding transformation law is given by

$$\begin{aligned}\gamma'^\mu &= \gamma^\mu, \\ \partial'_\mu &= (\Lambda^{-1})^\nu_\mu \partial_\nu \\ \Psi'(x') &= S(\Lambda)\Psi(x), \\ x' &= \Lambda x,\end{aligned}\tag{26.125}$$

and γ^μ should *not* be interpreted as 4th vector.

Chapter 27

Relativistic Particles: Weyl-Equation*

The Dirac equation and spinors can alternatively be represented in the Weyl form; For massless particles, this representation leads us to the Weyl equation and the independent left and right-handed 2-nd spinors Ψ^L and Ψ^R , the Weyl spinors. A finite mass couples the two equations for the chiral Weyl spinors. It is questionable whether there are massless fermions. A possible family of massless particles were the neutrinos; however, the discovery of the neutrino oscillations points to a finite mass (and physics 'beyond the standard model'). Where the small mass comes from (standard Higgs, different Higgs, different mechanism) is unclear, as is also unclear whether the neutrinos are Dirac fermions (with anti-particles) or Majorana fermions (particles is its own anti-particle).

We first investigate the consequence of the masslessness of spin-1/2 fermions before we come across the Weyl representation of the Dirac equation and the Weyl equations via an alternative derivation of the Dirac equation.

Let's put in the Dirac spinors

$$\begin{aligned} u(\vec{p}, s) &\equiv \sqrt{E_p + mc^2} \begin{pmatrix} 1 \\ \frac{c \vec{\sigma} \cdot \vec{p}}{E_p + mc^2} \end{pmatrix} \chi^{(s)}, \\ v(\vec{p}, s) &\equiv \sqrt{E_p + mc^2} \begin{pmatrix} \frac{c \vec{\sigma} \cdot \vec{p}}{E_p + mc^2} \\ 1 \end{pmatrix} (-i\sigma_2 \chi^{(s)}), \end{aligned} \quad (27.1)$$

the mass $m = 0$, we get (with $E_p = cp$)

$$\begin{aligned}
 u(\vec{p}, s) &= \sqrt{E_p} \begin{pmatrix} 1 \\ \vec{\sigma} \cdot \hat{p} \end{pmatrix} \chi^{(s)} = \sqrt{E_p} \begin{pmatrix} 1 \\ 2\lambda \end{pmatrix} \chi_\lambda \\
 &= u(\vec{p}, \lambda), \\
 v(\vec{p}, s) &= \sqrt{E_p} \begin{pmatrix} \vec{\sigma} \cdot \hat{p} \\ 1 \end{pmatrix} (-i\sigma_2 \chi^{(s)}) = \sqrt{E_p} \begin{pmatrix} -2\lambda \\ 1 \end{pmatrix} (-i\sigma_2 \chi_\lambda^*) \\
 &= v(\vec{p}, \lambda),
 \end{aligned} \tag{27.2}$$

with χ_λ the 2-component spin spinor with quantization axis \hat{p} (the transition $s \rightarrow \lambda$ should remind us of the fixed quantization axis along \hat{p}),

$$\begin{aligned}
 \vec{\sigma} \cdot \hat{p} \chi_\lambda &= 2\lambda \chi_\lambda, \quad \lambda = \pm \frac{1}{2}, \\
 \vec{\sigma} \cdot \hat{p} \sigma_2 \chi_\lambda^* &= -\sigma_2 \vec{\sigma}^* \cdot \hat{p} \chi_\lambda^* \\
 &= -2\lambda \sigma_2 \chi_\lambda^*.
 \end{aligned} \tag{27.3}$$

Since the (2-block) components of the 4^{er} spinors (27.2) are the same in terms of amount

$$\begin{aligned}
 \gamma^5 u(\vec{p}, \lambda) &= 2\lambda u(\vec{p}, \lambda), \\
 \gamma^5 v(\vec{p}, \lambda) &= -2\lambda v(\vec{p}, \lambda),
 \end{aligned} \tag{27.4}$$

with

$$\gamma^5 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \tag{27.5}$$

that means u and v are eigenvectors of the operator γ^5 . Further are

$$P_\pm = \frac{1}{2} (1 \pm \gamma^5) \tag{27.6}$$

the projectors for the eigenstates of γ^5 . The operator γ^5 defines the **Chirality** and acts equal to u and v ¹ like the helicity operator $\vec{\Sigma} \cdot \hat{p}$ with

$$\vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}. \tag{27.7}$$

¹The chirality coincides with helicity for massless particles; for massive particles, helicity changes when the observer 'overtakes' the particle.

With positive **helicity**, $\lambda = 1/2$, we get right-handed particles (described by $u(\vec{p}, 1/2)$) or left-handed anti-particles ($v(\vec{p}, -1/2)$) with the properties (cf. Abb. 27.1)

$$\begin{aligned}\gamma^5 u(\vec{p}, \tfrac{1}{2}) &= u(\vec{p}, \tfrac{1}{2}), \\ \gamma^5 v(\vec{p}, -\tfrac{1}{2}) &= v(\vec{p}, -\tfrac{1}{2}),\end{aligned}\tag{27.8}$$

$$\begin{aligned}P_+ u(\vec{p}, \tfrac{1}{2}) &= u(\vec{p}, \tfrac{1}{2}), \\ P_+ v(\vec{p}, -\tfrac{1}{2}) &= v(\vec{p}, -\tfrac{1}{2}),\end{aligned}\tag{27.9}$$

$$\begin{aligned}P_- u(\vec{p}, \tfrac{1}{2}) &= 0, \\ P_- v(\vec{p}, -\tfrac{1}{2}) &= 0,\end{aligned}\tag{27.10}$$

and left-handed particles ($u(\vec{p}, -1/2)$), right-handed anti-particles ($v(\vec{p}, 1/2)$) with negative helicity $\lambda = -1/2$ characterized via

$$\begin{aligned}\gamma^5 u(\vec{p}, -\tfrac{1}{2}) &= -u(\vec{p}, -\tfrac{1}{2}), \\ \gamma^5 v(\vec{p}, \tfrac{1}{2}) &= -v(\vec{p}, \tfrac{1}{2}),\end{aligned}\tag{27.11}$$

$$\begin{aligned}P_+ u(\vec{p}, -\tfrac{1}{2}) &= 0, \\ P_+ v(\vec{p}, \tfrac{1}{2}) &= 0,\end{aligned}\tag{27.12}$$

$$\begin{aligned}P_- u(\vec{p}, -\tfrac{1}{2}) &= u(\vec{p}, -\tfrac{1}{2}), \\ P_- v(\vec{p}, \tfrac{1}{2}) &= v(\vec{p}, \tfrac{1}{2}).\end{aligned}\tag{27.13}$$



Neutrinos im Standardmodell

Fig. 27.1: Particles of positive (left) and negative (right) helicity. The latter are relevant in the standard model where (massless) left-handed neutrinos (\vec{p} and $\vec{\sigma}$ anti-parallel) and right-handed anti-neutrinos appear (\vec{p} and $\vec{\sigma}$ parallel) — parity symmetry is broken.

Note that the right and left states can only keep their identity if $m = 0$: Massless particles have no rest system and their momentum cannot be changed by reversing the motion, with which the helicity is an invariant of the movement (the same situation of fixed helicity has already been found

with the photon). Since the impulse reverses under parity at P , while the spin is invariant under P , these states are not parity invariant: With the parity operator γ^0 it holds that $P_{\pm} \gamma_0 = \gamma_0 P_{\mp}$ and γ^0 transforms a right-handed state into a left-handed one and vice versa. It can be shown (see later) that the Dirac equation for $m = 0$ simplifies into the Weyl equations,

$$i\hbar\partial_t\Psi^{L/R} = \mp c \vec{p} \cdot \vec{\sigma} \Psi^{L/R}, \quad (27.14)$$

which is not manifestly covariant under space inversion. The Weyl equation was therefore rejected as a physically relevant equation until parity breakdown of the weak interaction was finally discovered in the 1960s. In the standard model only left-handed neutrinos and right-handed anti-neutrinos occur, with which the parity is explicitly broken.

27.1 Alternative derivation of the Dirac equation

We want to set up a relativistic equation for a spin 1/2- particle. For a charged particle (with charge q) the substitution $\vec{p} \rightarrow \vec{p} - q\vec{A}/c$ creates the coupling to the electromagnetic field. In particular, we then expect a (Spin-Zeeman) term of the form

$$H^{\text{Spin}} = -\frac{q\hbar}{2mc} \vec{\sigma} \cdot \vec{B} \quad (27.15)$$

in the Hamiltonian. The (non-relativistic) kinetic Hamiltonian

$$H^{\text{kin}} = \frac{p^2}{2m} \quad (27.16)$$

cannot generate this term, but the operator can

$$H^{\text{kin}} = \frac{1}{2m} (\vec{\sigma} \cdot \vec{p}) (\vec{\sigma} \cdot \vec{p}), \quad (27.17)$$

which also leads to the kinetic energy $p^2/2m$ there

$$p_i \sigma_i p_j \sigma_j = \sigma_i^2 p_i^2 + (1/2) \{\sigma_i, \sigma_j\} p_i p_j = p^2.$$

The minimal coupling to the electromagnetic field then results in the Hamiltonian²

$$\begin{aligned} H &= \frac{1}{2m} \vec{\sigma} \cdot (\vec{p} - q\vec{A}/c) \vec{\sigma} \cdot (\vec{p} - q\vec{A}/c) \\ &= \frac{1}{2m} (\vec{p} - q\vec{A}/c)^2 - \frac{q\hbar}{2mc} \vec{\sigma} \cdot \vec{B}. \end{aligned} \quad (27.18)$$

²We use the equation $(\vec{\sigma} \cdot \vec{A})(\vec{\sigma} \cdot \vec{B}) = \vec{A} \cdot \vec{B} + i \vec{\sigma} \cdot (\vec{A} \wedge \vec{B})$.

This gives us the gyromagnetic factor $g = 2$ for a spin-1/2 particle with charge q from the non-relativistic equation (27.17). Let us expand (27.17) to a relativistic theory, so the following should hold (with Ψ a 2-nd spinor)

$$\begin{aligned} E^2 - c^2 p^2 &= m^2 c^4 \\ \downarrow E &= i\hbar\partial_t, \quad \vec{p} = -i\hbar\vec{\nabla}, \quad p^2 \rightarrow (\vec{\sigma} \cdot \vec{p})^2, \\ (i\hbar\partial_t - c\vec{\sigma} \cdot \vec{p})(i\hbar\partial_t + c\vec{\sigma} \cdot \vec{p})\Psi &= m^2 c^4 \Psi. \end{aligned} \quad (27.19)$$

We create from (27.19) using the usual trick

$$\begin{aligned} \Psi^R &= \frac{1}{mc^2} (i\hbar\partial_t + c\vec{\sigma} \cdot \vec{p}) \Psi, \\ \Psi^L &= \Psi, \end{aligned} \quad (27.20)$$

a four-component differential equation of 1st order,

$$\begin{aligned} (i\hbar\partial_t - c\vec{\sigma} \cdot \vec{p})\Psi^R &= mc^2 \Psi^L, \\ (i\hbar\partial_t + c\vec{\sigma} \cdot \vec{p})\Psi^L &= mc^2 \Psi^R. \end{aligned} \quad (27.21)$$

With the 4-spinor

$$\Psi_W \equiv \begin{pmatrix} \Psi^R \\ \Psi^L \end{pmatrix} \quad (27.22)$$

we get the Dirac equation in the Weyl representation (we set $\hbar = c = 1$),

$$(i\gamma_W^\mu \partial_\mu - m) \Psi_W = 0 \quad (27.23)$$

with

$$\begin{aligned} \gamma_W^0 &= \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, & \gamma_W^i &= \begin{pmatrix} 0 & -\sigma_i \\ \sigma_i & 0 \end{pmatrix}, \\ \gamma_W^5 &= \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, & \{\gamma_W^\mu, \gamma_W^\nu\} &= 2g^{\mu\nu}. \end{aligned} \quad (27.24)$$

With the definition of the Dirac spinor

$$\Psi_D = \begin{pmatrix} \Psi^R + \Psi^L \\ \Psi^R - \Psi^L \end{pmatrix} = \begin{pmatrix} \mathbb{1} & \mathbb{1} \\ \mathbb{1} & -\mathbb{1} \end{pmatrix} \Psi_W, \quad (27.25)$$

the Dirac equation is obtained in the Dirac representation

$$(i\gamma_D^\mu \partial_\mu - m) \Psi_D = 0, \quad (27.26)$$

with the matrices

$$\begin{aligned}\gamma_{\text{D}}^{\mu} &= \frac{1}{2} \begin{pmatrix} \mathbb{1} & \mathbb{1} \\ \mathbb{1} & -\mathbb{1} \end{pmatrix} \gamma_{\text{W}}^{\mu} \begin{pmatrix} \mathbb{1} & \mathbb{1} \\ \mathbb{1} & -\mathbb{1} \end{pmatrix}, \\ \gamma_{\text{D}}^0 &= \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \quad \gamma_{\text{D}}^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}, \\ \gamma_{\text{D}}^5 &= \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}.\end{aligned}\tag{27.27}$$

27.2 Massless fermions in Weyl representation

The following projectors in Weyl representation

$$\begin{aligned}P_{\pm}^{\text{W}} &= \frac{1}{2}(\mathbb{1} \pm \gamma_{\text{W}}^5), \\ P_{+}^{\text{W}} &= \begin{pmatrix} \mathbb{1} & 0 \\ 0 & 0 \end{pmatrix}, \\ P_{-}^{\text{W}} &= \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{1} \end{pmatrix},\end{aligned}\tag{27.28}$$

project the 4-spinor Ψ_{W} onto right and left hand states,

$$\begin{aligned}P_{+}^{\text{W}} \Psi_{\text{W}} &= \Psi^{\text{R}}, \\ P_{-}^{\text{W}} \Psi_{\text{W}} &= \Psi^{\text{L}}.\end{aligned}\tag{27.29}$$

The two two-component coupled equations (27.21) decouple for the massless case $m = 0$ and we find the equations

$$\begin{aligned}\underbrace{i\hbar\partial_t}_{E=pc>0} \Psi^{\text{R/L}} &= \pm c \vec{\sigma} \cdot \vec{p} \Psi^{\text{R/L}}, \\ pc &= E > 0,\end{aligned}\tag{27.30}$$

with the eigenstates

$$\begin{aligned}\vec{\sigma} \cdot \hat{p} \Psi^{\text{R}} &= \Psi^{\text{R}}, \\ \vec{\sigma} \cdot \hat{p} \Psi^{\text{L}} &= -\Psi^{\text{L}},\end{aligned}\tag{27.31}$$

that is, Ψ^{R} and Ψ^{L} have the helicity +1 and -1.

Quarks					Leptonen		
		Farbe	Ladung	Masse		Ladung	Masse
1	up	RGB	2/3	$\sim 5 \text{ MeV}$	e	-1	$\sim 0.5 \text{ MeV}$
	down	RGB	-1/3	$\sim 7 \text{ MeV}$	ν_e	0	$< 18 \text{ eV}$
2	charme	RGB	2/3	$\sim 1.5 \text{ GeV}$	μ	-1	$\sim 105 \text{ MeV}$
	strange	RGB	-1/3	$\sim 200 \text{ MeV}$	ν_μ	0	$< 0.25 \text{ MeV}$
3	top	RGB	2/3	$\sim 174 \text{ GeV}$	τ	-1	$\sim 1.8 \text{ GeV}$
	bottom	RGB	-1/3	$\sim 4.5 \text{ GeV}$	ν_τ	0	$< 35 \text{ MeV}$

Fig. 27.2: The three generations of spin-1/2 fermions in the standard model.

27.2.1 Embedding in the standard model

There are three classes of fundamental particles within the standard model

- Spin 1/2 fermions; 3 generations of quarks and leptons.
- Spin 1 gauge bosons, γ , W^\pm , Z^0 of the electroweak force; Gluons in eight colors for Quantum-Chromo-Dynamics QCD.
- Spin 0 scalar Higgs particle, which gives the W^\pm and Z^0 a mass.

The three generations of spin-1/2 fermions are summarized in the table 27.2. We have dealt with the electrons and neutrinos that obey the Dirac equation, as well as the π -mesons resulting from two quarks that are described by the Klein-Gordon equation: two quarks $q \bar{q}$ plus gluons give colorless, bosonic mesons, three quarks $q_R \ q_B \ q_G$ plus mesons plus gluons give the colorless, fermionic baryons.

Index

- Abschirmung
 - Thomas-Fermi, 54
- acetylen, 107
- ammonia, 105
- anti-particle, 263
- anti-particles, 268
- anti-symmetric wave function, 19
- Antiparticle, 299
- approximation
 - Born-Oppernheimer, 86
- atoms
 - cold, 227
 - ground state, 70
 - term symbols, 70
- Auswahlregeln, 84
- barions, 20
- Basis, 221
- BCS, 24
- BEC, 24
- benzene, 107
- Boltzmann, 206
- bonding types, 101
- Born-Oppenheimer approximation,
 - 86, 87
- Bose Einstein condensation, 257
 - critical density, 257
 - critical temperature, 257
 - fugacity, 258
 - transition line, 259
- Bose gas
 - Stefan-Boltzmann law, 249
- Bose-condensation, 28
- boson occupation numbers, 113
- Bosonen
 - Streuung, 30
- Bosons
 - field theory, 149, 171
- bosons, 20
 - condensate, 28
 - excited states, 29
- Carbon C, 106
- charge conjugation, 279
- Chiralität, 318
- classical gas
 - Boltzmann gas, 228
- Clebsch-Gordan coefficients, 73
- coherent states, 163
- cold atoms, 227
- completeness, 164
- composite particles, 23
- conduction electrons, 27
- configuration, 63
- configurational integral, 234
- Cooper pairs, 28
- correspondence principle, 264
- Coulomb
 - Diversification, 281
- Coulomb gauge
 - tensor notation, 180
- Coulomb-gauge, 173
- critical density
 - BEC, 257

- critical temperature
 - BEC, 257
- Debye temperature, 254
- density matrix, 223
- Dirac
 - Adjungierte, 294
 - Gleichung, 292
 - Matrizen, 292
 - Norm, 294
 - Spinor, 292, 296
- dot
 - Klein-Gordon, 267
- Dulong-Petit, 254
- effect
 - Paschen-Back, 83
 - Zeeman, 68, 82
- Eichbosonen, 264
- Eichung
 - Coulomb, 173
- electromagnetic field, 171
- Elemente, 62
- ensemble
 - canonical, 225
 - micro canonical, 225
 - grand canonical, 225
- EPR, 1
- equations
 - Maxwell, 171
 - Proca-, 189
- equivalence class, 17
- ethylen, 107
- exclusion principle, 26, 300
- Fermi gas
 - degenerate, 240
 - dilute, 239
 - equation of state, 240, 241
 - Fermi energy, 240
 - Fermi wave vector, 241
 - in a metal, 242
 - in a white dwarf, 243
 - in neutron stars, 244
 - internal energy, 241
 - Sommerfeld expansion, 239
 - specific heat, 241
- Fermi Liquid Theory, 228
- Fermi sea, 138
- Fermienergie, 28
- fermion occupation numbers, 113
- fermions, 20
 - excited states, 29
 - scattering, 31
- Fermisee, 26
- Feshbach resonances, 227
- field
 - em-, 171
- field operators, 126
- fine structure, 67
- fine structure constant, 203
- Fock-space, 111
- Foldy-Wouthuysen-Transformation, 304
- fusion, 94
- γ -Matrizen, 292
- gas
 - classical, 228
 - classical limit, 232
 - density parameter, 238
 - quantum, 229
- Gase
 - Boltzmannzählen, 228
 - Bosegase
 - BE-Kondensation, 255
- gauge
 - Lorenz, 173
- gauge bosons, 20
- Gleichungen
 - Dirac, 292

- gluons, 20
- Grassman–algebra, 169
- ground state, 70
 - atomic, 63
- H_2^+ ion, 90
- H_2 molecule, 94
- Haber-Bosch, 106
- Hartree & Hartree-Fock Approximation, 49
- Hartree equations, 51
- Heitler-London, 95
- Helium, 24, 28
- Helizität, 183, 318, 320
- Higgs, 20
- Hilbertraum, 221
- Hund rules, 68
- Hund-Mullikan, 94
- hybridization, 103
- hydrogen
 - ortho-, 32
 - para-, 32
- hyperfine structure, 68
- identical nuclei, 99
- Identical Particles, 11
- identical particles, 14
- integro-differential equations
 - Hartree-Fock-, 53
- interaction
 - statistical, 234
- irreducible representations of S_n , 35
- jj -coupling, 73
- K mesons, 263
- Klein-Gordon
 - Adjoint, 276
 - Adjointed, 277
 - equation, 264
 - Scalar, 267
 - scalar product, 276, 277
 - Standard, 267
- Koopmans theorem, 53
- Kovarianz
 - Dirac Gleichung, 308
- Kramers-Heisenberg, 211
- Lamb-shift, 212
- Landé factor, 82
- Landau
 - Fermi Liquid Theory, 228
- leptons, 20
- light scattering, 211
- Liouville theorem, 224
- Lorentz
 - Gruppe, 308
- Lorenz gauge
 - tensor notation, 180
- Lorenz-gauge, 173
- LS -coupling, 73
- macroscopic quantum effect, 108
- magnons, 110
- many-body problem, 49
- Maser, 106
- Maxwell Boltzmann distribution, 239
- Maxwell equations, 171
- mesons, 20
- methane, 106
- MQP
 - macroscopic quantum phenomena, 232
- normal ordering, 163
- observable, 221
- Occupation number, 119
- occupation numbers, 113
- optical lattices, 227
- orbital wave function, 21

- Oxygen O₂, 103
- Pair generation, 269
- pairing of fermions, 24
- paramagnetic
 - O₂ Oxygen, 103
- particles, identical, 14
- Paschen-Back effect, 83
- Pauli
 - Matrizen, 291
- Pauli exclusion, 21, 26, 230
- Pauli Exclusion principle, 300
- Pauli principle, 113
- permutation
 - $\pi(i)$, 14
- permutation cycle, 16
- permutation group S_n , 11
- Pfadintegrale, 161
- Phononen, 218
- phonons, 87, 110
- photons, 110, 247
- Planck
 - radiation law, 251
- Poincaré Gruppe, 190
- Polaronen, 218
- probability
 - a priori, 223
- Proca-equation, 189
- processes
 - real, 216
 - virtual, 216
- product wave function, 50
- pseudopotential, 235
- Quantenkryptographie, 1
- Quantenrechner, 1
- quantum gas
 - bosons, 229
 - fermions, 229
 - partition sum, 230
- quantum simulation, 227
- quarks, 20
- race (german ‘Rasse’)
 - orbital-race, 33
 - spin-race, 33
- Raman scattering, 212
- random phases, 222
- Rayleigh scattering, 211
- Rayleigh-Jeans radiation law, 250
- reservoir, 222
- rotation modes
 - in molecules, 86
- Russel-Saunders coupling, 73
- saturation
 - of a bond, 100
- scaling
 - specific heat, 255
- scatter
 - Thomson, 211
- scattering
 - s-wave, 234
 - Coulomb, 281
 - Raman, 212
 - Rayleigh, 211
 - of bosons, 30
 - of fermions, 31
- selection rules, 204, 205
- Slater determinant, 26, 27, 29
- S_n , irreducible representation, 35
- S_n , permutation group, 16
- S_n , symmetric group, 16
- Sokhotski, 215
- specific heat
 - Debye, 254
 - Dulong-Petit, 254
- spin
 - and race (german ‘Rasse’), 33

- spin and race (german ‘Rasse’), 11, 22, 32, 34
- spin wave function, 21
 - anti-symmetric, 48
 - symmetric-, 48
- spin-orbit coupling, 79
- Spinore, Dirac-, 296
- state vector, 221
- Stefan-Boltzmann law, 249
- Streuung, 29
- superconductivity, 28
- superfluid, 24
- superfluidity, 24, 28
- superposition
 - incoherent, 223
- symmetric group S_n , 16
- symmetric wave function, 19
- symmetry
 - particle-hole, 242
 - permutation, 15
- tensor operator
 - vector, 76
- term symbol, 66
- tetrahedron, 105
- theorem
 - Koopmans, 53
 - multinomial, 228
 - of spin and race (german ‘Rasse’), 33
 - spin and statistics, 19
 - Wigner-Eckart, 73
- thermodynamics
 - 3rd law, 226
- Thomas-Fermi atom, 58
 - shells, 59
- Thomson scatter, 211
- time scales
 - in molecules, 87
- transition line
 - BEC, 259
- two-electron atom, 47
- Two-particle Hamilton, 133
- Verschränkung, 1
- vibrational amplitude
 - in a molecule, 86
- vibrational energy
 - in molecules, 86
- virtual process, 216
- water, 104
- wave function
 - anti-symmetric, 19
 - spin & orbit, 21
 - symmetric, 19
- Weyl
 - Darstellung, 321
 - Gleichungen, 320
- Wien
 - displacement law, 250
- Wigner-Eckart theorem, 73
- X-ray shells, 60
- Young tableau, 36
- Zeeman, 302
- Zeeman-effect
 - anomalous, 83
 - normal, 83
- Zeeman-splitting, 68, 82