Lecture Notes

Quantum Theory

by

Prof. Maximilian Kreuzer

Institute for Theoretical Physics
Vienna University of Technology

covering the contents of

132.056 Quantentheorie I

and

132.062 Quantentheorie II

Edition 07/08 — Version December 29, 2007

Links

The current version of the notes, as well as information on lectures and exams, is available at
`http://hep.itp.tuwien.ac.at/~kreuzer/QT.html`
Reports of typos and errors and suggestions for improvements are appreciated, e.g. by e-mail to `maximilian.kreuzer@tuwien.ac.at` (if possible after cross-checking with the current version).

Preface

The structure of these lecture notes is mainly motivated by the curricula of the bachelor's and master's programs of the faculty of physics at the Vienna University of Technology, which requires a division of quantum mechanics into two parts. The first part

- **Quantum Theory I: chapters 1 – 7**

should make available the prerequisites for the subsequent lecture on atomic physics and has to be covered in 45 units of 45 minutes each. After historic recollections in the introduction the principles of quantum theory are first illustrated for one-dimensional examples in chapter 2 and then presented in the proper formalism in chapter 3. In chapters 4 and 5 we solve the Schrödinger equation for the spherically symmetric hydrogen atom and treat the quantization and the addition of general angular momenta, respectively. Chapter 6 introduces approximation techniques and chapter 7 initiates relativistic quantum mechanics and derives the Pauli equation and the fine structure corrections in the non-relativistic limit of the Dirac equation.

The systematic discussion of symmetries as well as identical particles and many particle theory had to be postponed to part 2,

- **Quantum Theory II: chapters 8 – 11.**

In chapter 8 we start with 3-dimensional scattering theory. Transformations, symmetries and conservation laws are discussed in chapter 9 and applied to non-relativistic and relativistic contexts. In chapter 10 we discuss many particle systems. The Hartree–Fock approximation is used as a motivation for the introduction of the occupation number representation and the quantization of the radiation field. These three chapters are largely independent so that their order could be permuted with little modifications. In the last chapter we discuss semiclassical methods and the path integral.

Acknowledgements

A first draft of these lecture notes was created by Katharina Dobes (chap. 1,6,10), Wolfgang Dungel (chap. 3,11), Florian Hinterschuster (chap. 4,5,9) and Daniel Winklehner (2,7,8,9) as a project work. While the text was then largely rewritten by the lecturer, the draft provided many valuable ideas for the structure and the presentation of the contents.
My acknowledgements also go to my colleagues at the Institute for Theoretical Physics for sharing their knowledge and ideas, with special thanks to Harald Grosse (Vienna University), Anton Rebhan and Karl Svozil, whose expertise was of great help, and to the late Wolfgang Kummer, from whom I learned quantum mechanics (and quantum field theory) in the first place. In addition to input from many of the books in the references I took advantage of the excellent lecture notes of Profs. Burgdörfer, Hafner and Kummer. Often as a first and sometimes as a last resort I used Wikipedia and Google. Last but not least, many thanks to the students who are helping to improve these lecture notes by reporting errors and typos.

Contents

Chapter 1

Introduction

1.1 Historical notes

In the nineteenth century the profession of a specialized scientist was created and the main scientific activity moved to university-like institutions. As a result scientific research flourished. One of the major and at the same time one of the oldest branches of physics was *mechanics*. Its foundation dates back to 1687, when *Isaac Newton* (1642–1727) formulated the principles of mechanics and the gravitational law. The theory was further developed, among others, by *Joseph Louis Lagrange* (1736–1813), who formulated the dynamical equations, *Carl Friedrich Gauss* (1777–1855), who introduced the 'principle of least constraints', as well as *William Rowan Hamilton* (1805–1865) and *Carl Gustav Jacob Jacobi* (1804–1851), who worked out a new scheme of mechanics. They stated that motions of objects in nature always occur with least action, which was defined as the time integral over the so-called Lagrange function.

On the basis of these discoveries *thermodynamics* was developed as a new branch of physics. *Julius Robert Mayer* (1814–1878) and *James Prescott Joule* (1818–1889) found out that heat fully corresponds to energy. The first and the second law of thermodynamics were first explicitly stated in a book by *Rudolf Emanuel Clausius* (1822–1888) in 1850. Clausius also shaped the concept of entropy in 1865. Maxwell's velocity distribution for the kinetic theory of gases was then explained by Boltzmann (1844–1906) with statistical mechanics. At the end of the 19^{th} century this lead to the important problem of blackbody radiation, i.e. the quest for a theoretical understanding of the spectrum emitted by a perfect absorber (see chapter 1.2.1).

Electrodynamics and optics were two separate disciplines until *Heinrich Hertz* (1857–1894) proved in 1888 that light possesses all characteristics of an electromagnetic wave. The first quantitative description of an electrical force (attractive or repulsive) was made by *Charles Auguste de Coulomb* (1736–1806) in 1785. *André Marie Ampère* (1775–1836) was the first to speak of electrodynamics in 1822. In 1826 *Georg Simon Ohm* (1787–1854) formulated what is

nowadays known as Ohm's law. In 1833 Gauss and *Wilhelm Weber* (1804–1891) invented the telegraph. One of the most important contributions was made by *Michael Faraday* (1792–1867) who discovered electromagnetic induction and electrolysis. Based on this work *James Clerk Maxwell* (1831–1879) found a complete system of equations that describes all electromagnetic phenomena.

We conclude our excursion into the evolution of physics till the beginning of the 20^{th} century with a short glance at *atomism*. In ancient Greece, *Demokritus* introduced the idea of atoms as indivisible building block of matter. This idea was reintroduced in the 17^{th} century after it had been mostly forgotten throughout the middle ages. Chemists focused on matter that could not be separated by chemical methods. Physicists, on the other hand, tried to explain phenomena such as pressure, temperature, specific heat and viscosity in terms of the particles (molecules) that gases consist of. This approach is called the kinetic theory of gases. Out of this statistical mechanics evolved. At the beginning of the 20^{th} century the atomic hypothesis was at last widely accepted among the scientific community. It was not until 1905, however, that a theoretical proof for the existence of atoms was made simultaneously by *Albert Einstein* (1879–1955) and *Marian Smoluchowski* (1872–1917) in their work on Brownian motion. Still the structure of an atom and the ways in which the atoms of different elements differ were not yet understood at all. All in all, one can say that atomic physics was in its infancy at the turn of the century.

In the late 19^{th} century some very important discoveries were made: In 1885 *Wilhelm Conrad Röntgen* (1845–1923) discovered what he called X-rays. This phenomenon reminded *Antoine-Henri Becquerel* (1852-1908) of his work on phosphorescent stone and he began to search for a stone with similar properties. He finally found one – a uranium salt – and realized that he had observed a new kind of radiation emitted by radioactive material. This radiation later on turned out to be a very powerful tool for investigating atomic structure. In 1897 *Joseph John Thomson* (1856–1940) was able to identify the first elementary particle, the electron, and to determine its charge to mass ratio. The reaction of the scientific world was rather unenthusiastic. Some physicists didn't even believe in the concept of atoms. Others thought that atom and electrons were too small to be made objects of speculation. Later, Lord Kelvin and J.J. Thompson together developed a theory of atomic structure.

The 20^{th} century

There were some physicists at the end of the 19^{th} century who believed that physics had come to some kind of an "end of evolution" and that there was hardly anything interesting left to be found out. Classical mechanics was able to describe almost all phenomena that had been detected and thus seemed to be satisfactory. It was a simple and unified theory.

Physicists distinguished two completely different categories of objects – matter and radiation: According to Newtonian mechanics *matter* is built out of localizable corpuscles with a well-defined position and velocity. One can thus compute the time evolution of a system as soon as one knows this data at a given moment. The corpuscular theory could even be extended to the microscopic scale of solid bodies (i.e. to molecules or atoms). According to thermodynamics and statistical mechanics macroscopic parameters thus derive from the motion of the (microscopic) particles. *Radiation*, on the other hand, could well be explained with Maxwell's laws that are able to link electromagnetism, optics and acoustics. As light was capable of interference and diffraction, which are clearly associated with waves, light was eventually considered to be a form of radiation.

At the beginning of the 20^{th} century some experiments and theoretical problems implied, however, that this distinction between radiation and matter was not entirely valid. Physicists were confronted with a bunch of data that seemed hard to explain within the framework of what we now call classical physics and were even forced to look for different and at first strange new concepts. This lead to the idea of *quantization* of physical entities and to *wave-particle dualism*. The important achievements of quantum physics in the first three decades of the new century include the following:

- **1900** Max Planck derives his formula for blackbody radiation by introducing a constant h that determines the sizes of energy packages, called *quanta*, of electromagnetic radiation.
- **1905** Albert Einstein explaines the photoelectric effect in terms of the same constant.
- **1906** J.J. Thompson discovers the proton.
- **1910** Robert Millikan measures the elementary electric charge.
- **1911** After observations on the scattering of alpha particles caused by atoms, Ernest Rutherford introduces the first modern picture of the atom.
- **1913** Niels Bohr explains spectral lines and the stability of atoms by postulating quantization of angular momentum.
- **1923** Arthur Compton gives an explanation for the scattering of photons on electrons by assigning the momentum $\vec{p} = \hbar\vec{k}$ to photons.
- **1924** Wolfgang Pauli formulates his exclusion principle.
- **1925** Louis de Broglie's doctoral thesis states that matter particles like photons are associated to waves of wavelength $\lambda = h/p$.
- **1925** Werner Heisenberg invents matrix mechanics, which assigns noncommuting matrix operators to dynamical variables.

- **1926** Erwin Schrödinger finds his equation, which describes wave mechanics.
- **1927** Werner Heisenberg derives the uncertainty relation.
- **1927** Max Born suggests the probabilistic interpretation of the wavefunction.
- **1928** Paul Adrien Maurice Dirac discovers the Dirac equation, which combines quantum mechanics with special relativity. This lead him to predict the existence of antimatter.
- **1932** Anderson's discovery of positrons in cosmic ray showers confirms Dirac's prediction.
- **1932** Chadwick observes a neutron (predicted by Rutherford in 1920).

We next discuss some of the problems mentioned above in more detail.

1.2 Limitations of classical physics

1.2.1 Blackbody radiation

A blackbody is by definition a surface that absorbs radiation entirely. One can imagine a blackbody to be a closed container with a well-absorbing surface and with a small window brought to a uniform temperature, i.e in thermal equilibrium. Radiation entering the container through the small window is reflected several times within the blackbody (see figure 1.1) and has a negligible chance for reemerging through the window. Hence this container is a perfect absorber. According to Kirchhoff's law the ratio of the emission power, or emittance, to the absorption coefficient is the same for all bodies at the same temperature. Since a blackbody has a maximum absorption coefficient it must therefore also be the most efficient emitter.

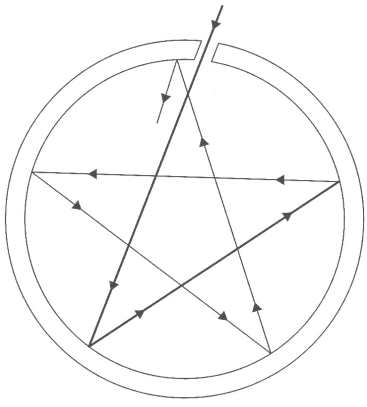


Figure 1.1: Schematic illustration of a blackbody

Rayleigh and Jeans used electrodynamics and thermodynamics to deduce a formula for the energy $u(\nu)$ per frequency interval that is emitted by such a blackbody:

$$u_{RJ} = \frac{8\pi\nu^2}{c^3} k_B T, \tag{1.1}$$

where $k_B = 1.381 \cdot 10^{-23} J/K$ is Boltzmann's constant and c is the speed of light. This formula fits the experimentally observed curve for low frequencies quite well but it deviates from the experimental value and diverges at larger ones (cf. figure 1.2)! The formula predicts an infinite total energy emission and hence cannot possibly be correct. This indicates an inconsistency between statistical mechanics and electrodynamics.

Wien also tried to describe the radiation of a blackbody. Upon general considerations he came to the conclusion that the proper term for $u(\nu)$ must be of the form

$$u(\nu, T) = \nu^3 g\left(\frac{\nu}{T}\right), \tag{1.2}$$

where g is a function that cannot be determined from thermodynamics. In order to specify this function one has to go beyond thermodynamical reasoning and use a more detailed theoretical approach. Finally Wien, Lord Rayleigh and J. Jeans managed to derive an expression for g that could explain the experimental data for higher frequencies quite well.

Planck tried to interpolate the two approximations of Wien and Rayleigh & Jeans. By guesswork he found a perfect fit to the experimental data, but he was confronted with the problem that he was lacking a theoretical derivation for this formula. Thirty-one years after this discovery Planck described this situation as follows:

> I can characterize the whole procedure as an act of desperation, since, by nature, I am peaceable and opposed to doubtful adventures. I had fought for six years with the problem [...] without arriving at any successful result. [...] I knew the formula describing the energy distribution [...] hence a theoretical interpretation had to be found at any price, however high it might be.

He made an assumption that might at first seem strange (and therefore at first was not accepted by the physicists of his time): He postulated that the energy for radiation with the frequency ν exists only in multiples of $h\nu$, where h is a constant of nature, the so called *Planck's constant*

$$h = 6.6260755 \cdot 10^{-34} Js. \tag{1.3}$$

According to this hypothesis energy is no longer a continuous quantity, but it consists of small quanta of energy $h\nu$, called *photons*. Planck thus arrived at the following expression for the energy per frequency interval $u(\nu)$:

$$\boxed{u(\nu) = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{\frac{h\nu}{k_B T}} - 1}}. \tag{1.4}$$

This formula fits strikingly well to the experimentally obtained curves. It looks similar to the Rayleigh-Jeans approximation, but the factor $[e^{\frac{h\nu}{k_B T}} - 1]^{-1}$ prevents the expression from diverging at higher frequencies (see figure 1.2).

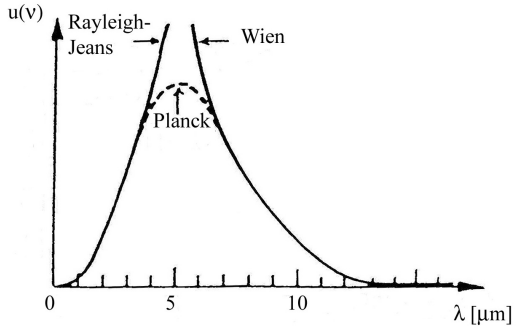


Figure 1.2: Comparison of the results for the spectrum of a blackbody according to Wien, Rayleigh-Jeans and Planck

Although Planck received a Nobel prize in 1918 for his ideas, his explanation of the spectrum of blackbody radiation did not take the world by storm at first. It seemed as if he had constructed a theory derived from experiment, but based on a hypothesis with no experimental basis.

1.2.2 The photoelectric effect

Five years later *Einstein* built on the ad hoc hypothesis of the quantization of energy to explain the phenomenon of the photoelectric effect. This effect was first observed by Hertz in 1887: If an alkali metal is irradiated by light with a frequency larger than a certain minimum frequency (which depends on the metal) electrons are emitted by this metal. It is interesting that the velocity of the electrons (and thus their energy) is only dependent on the frequency of the light beam hitting the metal, but not on its intensity. Classical physics is not able to explain the ν–proportionality of this effect. Assuming light to be an electromagnetic wave, the electrons of the metal should absorb an energy that is increasing with the intensity of the light beam until their velocity is high enough to overcome the potential well. According to this, we should be able to observe a delay between the start of the irradiation and the onset of the emission of electrons. This delay has not been measured until today, even though by now we would be able to do so (if it existed). Classical physics thus fails to explain this effect correctly.

Einstein took up the idea of Planck and even went a bit further. He assumed that light consisted of particles, called photons, with the energy $h\nu$. When one of these corpuscles encounters an electron of the metal, it is absorbed and the electron receives its energy $h\nu$ (at one instant). If this energy is large enough for the electron to overcome the potential of the atom,

it escapes. The energy of such an electron would be

$$\frac{1}{2}mv^2 = h\nu - W, \tag{1.5}$$

where W is the work needed to free an electron from the potential well. This theory is in complete accord with the experiment.

At this time the whole extent of the idea of energy or light quanta could not yet be perceived. Planck thought that his hypothesis was a mere complement to the theories known so far. Years later it became evident that they were in fact revolutionary. Nernst wrote in 1911:

> It appears that we find ourselves at present in the midst of an all-encompassing re-formulation of the principles on which the erstwhile kinetic theory of matter has been based.

Although Einstein himself contributed to the development of this new theory, he turned out to be a strict opponent to some of its consequences. In 1944 he wrote in a letter to Max Born:

> You believe in the God who plays dice, and I in complete law and order in a world which objectively exists, and which I, in a wildly speculative way, am trying to capture. I hope that someone will discover a more realistic way [...] than it has been my lot to find. Even the great initial success of Quantum Theory does not make me believe in the fundamental dice-game, although I am well aware that our younger colleagues interpret this as a consequence of senility. No doubt the day will come when we will see whose instinctive attitude was the correct one.

Einstein was appreciated for his work with a nobel prize in 1921.

1.2.3 Bohr's theory of the structure of atoms

At the end of the 19^{th} century Gustav Kirchhoff and Robert Bunsen examined the spectrum of gas atoms. If you energize a tube filled with gas of atoms of a certain kind, the gas begins to glow at a sufficient voltage. It emits a line spectrum, i.e. the emerging light has a discrete set of wavelengths. It turned out that every atom has a characteristic spectrum. The atomic number Z and the wavelengths of the spectrum are related by the *Rydberg-Ritz-formula*:

$$\boxed{\frac{1}{\lambda} = RZ^2\left(\frac{1}{m^2} - \frac{1}{n^2}\right)} \tag{1.6}$$

λ	...	wavelength of spectral line
R	...	Rydberg's constant, for big Z; $R_\infty = 10,97373\frac{\mu}{m}$
Z	...	atomic number
n,m	...	whole numbers with $n > m$

At first there was no theoretical explanation for this formula. In 1911 *Rutherford* and his coworkers Hans Geiger and Ernest Marsden deduced from scattering experiments of α-particles off a golden foil that the positive charge of the atom is cumulated in a small center, the nucleus. They imagined that the electrons move along circular or elliptical orbits around the nucleus, just like the planets move around the sun. Within the framework of classical physics, the moving electron would radiate (because its circular trajectory is equivalent to an accelerated movement) and thus loose energy until it would eventually fall into the nucleus within 10^{-8} seconds.

Many attempts were made to overcome these and similar difficulties without any significant success. Physicists tried to find a solution to this problem within the framework of the newly arisen quantum theory. It appeared natural to do so since the discrete lines in the spectra of atoms seemed to be related to the fact that the energy of an oscillator assumed values that were integral multiples of the energy packets $h\nu$. In 1913 a so far unknown physicist, *Niels Bohr*, who worked with Rutherford in Manchester and had therefore come to know his model of the atom, had an idea to avoid this 'disaster'. He set up two *postulates*:

- The electron moves around the nucleus in discrete circles according to classical mechanics. In these (stationary) states with energy E_n the atom does not radiate and the momentum is given by:
$$\oint p\,dr = nh \tag{1.7}$$
The line integral extends over the electron's orbit around the nucleus.

- When an atom undergoes a change from energy E_n to E_m it emits a photon with the energy
$$E = E_n - E_m \tag{1.8}$$
and correspondingly with the frequency
$$\nu = \frac{E_n - E_m}{h}. \tag{1.9}$$

Let us consider the first postulate in more detail. If the electron moves along a circular trajectory, the line integral is
$$2\pi r p = nh \tag{1.10}$$
or, with $p = \hbar k = \frac{h}{\lambda}$,
$$2\pi r = n\lambda. \tag{1.11}$$
The circumference of the electron's orbit thus is a multiple of the wavelength λ of the electron and the orbits are quantized. We will now calculate the radius and the energy for such an orbit.

The electron moves in a circular orbit around the nucleus. The centripetal force thus balances the Coulomb force between the electrons and the protons,

$$\frac{mv^2}{r} = \frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r^2}. \tag{1.12}$$

So the radius of the atom is

$$r = \frac{1}{4\pi\epsilon_0}\frac{Ze^2}{mv^2}. \tag{1.13}$$

With $\vec{p} = m\vec{v}$ we find

$$r = \frac{1}{4\pi\epsilon_0} m \frac{Ze^2}{p^2} \tag{1.14}$$

Using the above quantization rule,

$$p = \frac{nh}{2r\pi}, \tag{1.15}$$

the radius becomes

$$r_n = \frac{n^2}{Z}\frac{\epsilon_0 h^2}{me^2\pi} = \frac{n^2}{Z}a_0 \tag{1.16}$$

$$a_0 = \frac{\epsilon_0 h^2}{me^2\pi} \tag{1.17}$$

r_n ... radius of the electron's orbit, for $n = 1, 2, 3, ...$ different radii
a_0 ... *Bohr radius*

Each radius belongs to a certain energy E_n. The energy for an electron in an orbit with the radius r_n is

$$E_n = \underbrace{\frac{mv^2}{2}}_{E_{kin}} - \underbrace{\frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r_n}}_{E_{pot}} \tag{1.18}$$

Using equation (1.12) we find

$$mv^2 = \frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r_n}. \tag{1.19}$$

Inserting this and formula (1.16) into the expression for E_n we find

$$E_n = \frac{1}{8\pi\epsilon_0}\frac{Ze^2}{r_n} - \frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r_n} = -\frac{1}{8\pi\epsilon_0}\frac{Ze^2}{r_n}, \tag{1.20}$$

$$E_n = -\frac{me^4}{8\epsilon_0^2 h^2}\frac{Z^2}{n^2}. \tag{1.21}$$

Let us now return to the initial problem: the spectrum emitted by atoms and the Rydberg-Ritz formula (1.6). If an electron falls from the energy level E_n to a lower level E_m it emits a photon with a wavelength λ corresponding to $E_n - E_m$. According to (1.21):

$$\frac{hc}{\lambda} = \Delta E = E_n - E_m = \frac{me^4}{8\epsilon_0^2 h^2} Z^2 \left(\frac{1}{m^2} - \frac{1}{n^2}\right) \tag{1.22}$$

So we end up with formula (1.6):

$$\frac{1}{\lambda} = \frac{me^4}{8\epsilon_0^2 h^3 c} Z^2 \left(\frac{1}{m^2} - \frac{1}{n^2} \right) = RZ^2 \left(\frac{1}{m^2} - \frac{1}{n^2} \right) \tag{1.23}$$

$R = \frac{me^4}{8\epsilon_0^2 h^3 c}$ ... *Rydberg's constant*

We thus find the following picture of the structure of an atom:

- The *bound electrons* of an atom move along circular orbits with different radii. The radii are quantized and correspond to *discrete energy values*. These values are all *negative*.
- There is a *minimum energy* $E_0 = -\frac{me^4}{8\epsilon_0^2 h^2} Z^2$ (formula (1.21) with $n = 1$), the *ground state* of the atom. If an electron is excited to a higher energy level ($n = 2, 3, 4 \ldots$), it always returns to an energy as low as possible, whereby it emits light of a certain frequency.
- For $r_n \to \infty$ the energy of an electron becomes $\lim_{n\to\infty} E_n = 0$. For $E > 0$ the atom is ionized and *all (continuous) values of the energy* are allowed.

Many years later, Werner Heisenberg recalled the work on the development of the atomic model:

> I remember discussions with Bohr which went through many hours till very late at night and ended almost in despair; and when at the end of the discussion I went alone for a walk in the neighbouring park I repeated to myself again and again the question: Can nature possibly be so absurd as it seemed to us in these atomic experiments?

Niels Bohr was awarded the nobel prize in 1922.

1.2.4 The Compton effect

The Compton effect also confirms the photon theory. Consider free electrons irradiated by x-rays (see figure 1.3). One observes that the wavelength of the incoming x-rays is different from the wavelength of the outgoing ones.

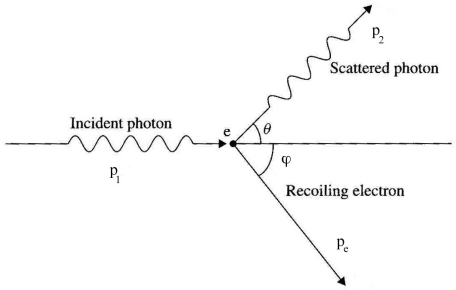


Figure 1.3: The experimental setup for the Compton effect

$$\lambda_{in} \neq \lambda_{out} \tag{1.24}$$

The difference $\Delta\lambda$ is related to the angle θ between the direction of propagation of the x-rays and of the scattered beam according to

$$\boxed{\Delta\lambda = 2\frac{h}{mc}sin^2\frac{\theta}{2}} \tag{1.25}$$

It is not possible to understand the shift of the wavelength of the radiation from a classical point of view. If we regard the x-rays as waves, the electrons should absorb energy and then re-emit radiation of the same wavelength λ. So, what is the origin of this $\Delta\lambda$?

Compton managed to explain this effect using the idea of photons. The irradiation of the electrons can thus be understood as an elastic collision between a photon and an electron. The photon loses energy to the electron and, since its wavelength is inversely proportional to the energy, it has to increase.

Since photons travel at the speed of light their energy and momentum are related by the relativistic formula $E^2 = m_0^2c^4 + p^2c^2$ with rest mass $m_0 = 0$, i.e. $|p| = E/c$. The *Planck-Einstein relation* $E = h\nu$ and the relation between frequency ν and wave vector $\vec{k}$ in vacuum thus imply

$$E = h\nu = \hbar\omega, \tag{1.26}$$

$$\vec{p} = \hbar\vec{k}. \tag{1.27}$$

Considering the elastic collision of a photon with an electron we can use the conservation of *momentum*

$$\vec{p_1} = \vec{p_2} + \vec{p_e}, \tag{1.28}$$

or

$$\hbar\vec{k_1} = \hbar\vec{k_2} + \vec{p_e} \tag{1.29}$$

$\vec{p}_1$,$\vec{k}_1$	...	momentum, wave vector before the impact
$\vec{p}_2$,$\vec{k}_2$	...	momentum, wave vector after the impact
$\vec{p}_e$	...	momentum of the electron

and the conservation of *energy*

$$\underbrace{p_1 c}_{\text{moving photon}} + \overbrace{m_e c^2}^{\text{resting electron}} = \underbrace{p_2 c}_{\text{moving photon}} + \overbrace{\sqrt{p_e^2 c^2 + m_e^2 c^4}}^{\text{moving electron}} \tag{1.30}$$

or, with $pc = E = \hbar\omega$ and $\omega = kc$

$$\hbar k_1 + m_e c = \hbar k_2 + \sqrt{p_e^2 + m_e^2 c^2} \tag{1.31}$$

Combining (1.29) and (1.31) and eliminating $\vec{p}_e$, where the scalar product of $\vec{k}_1$ and $\vec{k}_2$ is

$$\vec{k}_1 \vec{k}_2 = k_1 k_2 cos\theta \tag{1.32}$$

with θ being the angle between $\vec{k}_1$ and $\vec{k}_2$, we finally end up with formula (1.25).

1.2.5 Interference phenomena

So far, we have considered situations of electromagnetic waves behaving in a corpuscular manner. We have come to the conclusion that it is problematic to describe some phenomena in a classical way. In the following we will see that the new corpuscular theory is insufficient too and that a combination of wave and particle aspects of matter is needed.

Problems with the newly introduced photon theory arise when we observe phenomena such as diffraction or interference. Is there a way to find an explanation for these things based upon the photon theory? Consider *Young's double-slit experiment* (see figure 1.4), in which light falls on a wall with two slits. Behind that wall there is a detector like a photographic plate in order to observe the interference pattern that is produced by the wall. The blackening of the photographic plate is proportional to the distribution of the light intensity.

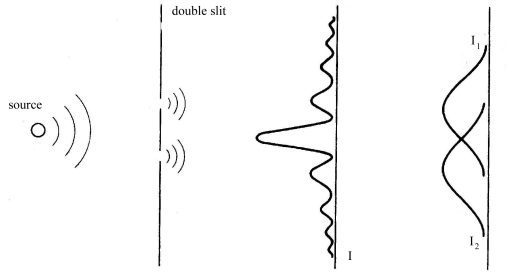


Figure 1.4: Young'd double slit experiment

The two beams produced by slit one and slit two interfere and thus the total intensity on the screen depends on the phase between the two beams. If these beams are represented by the two wave functions

$$\psi_1 = |\psi_1| e^{i\varphi_1} \tag{1.33}$$

$$\psi_2 = |\psi_2| e^{i\varphi_2} \tag{1.34}$$

where φ_1 and φ_2 are the phases of the two waves, and thus functions of $(\vec{r}, t)$, the overall intensity on the photographic plate is

$$I = |\psi|^2 = |\psi_1 + \psi_2|^2 = |\psi_1|^2 + |\psi_2|^2 + \underbrace{|\psi_1 \psi_2|[e^{i(\varphi_1 - \varphi_2)} + e^{i(\varphi_2 - \varphi_1)}]}_{\text{interference term}}, \tag{1.35}$$

which is not only the sum of the two intensities I_1 and I_2,

$$I \neq I_1 + I_2 = |\psi_1|^2 + |\psi_2|^2. \tag{1.36}$$

One could try to explain this result with the interaction of the photons that passed through slit one and those that passed through slit two. If we diminish the intensity of the light beam that falls on the wall and increase the exposure time so that the overall amount of photons that are detected on the plate behind the wall remains the same, the photons eventually pass the two slits one after another and thus cannot interact. But the interference pattern on the photographic plate is found to stay the same!

It seems as if in this case the wave-aspects of light would dominate. But if we diminish the intensity of the light beam and keep the exposure time short, we are still able to detect localized impacts on the photographic plate, i.e. single photons. Here the wave theory is insufficient. On the other hand, even if these photons pass the double slit one by one (without possible interaction) they still generate the interference pattern. The result of this experiment leads to a paradox: As mentioned before the intensity distribution of a double slit is *not* simply the sum of two single slits. Although a photon is far too small to "know" whether there is a second slit or not, it nevertheless seems to be aware of it and moves accordingly. While all photons are emitted under essentially the same conditions, their trajectories are different. The initial state of a system thus no longer determines its evolution in time. There is only a *statistical probability* for different locations (for example, photons are more likely to hit the photographic plate at a maximum of the intensity of the interference pattern than at a minimum).

Chapter 2

Wave Mechanics and the Schrödinger equation

Falls es bei dieser verdammten Quantenspringerei bleiben sollte, so bedauere ich, mich jemals mit der Quantentheorie beschäftigt zu haben!
-Erwin Schrödinger

In this chapter we introduce the Schrödinger equation and its probabilistic interpretation. We then discuss some basic physical phenomena like wave packets, discrete bound state energies and scattering on the basis of one-dimensional examples.

2.1 The Schrödinger equation

Schrödingers wave mechanics originates in the work of Louis de Broglie on matter waves. De Broglie postulated that all material particles can have corpuscular as well as wavelike aspects and that the correspondence between the dynamical variables of the particle and the characteristic quantities of the associated wave,

$$E = \hbar\omega, \qquad \text{and} \qquad \vec{p} = \hbar\vec{k}, \tag{2.1}$$

which was established for photons by the Compton effect, continues to hold for all *matter waves*. Schrödinger extended these ideas and suggested that the dynamical state of a quantum system is completely described by a wave function ψ satisfying a homogeneous linear differential equation (so that different solutions can be superimposed, which is a typical property of waves). In particular, we can express ψ as a continuous superposition of plane waves,

$$\psi(\vec{x}, t) = \int d^3k \; f(\vec{k})\, e^{i(\vec{k}\vec{x} - \omega(k)t)}. \tag{2.2}$$

For the plane waves $e^{i(\vec{k}\vec{x}-\omega t)}$ the relation (2.1) suggests the correspondence rule

$$\boxed{E \to i\hbar\frac{\partial}{\partial t}, \qquad \vec{p} \to \frac{\hbar}{i}\vec{\nabla}.} \tag{2.3}$$

Energy and momentum of a free classical particle are related by $E = p^2/2m$. When a particle moves in a potential $V(x)$ its conserved energy is given by the Hamilton function $H(x,p) = \frac{p^2}{2m} + V(x)$. Setting $E\psi = H\psi$ with $E \to i\hbar\partial_t$ and $\vec{p} \to \frac{\hbar}{i}\vec{\nabla}$ we arrive at the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi(x,t) = H\psi(x,t) \qquad \text{with} \qquad H = -\frac{\hbar^2}{2m}\Delta + V(x), \tag{2.4}$$

where $\Delta = \vec{\nabla}^2$ is the Laplace operator and $V = e\phi$ for an electron moving in an electric field $\vec{E}(x) = -\text{grad}\,\phi(x)$.

More generally, a classical point particle with mass m and charge e moving in an electromagnetic field

$$\vec{E} = -\vec{\nabla}\phi - \frac{1}{c}\partial_t\vec{A}, \qquad \vec{B} = \vec{\nabla}\times\vec{A} \tag{2.5}$$

with gauge potential $A^\mu = (\phi, \vec{A})$ feels a Lorentz force $\vec{F} = e(\vec{E} + \frac{1}{c}\,\vec{v}\times\vec{B})$. The Hamilton function describing this dynamics is[1]

$$H(x,p;t) = \frac{1}{2m}(\vec{p} - \frac{e}{c}\vec{A}(\vec{x},t))^2 + e\phi(\vec{x},t). \tag{2.7}$$

With the correspondence rule (2.3) we thus find the general Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi = \left[\frac{1}{2m}\left(\frac{\hbar}{i}\vec{\nabla} - \frac{e}{c}\vec{A}\right)^2 + e\phi\right]\psi, \tag{2.8}$$

which describes the motion of a quantum mechanical scalar point particle in a classical external electromagnetic field. This is an approximation in several respects. First we have neglected the spin of elementary point particles like electrons, which we will discuss in chapter 5. In chapter 7 we will discuss the Dirac equation, which is the relativistic generalization of the Schrödinger equation. The relativistic treatment is necessary for a proper understanding of the magnetic interactions, and hence of the fine structure of the energy levels of hydrogen, and it will lead to the prediction of anti-matter. Eventually we should note that also the environment, including

[1] In order to derive the Lorentz force from this Hamiltonian we consider the canonical equations of motion

$$\dot{x}_i = \frac{\partial H}{\partial p_i} = \frac{p_i - \frac{e}{c}A_i}{m}, \qquad \dot{p}_j = -\frac{\partial H}{\partial x_j} = \frac{e}{c}\frac{\partial A_i}{\partial x_j}\frac{p_i - \frac{e}{c}A_i}{m} - e\frac{\partial\phi}{\partial x_j} = \frac{e}{c}(\partial_j A_i)\dot{x}_i - e\partial_j\phi, \tag{2.6}$$

which imply $\vec{F} = m\ddot{\vec{x}} = \frac{d}{dt}(p_i - \frac{e}{c}A_i) = \dot{\vec{p}} - \frac{e}{c}(\partial_t + \dot{x}_i\partial_i)\vec{A} = \frac{e}{c}(v_i\vec{\nabla}A_i - v_i\partial_i\vec{A}) - e(\frac{1}{c}\dot{\vec{A}} + \vec{\nabla}\phi) = \frac{e}{c}\vec{v}\times\vec{B} + e\vec{E}$. Note that the relation between the *canonical* momentum $p_j = m\dot{x}_j + \frac{e}{c}A_j$ and the velocity $\vec{v} = \dot{\vec{x}}$ depends on the gauge-dependent vector potential $\vec{A}$. The gauge-independent quantity $\vec{\pi} = m\dot{\vec{x}} = \vec{p} - \frac{e}{c}\vec{A}$ is sometimes called *physical* or *mechanical* momentum. According to the general quantization rule (see below) the operator $\frac{\hbar}{i}\vec{\nabla}$ has to replace the *canonical* momentum.

the electromagnetic field, consists of quantum systems. This leads to the "second quantization" of quantum field theory. First, however, we restrict our attention to the quantum mechanical description of a single non-relativistic point particle in a classical environment.

It is an important and surprising property of the Schrödinger equation that it explicitly depends on the electromagnetic potentials A^μ, which are unobservable and whose values depend on the coice of a gauge. This is in contrast to classical physics, where the Lorentz force is a function of the gauge invariant field strengths. A straightforward calculation shows that a gauge transformation

$$\phi \to \phi' = \phi - \frac{1}{c}\frac{\partial}{\partial t}\Lambda, \qquad A \to A' = A + \vec{\partial}\Lambda \tag{2.9}$$

of the scalar and vector potentials, which leaves the observable fields $\vec{E}$ and $\vec{B}$ invariant for an arbitrary function $\Lambda(t, \vec{x})$, can be compensated by an space- and time-dependent phase rotation of the wave function[2]

$$\psi \to \psi' = e^{\frac{ie}{\hbar c}\Lambda}\psi, \tag{2.10}$$

i.e. if ψ solves the Schrödinger equation (2.8) then ψ' solves the same equation for potentials ϕ' and $\vec{A}'$. Since the phase of the wave function ψ can be changed arbitrarily by such a gauge transformation we might expect that only its modulus $|\psi(t, x)|$ is observable. This conclusion is indeed consistent with the physical interpretation of the wave function that was suggested by Max Born in 1927: $|\psi|^2(x) = (\psi^*\psi)(x)$ is the probability density for finding an electron with wave function $\psi(x)$ at a position $x \in \mathbb{R}^3$. It is a perplexing but characteristic feature of quantum physics that a local description of particle interactions requires the introduction of mathematical objects like gauge potentials $(\phi, \vec{A})$ and complex wave functions ψ that are not directly observable and only certain functions of which can be related to "the real world".[3]

2.1.1 Probability density and probability current density

Born's interpretation of the wave function $\psi(\vec{x}, t)$ implies that the integral over the probability density, i.e. the total probability to find the electron somewhere in space, has to be one:

$$\int d^3x\ \rho(\vec{x}, t) = 1, \qquad \text{with} \qquad \rho(\vec{x}, t) = |\psi(\vec{x}, t)|^2. \tag{2.11}$$

This fixes the normalization of the wave function, which is also called *probability amplitude*, at some initial time up to a phase. Consistency of the interpretation requires that the total probability stays one under time evolution. To check this we compute the time derivative of ρ

[2] This follows from $(\frac{\hbar}{i}\vec{\nabla} - \frac{e}{c}\vec{A}')e^{\frac{ie}{\hbar c}\Lambda} = e^{\frac{ie}{\hbar c}\Lambda}(\frac{\hbar}{i}\vec{\nabla} - \frac{e}{c}\vec{A})$ and $(i\hbar\partial_t - e\phi')e^{\frac{ie}{\hbar c}\Lambda} = e^{\frac{ie}{\hbar c}\Lambda}(i\hbar\partial_t - e\phi)$.

[3] For the electromagnetic potentials this necessity manifests itself in the Aharonov-Bohm effect, which predicts an "action at a distance" of a magnetic field on interference patterns of electrons (see below). This effect was predicted in 1959 and first confirmed experimentally in 1960 [Schwabl].

for a solution of the Schrödinger equation. With $\frac{\hbar}{i}\vec{\nabla} - \frac{e}{c}\vec{A} = \frac{\hbar}{i}(\vec{\nabla} - ig\vec{A})$ for $g = e/(\hbar c)$ and the anti-commutator $\{\vec{\nabla}, \vec{A}\} \equiv \vec{\nabla}\vec{A} + \vec{A}\vec{\nabla} = (\vec{\nabla}\vec{A}) + 2\vec{A}\vec{\nabla}$ we find

$$\begin{aligned}
\dot{\rho}(\vec{x},t) &= \dot{\psi}^*\psi + \psi^*\dot{\psi} = (\frac{1}{i\hbar}H\psi)^*\psi + \psi^*\frac{1}{i\hbar}H\psi \\
&= \frac{1}{i\hbar}\frac{\hbar^2}{2m}\Big(\psi(\vec{\nabla} + ig\vec{A})^2\psi^* - \psi^*(\vec{\nabla} - ig\vec{A})^2\psi\Big) \\
&= \frac{\hbar}{2im}\Big(\psi(\Delta + ig\{\vec{\nabla}, \vec{A}\} - g^2\vec{A}^2)\psi^* - \psi^*(\Delta - ig\{\vec{\nabla}, \vec{A}\} - g^2\vec{A}^2)\psi\Big) \\
&= \frac{\hbar}{2im}\Big(\psi\Delta\psi^* - \psi^*\Delta\psi + 2ig\Big((\vec{\nabla}\vec{A})\psi^*\psi + \psi\vec{A}\vec{\nabla}\psi^* + \psi^*\vec{A}\vec{\nabla}\psi\Big)\Big) \\
&= -\vec{\nabla}\left(\frac{\hbar}{2im}(\psi^*\vec{\nabla}\psi - \psi\vec{\nabla}\psi^*) - \frac{e}{mc}\vec{A}\psi^*\psi\right) \qquad (2.12)
\end{aligned}$$

We thus obtain a continuity equation (similar to the one we know for incompressible fluids)

$$\frac{\partial}{\partial t}\rho(\vec{x},t) + \vec{\nabla}\vec{j}(\vec{x},t) = 0 \qquad (2.13)$$

with the *probability current density*

$$\boxed{\vec{j}(\vec{x},t) = \frac{\hbar}{2im}(\psi^*\vec{\nabla}\psi - (\vec{\nabla}\psi^*)\psi) - \frac{e}{mc}\vec{A}\psi^*\psi} \qquad (2.14)$$

(It is instructive to compare this formula with the classical particle current $\dot{\vec{x}} = \frac{1}{m}(\vec{p} - \frac{e}{c}\vec{A})$.) By Gauss' theorem, the change in time of the probability to find the particle in a finite volume V equals the flow of the probability current density through the bounding surface ∂V of that domain,

$$\frac{\partial}{\partial t}\int\limits_V \rho(\vec{x},t)d^3x = -\int\limits_V \vec{\nabla}\vec{j}(\vec{x},t)d^3x = -\oint\limits_{\partial V} \vec{j}(\vec{x},t)d\vec{f} \qquad (2.15)$$

Normalizability of ψ implies that the fields fall off at infinity so that the surface integral vanishes as $V \to \mathbb{R}^3$. This establishes conservation of the total probability $\int\limits_{\mathbb{R}3} d^3x\,\rho(x) = 1$ for all times.

2.1.2 Axioms of quantum theory

In order to gain some intuition for the physical meaning of the Schrödinger equation we next work out its solutions for a number of simple one-dimensional examples. Before going into the details of the necessary calculations we list here, for later reference, the basic assumptions of quantum mechanics:

1. The state of a quantum system is *completely* determined by a wave function $\psi(x)$.

2. Observables correspond to self-adjoint operators A (these can be diagonalized and have real eigenvalues).

3. Expectation values of observables (i.e. mean values for repeated measurements of A in the same quantum state) are given by the "scalar product" $\langle A\rangle = \langle\psi|A\psi\rangle = \int \psi^* A\psi$.

4. The time evolution of the system is determined by the Schrödinger equation $i\hbar\frac{\partial\psi}{\partial t} = H\psi$.

5. When the measurement of an observable A yields an eigenvalue a_n then the wave function immediately turns into the corresponding eigenfunction ψ_n of A (this is called collapse of the wave function).

It can be shown that axioms 2 and 3 imply that the result of the measurement of an observable A can only be an eigenvalue a_n of that operator and that the probability for measuring a_n is given by $|c_n|^2$, where c_n is the coefficient of the eigenfunction ψ_n in the expansion $\psi = \sum c_n\psi_n$. In particular, this will imply Born's probability density interpretation of $|\psi(x)|^2$.

2.1.3 Spreading of free wave packets and uncertainty relation

The position and the momentum of a quantum mechanical particle are described by the linear operators

$$\vec{X}\psi(x) = \vec{x}\psi(x) \qquad \text{and} \qquad \vec{P}\psi(x) = \frac{\hbar}{i}\vec{\nabla}\psi(x), \tag{2.16}$$

respectively. The uncertainty ΔA of a measurement of an observable A in a state ψ is defined as the square root of its mean squared deviation from its expectation value,

$$(\Delta A)^2 = \langle\psi|\ (A - \langle A\rangle_\psi)^2\,|\psi\rangle = \langle A^2\rangle_\psi - (\langle A\rangle_\psi)^2, \tag{2.17}$$

where $\langle A\rangle_\psi = \langle\psi|A|\psi\rangle = \int d^3x\,\psi^* A\psi$ denotes the expectation value of A in the state $\psi(x)$ and $A^2\psi = A(A(\psi))$; to be more precise, within the expectation value the number $\langle A\rangle_\psi$ is identified with that number times the unit operator.

For a free particle it can be shown that the uncertainty ΔX of the position increases at late times, i.e. that the wave packets describing localized free particles delocalize and spread out. We now illustrate this phenomenon for a Gaussian wave packet and consider the time evolution of the wave function of a free particle in one dimension, which satisfies the Schrödinger equation with vanishing potential

$$i\hbar\dot{\psi} = -\frac{\hbar^2}{2m}\psi'' \tag{2.18}$$

Since the Fourier transform $\tilde{\psi}(k)$ of a Gaussian distribution is again Gaussian we start with a Fourier integral

$$\psi(x,0) = \frac{1}{\sqrt{2\pi}}\int dk\ e^{ikx}\,\tilde{\psi}(k) \tag{2.19}$$

with

$$\tilde{\psi}(k) = \alpha\, e^{-d^2(k-k_0)^2}, \tag{2.20}$$

so that the wave numbers are centered about k_0 with width $1/d$. The normalization constant α will be determined later. Since plane waves $e^{i(kx-\omega t)}$ satisfy the free Schrödinger equation (2.18) if $\omega = \omega(k) = \hbar k^2/(2m)$ we can directly write down the solution for arbitrary times as a Fourier integral

$$\psi(x,t) = \frac{1}{\sqrt{2\pi}} \int dk\ \tilde{\psi}(k) e^{i(kx-\omega t)} = \frac{\alpha}{\sqrt{2\pi}} \int dk\ e^{i(kx-\frac{\hbar k^2}{2m}t)}\ e^{-(k-k_0)^2\, d^2} \tag{2.21}$$

In order to evaluate this integral we bring the exponent into a quadratic form

$$\psi(x,t) = \frac{\alpha}{\sqrt{2\pi}} \int dk \quad e^{-ak^2+2bk-c}, \tag{2.22}$$

where we introduced the combinations

$$a = d^2 + \frac{i\hbar t}{2m}, \qquad b = k_0 d^2 + \frac{ix}{2}, \qquad c = k_0^2 d^2. \tag{2.23}$$

Due to the exponential falloff of the integrand the integration path $-\infty < k < \infty$ can be shifted in the complex plane by the imaginary part of b/a and rotated by the argument of $\sqrt{a}$ without picking up a contribution from the arcs at infinity. Introducing the new integration variable $\kappa = \sqrt{a}(k - \frac{b}{a})$ we can thus again integrate over the real axis and find

$$\psi(x,t) = \frac{\alpha}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk\ e^{-a\left(k-\frac{b}{a}\right)^2+\frac{b^2}{a}-c} = \frac{\alpha}{\sqrt{2\pi}}\ e^{\frac{b^2}{a}-c} \int_{-\infty}^{\infty} \frac{d\kappa}{\sqrt{a}}\ e^{-\kappa^2} = \frac{\alpha}{\sqrt{2\pi}} \cdot e^{\frac{b^2}{a}-c} \sqrt{\frac{\pi}{a}}. \tag{2.24}$$

For the probability density we obtain

$$|\psi(x,t)|^2 = \frac{|\alpha|^2}{2|a|} \cdot e^{2Re\left(\frac{b^2-ac}{a}\right)} = \frac{|\alpha|^2}{2d^2\sqrt{(1+T^2)}} \cdot e^{\frac{-(x-v_0 t)^2}{2d^2(1+T^2)}}, \tag{2.25}$$

where we introduced the velocity v_0 and a rescaled time T as

$$\boxed{v_0 = \frac{\hbar k_0}{m}, \qquad T = \frac{\hbar t}{2md^2}.} \tag{2.26}$$

As expected, the integrated probability density

$$\int dx\ |\psi(x,t)|^2 = \frac{|\alpha|^2}{2d^2\sqrt{(1+T^2)}} \sqrt{2\pi d^2(1+T^2)} = \frac{|\alpha|^2}{d}\sqrt{\frac{\pi}{2}} \tag{2.27}$$

becomes time independent and we find the normalization constant

$$|\alpha|^2 = \sqrt{\frac{2}{\pi}} d. \tag{2.28}$$

$v_0 = \hbar k_0/m$ is the group velocity of the wave packet and for large times $t \gg 2md^2/\hbar$ the width of the wave packet in *position space* becomes proportional to $d\,T = t\hbar/(2md)$ as shown in figure 2.1. Inserting the expressions eq. (2.23) we find the explicit form

$$\psi(x,t) = \frac{\sqrt{1-iT}}{\sqrt{2\pi d(1+T^2)}} \cdot e^{\frac{-(x-v_0 t)^2+i(Tx^2+4xk_0d^2-4Tk_0^2d^4)}{4d^2(1+T^2)}} \tag{2.29}$$

for the solution to the Schrödinger equation with initial data $\psi(x,0)$.

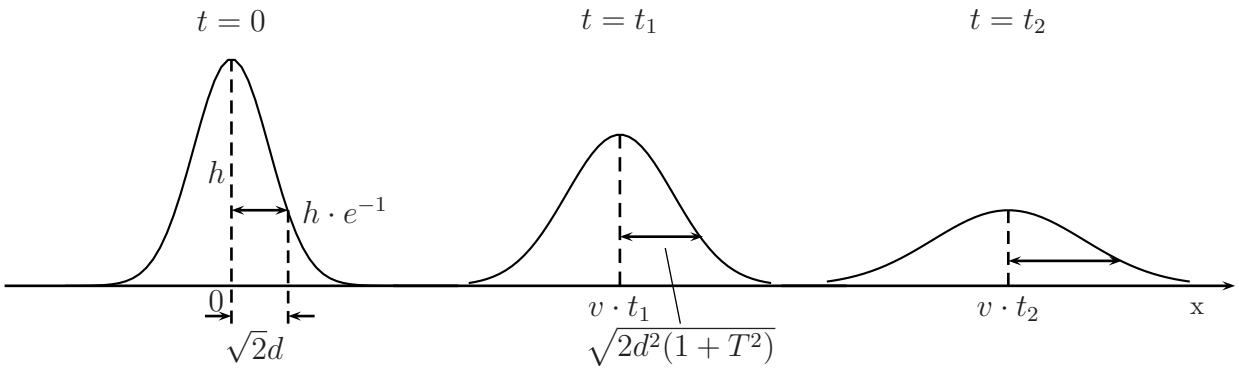


Figure 2.1: Schematic graph of the delocalization of a Gaussian wave packet

Heisenberg's uncertainty relation for position and momentum

In chapter 3 we will derive the general form of Heisenberg's uncertainty relation which, when specialized to position and momentum, reads $\Delta X\,\Delta P \geq \frac{1}{2}\hbar$. Here we check that this inequality is satisfied for our special solution. We first compute the expectation values that enter the uncertainty $\Delta X^2 = \langle (x - \langle x\rangle)^2\rangle$ of the position.

$$\langle x\rangle = \int x|\psi(x,t)|^2 dx = \int (x - v_0 t)|\psi(x,t)|^2 dx + v_0 t \int |\psi(x,t)|^2 dx = v_0 t. \tag{2.30}$$

The first integral on the r.h.s. is equal to zero because the integration domain is symmetric in $x' = x - v_0 t$ and ψ is an even function of x' so that its product with x' is odd. The second integral has been normalized to one. For the uncertainty we find

$$(\Delta x)^2 = \langle (x - \langle x\rangle)^2\rangle = \int (x - v_0 t)^2 |\psi(x,t)|^2 dx = d^2(1 + T^2), \tag{2.31}$$

where we have used

$$\int_{-\infty}^{+\infty} x^2 e^{-bx^2} dx = -\frac{\partial}{\partial b}\int_{-\infty}^{+\infty} e^{-bx^2} dx = -\frac{\partial}{\partial b}\sqrt{\frac{\pi}{b}} = \sqrt{\frac{\pi}{b}}\frac{1}{2b}, \tag{2.32}$$

i.e. the expectation value of x^2 in a normalized Gaussian integral, as in eq. (2.25), is $\frac{1}{2}$ times the inverse coefficient of $-x^2$ in the exponent.

The uncertainty of the momentum can be computed similarly in terms of the Fourier transform of the wave function since $P = \frac{\hbar}{i}\partial_x = \hbar k$ in the integral representation. For $n = 0, 1, 2, \ldots$

$$\int dx\ \psi^* P^n \psi = \int dx \iint \frac{dk\,dk'}{2\pi}\ e^{-i(k'x - \omega' t)}\tilde{\psi}^*(k)(\hbar k)^n e^{i(kx - \omega t)}\tilde{\psi}(k) = \int dk\,|\tilde{\psi}(k)|^2(\hbar k)^n, \tag{2.33}$$

where $\int dx\, e^{ix(k-k')} = 2\pi\delta(k - k')$ was used to perform the k' integration. Like above, symmetric integration therefore implies $\langle P\rangle = \hbar\langle k\rangle = \hbar k_0$, and by differentiation with respect to the

coefficient of $-k^2$ in the exponent of $|\tilde{\psi}(k)|^2$ we find

$$\frac{1}{\hbar^2}(\Delta P)^2 = (\Delta k)^2 = \langle (k-k_0)^2 \rangle = \int (k-k_0)^2 |\tilde{\psi}(k)|^2 = (4d^2)^{-1}. \tag{2.34}$$

The product of the uncertainties is

$$\Delta X \Delta P = \hbar \Delta x \Delta k = \frac{\hbar}{2}\sqrt{1+T^2} \tag{2.35}$$

which assumes its minimum at the initial time $t = 0$. Hence

$$\boxed{\Delta X \Delta P \geq \frac{\hbar}{2}}. \tag{2.36}$$

Relation (2.36) is known as the *Heisenberg uncertainty relation* and, for this special case, it predicts that one cannot measure position and momentum of a particle at the same time with arbitrary precision. In chapter 3 we will derive the general form of the uncertainty relations for arbitrary pairs of observables and for arbitrary states.

2.2 The time-independent Schrödinger equation

If the Hamiltonian does not explicitly depend on time we can make a separation ansatz

$$\Psi(\vec{x}, t) = u(\vec{x})v(t). \tag{2.37}$$

The Schrödinger equation now reads

$$v(t)\left(-\frac{\hbar^2}{2m}\Delta + V(\vec{x})\right)u(\vec{x}) = u(\vec{x})\; i\hbar\frac{\partial}{\partial t}v(t). \tag{2.38}$$

$u(\vec{x})$ and $v(t)$ cannot vanish identically, and except for isolated zeros of these functions we can divide by their product,

$$\frac{1}{u(\vec{x})}\left(-\frac{\hbar^2}{2m}\Delta + V(\vec{x})\right)u(\vec{x}) = \frac{1}{v(t)}\left(i\hbar\frac{\partial v(t)}{\partial t}\right) = E. \tag{2.39}$$

The left hand side $(Hu)/u$ depends only on $\vec{x}$ and the right hand side $i\hbar\dot{v}/v$ only on t, therefore both sides of this equation must be equal to a constant E. We thus obtain two separate eigenvalue equations:

$$\left[-\frac{\hbar^2}{2m}\Delta + V(\vec{x})\right]u(\vec{x}) = Eu(\vec{x}) \tag{2.40}$$

and

$$i\hbar\frac{\partial}{\partial t}v(t) = Ev(t). \tag{2.41}$$

Equation (2.40) is known as the *time-independent* or *stationary* Schrödinger equation. Up to a constant factor, which is absorbed into a redefinition of $u(x)$, the unique solution to (2.41) is

$$v(t) = e^{-\frac{i}{\hbar}Et} = e^{-i\omega t} \tag{2.42}$$

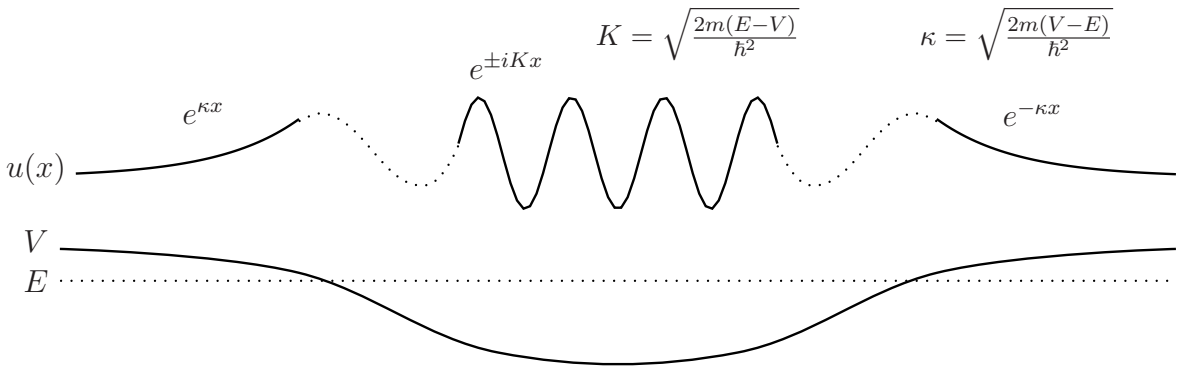


Figure 2.2: Bound state solutions for the stationary Schrödinger equation.

with the Einstein relation $E = \hbar\omega$. The stationary solutions $\psi(x,t)$ to the Schrödinger equation thus have the form

$$\psi(\vec{x},t) = u(\vec{x})e^{-i\omega t}. \tag{2.43}$$

Their time dependence is a pure phase so that probability densities are time independent.

In order to get an idea of the form of the wave function $u(x)$ we consider a slowly varying and asymptotically constant attractive potential as shown in figure 2.2. Since the stationary Schrödinger equation in one dimension

$$-\frac{\hbar^2}{2m}\,u''(x) = (E - V(x))\,u(x) \tag{2.44}$$

is a second order differential equation it has two linearly independent solutions, which for a slowly varying $V(x)$ are (locally) approximately exponential functions

$$u(x) \approx \begin{cases} Ae^{iKx} + Be^{-iKx} = A'\sin(Kx) + B'\cos(Kx), & K = \sqrt{\frac{2m(E-V)}{\hbar^2}} \quad \text{for } E > V, \\ Ce^{\kappa x} + De^{-\kappa x} = C'\sinh(\kappa x) + D'\cosh(\kappa x), & \kappa = \sqrt{\frac{2m(V-E)}{\hbar^2}} \quad \text{for } E < V. \end{cases} \tag{2.45}$$

In the classically allowed realm, where the energy E of the electron is larger then the potential, the solution is oscillatory, whereas in the classically forbidden realm of $E < V(x)$ we find a superposition of exponential growth and of exponential decay. Normalizability of the solution requires that the coefficient C of exponential growth for $x \to \infty$ and the coefficient D of exponential decay for $x \to -\infty$ vanish. If we require normalizability for negative x and increase the energy, then the wave function will oscillate with smaller wavelength in the classically allowed domain, leading to a component of exponential growth of $u(x)$ for $x \to \infty$, until we reach the next energy level for which a normalizable solution exists. We thus find a sequence of wave functions $u_n(x)$ with energy eigenvalues $E_1 < E_2 < \ldots$, where $u_n(x)$ has $n-1$ nodes (zeros). The normalizable eigenfunctions u_n are the wave functions of bound states with a discrete spectrum of energy levels E_n.

It is clear that bound states should exist only for $V_{min} < E < V_{max}$. The lower bound follows because otherwise the wave function is convex, and hence cannot be normalizable.

These bounds already hold in classical physics. In quantum mechanics we will see that the energy can be bounded from below even if $V_{min} = -\infty$ (like for the hydrogen atom). We also observe that *in one dimension* the energy eigenvalues are nondegenerate, i.e. for each E_n any two eigenfunctions are proportional (the vector space of eigenfunctions with eigenvalue E_n is one-dimensional). Normalization of the integrated probability density moreover fixes $u_n(x)$ up to a phase factor (i.e. a complex number ρ with modulus $|\rho| = 1$). Since the differential equation (2.40) has real coefficients, real and imaginary parts of every solution are again solutions. The bound state eigenfunctions $u(x)$ can therefore be chosen to be real.

Parity is the operation that reverses the sign of all space coordinates. If the Hamilton operator is invariant under this operation, i.e. if $H(-\vec{x}) = H(\vec{x})$ and hence the potential is symmetric $V(-\vec{x}) = V(\vec{x})$, then the $u(-\vec{x})$ is an eigenfunction for an eigenvalue E whenever $u(\vec{x})$ has that property because $(H(\vec{x}) - E)u(\vec{x}) = 0$ implies $(H(\vec{x}) - E)u(-\vec{x}) = (H(-\vec{x}) - E)u(-\vec{x}) = 0$. But every function u can be written as the sum of its even part u_+ and its odd part u_-,

$$u(\vec{x}) = u_+(\vec{x}) + u_-(\vec{x}) \qquad \text{with} \qquad u_\pm(\vec{x}) = \frac{1}{2}(u(\vec{x}) \pm u(-\vec{x}) = \pm u_\pm(-\vec{x}). \tag{2.46}$$

Hence $u_\pm$ also solve the stationary Schrödinger equation and all eigenfunctions can be chosen to be either even or odd. In one dimension we know that, in addition, energy eigenvalues are non-degenerate so that u_+ and u_- are proportional, which is only possible if one of these functions vanishes. We conclude that parity symmetry in *one dimension* implies that all eigenfunctions are automatically either even or odd. More precisely, eigenfunctions with an even (odd) number of nodes are even (odd), and, in particular, the ground state u_1 has an even eigenfunction, for the first excited state u_2 is odd with its single node at the origin, and so on.

2.2.1 One-dimensional square potentials and continuity conditions

In the search for stationary solutions we are going to solve equation (2.40) for the simple one-dimensional and time independent potential

$$V(x) = \begin{cases} 0 & \text{for} \quad |x| \geq a \\ V_0 & \text{for} \quad |x| < a \end{cases} . \tag{2.47}$$

For $V_0 < 0$ we have a potential well (also known as potential pot) with an attractive force and for $V_0 > 0$ a repulsive potential barrier, as shown in figure 2.3. Since the force becomes infinite (with a δ-function behavior) at a discontinuity of $V(x)$ such potentials are unphysical idealizations, but they are useful for studying general properties of the Schrödinger equation and its solutions by simple and exact calculations.

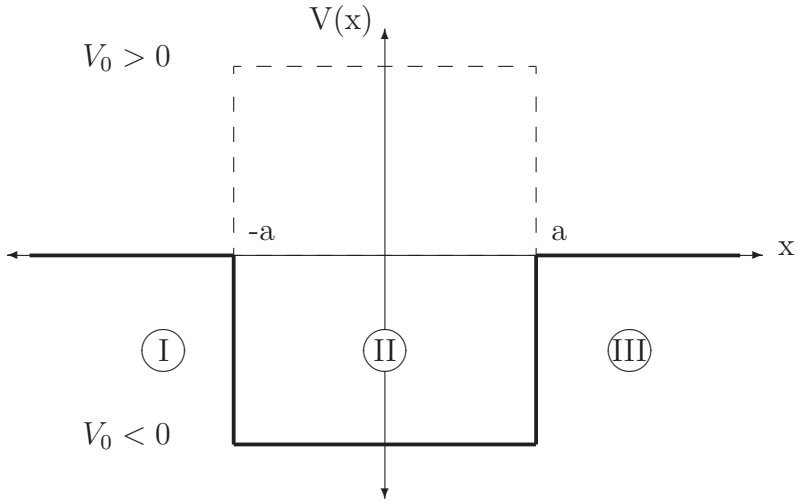


Figure 2.3: One-dimensional square potential well and barrier

Continuity conditions

We first need to study the behavior of the wave function at a discontinuity of the potential. Integrating the time-independent Schrödinger equation (2.40) in the form

$$u''(x) = \frac{2m}{\hbar^2}(V - E)u(x) \tag{2.48}$$

over a small interval $[a - \varepsilon, a + \varepsilon]$ about the position a of the jump we obtain

$$\int_{a-\varepsilon}^{a+\varepsilon} u''(x)\,dx = u'(a+\varepsilon) - u'(a-\varepsilon) = \frac{2m}{\hbar^2}\int_{a-\varepsilon}^{a+\varepsilon}(V - E)u(x)\,dx. \tag{2.49}$$

Assuming that $u(x)$ is continuous (or at least bounded) the r.h.s. vanishes for $\varepsilon \to 0$ and we conclude that the first derivative $u'(x)$ is continuous at the jump and only $u''(x)$ has a discontinuity, which according to eq. (2.48) is proportional to $u(a)$ and to the discontinuity of $V(x)$. With $u(a_\pm) = \lim_{\varepsilon\to 0} u(a \pm \varepsilon)$ the matching condition thus becomes

$$u(a_+) = u(a_-) \qquad \text{and} \qquad u'(a_+) = u'(a_-)\,, \tag{2.50}$$

confirming the consistency of our assumption of u being continuous. Even more unrealistic potentials like an infinitely high step for which finiteness of (2.49) requires

$$V(x) = \begin{cases} V_0 & \text{for } x < a \\ \infty & \text{for } x > a \end{cases} \qquad \Rightarrow \qquad u(x) = 0 \ \text{ for } x \geq a\,, \tag{2.51}$$

or δ-function potentials, for which (2.49) implies a discontinuity of u'

$$V(x) = V_{cont.} + A\,\delta(a) \qquad \Rightarrow \qquad \begin{cases} u(a_+) - u(a_-) = 0 \\ u'(a_+) - u'(a_-) = A\,\frac{2m}{\hbar^2}\,u(a) \end{cases}, \tag{2.52}$$

are used for simple and instructive toy models.

2.2.2 Bound states and the potential well

For a bound state in a potential well of the form shown in figure 2.3 we need

$$V_0 < E < 0. \tag{2.53}$$

The stationary Schrödinger equation takes the form

$$\begin{array}{lll} \frac{d^2}{dx^2}u(x) + k^2 u(x) = 0, & k^2 = \frac{2m}{\hbar^2}E = -\kappa^2 & \text{for } |x| > a \\ \frac{d^2}{dx^2}u(x) + K^2 u(x) = 0, & K^2 = \frac{2m}{\hbar^2}(E - V_0) & \text{for } |x| < a \end{array} \tag{2.54}$$

in the different sectors and the respective ansätze for the general solution read

$$\begin{array}{rllll} u_I & = A_1 e^{\kappa x} & + B_1 e^{-\kappa x} & for & x \le -a, \\ u_{II} & = A_2 e^{iKx} & + B_2 e^{-iKx} & for & |x| < a, \\ u_{III} & = A_3 e^{\kappa x} & + B_3 e^{-\kappa x} & for & x \ge a. \end{array} \tag{2.55}$$

For $x \to \pm\infty$ normalizability of the wave function implies $B_1 = A_3 = 0$. Continuity of the wave function and of its derivative at $x = \pm a$ implies the four matching conditions

$$u_I(-a) = u_{II}(-a) \qquad u_{II}(a) = u_{III}(a) \tag{2.56}$$

$$u_I'(-a) = u_{II}'(-a) \qquad u_{II}'(a) = u_{III}'(a) \tag{2.57}$$

or

$$u(-a) = \quad A_1 e^{-\kappa a} = A_2 e^{-iKa} + B_2 e^{iKa}, \qquad u(a) = A_2 e^{iKa} + B_2 e^{-iKa} = B_3 e^{-\kappa a}, \tag{2.58}$$

$$\tfrac{1}{iK}u'(-a) = \tfrac{\kappa}{iK}A_1 e^{-\kappa a} = A_2 e^{-iKa} - B_2 e^{iKa}, \quad \tfrac{1}{iK}u'(a) = A_2 e^{iKa} - B_2 e^{-iKa} = \tfrac{i\kappa}{K}B_3 e^{-\kappa a}. \tag{2.59}$$

These are 4 homogeneous equations for 4 variables, which generically imply that all coefficients vanish $A_1 = A_2 = B_2 = B_3 = 0$. Bound states (i.e. normalizable energy eigenfunctions) therefore only exist if the equations become linearly dependent, i.e. if the determinant of the 4×4 coefficient matrix vanishes. This condition determines the energy eigenvalues because κ and K are functions of the variable E.

Since the potential is parity invariant we can simplify the calculation considerably by using that the eigenfunctions are either even or odd, i.e. $B_2 = \pm A_2$ and $B_3 = \pm A_1$, respectively. With $A_2 = B_2 = \frac{1}{2}A_2'$ for u_{even} and $B_2 = -A_2 = \frac{i}{2}B_2'$ for u_{odd} the simplified ansatz becomes

$$\begin{array}{rcll} u_{even} & = & A_2' \cdot \cos(Kx) & \text{for } 0 < x < a, \\ u_{even} & = & B_3 \cdot e^{-\kappa x} & \text{for } a < x \end{array} \tag{2.60}$$

and

$$\begin{array}{rcll} u_{odd} & = & B_2' \cdot \sin(Kx) & \text{for } 0 < x < a, \\ u_{odd} & = & B_3 \cdot e^{-\kappa x} & \text{for } a < x. \end{array} \tag{2.61}$$

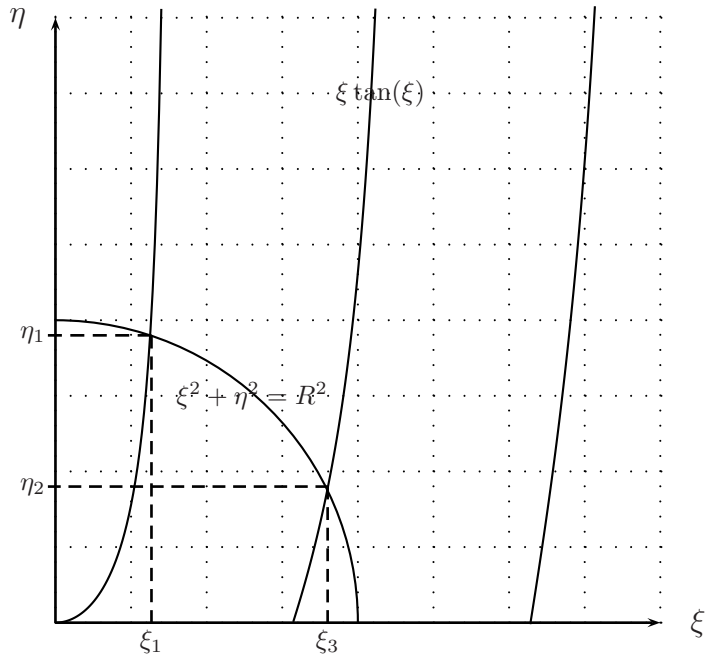


Figure 2.4: Graphical solution of the bound state energy equation for even eigenfunctions.

In both cases it is sufficient to impose the matching conditions for $x \geq 0$, i.e. at $x = a$. For the even solutions continuity of u and u' implies

$$A_2' \cdot \cos(Ka) = B_3 \cdot e^{-\kappa a} \tag{2.62}$$

$$-KA_2' \cdot \sin(Ka) = -\kappa B_3 \cdot e^{-\kappa a} \tag{2.63}$$

Taking the quotient we observe that the two equations are linearly dependent if

$$\tan(Ka) = \frac{\kappa}{K} \qquad \text{for } u_{even}. \tag{2.64}$$

For the odd case $u(a) = B_2' \sin(Ka) = B_3 \cdot e^{-\kappa a}$ and $u'(a) = KB_2' \cos(Ka) = -\kappa B_3 \cdot e^{-\kappa a}$ imply

$$\cot(Ka) = -\frac{\kappa}{K} \qquad \text{for } u_{odd}. \tag{2.65}$$

The respective wave functions are

$$u_{even}(x) = A_1 \cdot \begin{cases} e^{\kappa x} & x < -a, \\ e^{-\kappa a} \cdot \frac{\cos(Kx)}{\cos(Ka)} & |x| \leq a, \\ e^{-\kappa x} & x > a \end{cases} \tag{2.66}$$

and

$$u_{odd}(x) = A_1 \cdot \begin{cases} -e^{\kappa x} & x < -a, \\ e^{-\kappa a} \cdot \frac{\sin(Kx)}{\sin(Ka)} & |x| \leq a, \\ e^{-\kappa x} & x > a \end{cases} \tag{2.67}$$

with $|A_1|$ determined by the normalization integral $\int |u|^2 = 1$.

The transcendental equations (2.64) and (2.65), which determine the energy levels, cannot be solved explicitly. The key observation that enables a simple graphical solution is that

$K^2 + \kappa^2 = -2mV_0/\hbar^2$ is independent of E. In the (K, κ)–plane the solutions to the above equations therefore correspond to the intersection points of the graphs of these equations with a circle of radius $\sqrt{-2mV_0/\hbar^2}$. It is convenient to set $\xi = Ka$ and $\eta = \kappa a$, hence

$$\eta^2 + \xi^2 = (\kappa a)^2 + (Ka)^2 = -\frac{2mE}{\hbar^2}a^2 + \frac{2m}{\hbar^2}(E - V_0)a^2 = a^2\frac{2m}{\hbar^2}|V_0| = R^2 \tag{2.68}$$

The transcendental equations become

$$\eta = \begin{cases} \xi \tan(\xi) & \text{for } u_{even} \\ -\xi \cot(\xi) & \text{for } u_{odd} \end{cases} \tag{2.69}$$

where only values in the first quadrant $\xi, \eta > 0$ are relevant because K and κ were defined as positive square roots. Figure 2.4 shows the graph of the equations for even wave functions. We observe that there is always at least one solution with $0 < \xi < \pi/2$. The graph of the equation for the odd solutions looks similar with the branches of $-\cot\xi$ shifted by $\pi/2$ as compared to the branches of $\tan\xi$, so that indeed even and odd solutions alternate with increasing energy levels in accord with the *oscillation theorem*. An odd solution only exists if $R > \pi/2$ and for large R the number of bound states is approximately $\frac{2}{\pi}\frac{a}{\hbar}\sqrt{-2mV_0}$. The energy eigenvalues are related by

$$E_n = -\frac{(\hbar\kappa_n)^2}{2m} = -\frac{(\hbar\eta_n)^2}{2ma^2} = V_0 + \frac{(\hbar\xi_n)^2}{2ma^2} \tag{2.70}$$

to the common solutions (ξ_n, η_n) of equations (2.68) and (2.69).

2.2.3 Scattering and the tunneling effect

We now turn to the consideration of free electrons, i.e. to electrons whose energy exceeds the value of the potential at infinity. In this situation there are no normalizable energy eigenstates and a realistic description would have to work with wave packets that are superpositions of plane waves of different energies. A typical experimental situation is a accelerator where a beam of particles hits an interaction region, with particles scattered into different directions (for the time being we have to ignore the possibility of particle creation or annihilation).

In our one-dimensional situation the particles are either reflected or transmitted by their interaction with a localized potential. If we consider a stationary situation with an infinitely large experminent this means, however, that we do not need a normalizable wave function because the total number of particles involved is infinite, with a certain number coming out of the electron source per unit time. Therefore we can work with a finite and for $x \to \pm\infty$ constant *current density*, which describes the flow of particles. According to the correspondence $p = mv = \frac{\hbar}{i}\partial_x$ we expect that the wave functions

$$u_{right} = Ae^{ikx} \qquad \text{and} \qquad u_{left} = Be^{-ikx} \tag{2.71}$$

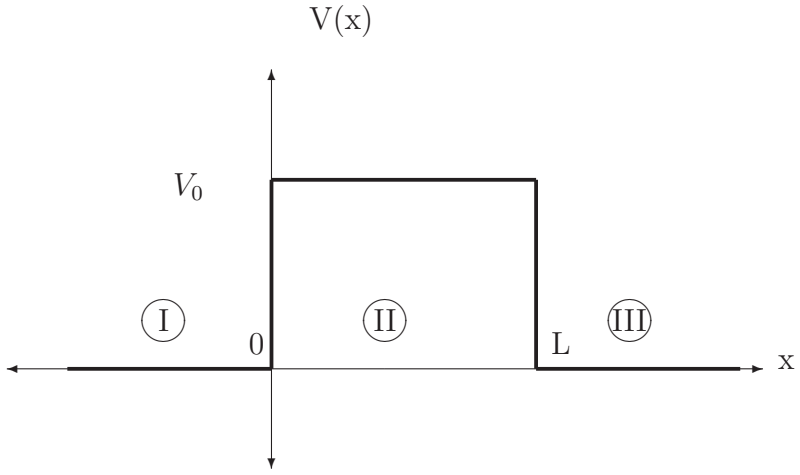


Figure 2.5: Potential barrier

describe right-moving and left-moving electron rays with velocities $v = \pm\hbar k/m$, respectively. Indeed, inserting into the formula (2.14) for the probability current density we find

$$j_{right} = \frac{\hbar k}{m}|A|^2 \qquad \text{and} \qquad j_{left} = -\frac{\hbar k}{m}|B|^2. \tag{2.72}$$

As a concrete example we again consider the square potential. For $V_0 > 0$ we have a potential barrier and for $V_0 < 0$ a potential well. Classically all electrons would be transmitted as long as $E > V_0$ and all electrons would be reflected by the potential barrier if $E < V_0$. Quantum mechanically we generically expect to find a combination of reflection and transmission, like in optics. For a high barrier $V_0 > E$ we will find an exponentially suppressed but non-zero probability for electrons to be able to penetrate the classically forbidden region, which is called tunneling effect. Our ansatz for the stationary wave function in the potential of figure 2.5 is

$$u_I = Ae^{ikx} + Be^{-ikx} \qquad \text{for} \quad x < 0 \qquad \text{with} \quad k = \sqrt{\tfrac{2mE}{\hbar^2}}, \tag{2.73}$$

$$u_{II} = \begin{cases} Fe^{-\kappa x} + Ge^{+\kappa x} & \text{for} \quad E < V_0 \quad \text{with} \quad \kappa = \sqrt{\tfrac{2m(V_0-E)}{\hbar^2}}, \\ Fe^{iKx} + Ge^{-iKx} & \text{for} \quad E > V_0 \quad \text{with} \quad K = \sqrt{\tfrac{2m(E-V_0)}{\hbar^2}} = i\kappa, \end{cases} \tag{2.74}$$

$$u_{III} = Ce^{ikx} + De^{-ikx} \qquad \text{for} \quad x > L. \tag{2.75}$$

Since for tunneling $E < V_0$ and for the case $E > V_0$ the ansätze in the interaction region II as well as the resulting continuity equations are formally related by $K = i\kappa$, both cases can be treated in a single calculation. Moreover, the ansatz for $E > V_0$ covers scattering at a potential barrier $V_0 > 0$ as well as the scattering at a potential well $V_0 < 0$.

Considering the physical situation of an electron source at $x \ll 0$ and detectors measuring the reflected and the transmitted particles we observe that A is the amplitude for the incoming ray, B is the amplitude for reflection, C is the amplitude for transmission and we have to set

$D = 0$ because there is no electron source to the right of the interaction region. We define the two quantities

$$\boxed{\textit{reflection coefficient } R = \left|\frac{j_{ref}}{j_{in}}\right|,} \qquad \boxed{\textit{transmission coefficient } T = \left|\frac{j_{trans}}{j_{in}}\right|,} \tag{2.76}$$

where the *reflection coefficient* R is defined as the ratio of the intensity of the reflected current over the intensity of the incident current and conservation of the total number of electrons implies $T = 1 - R$. Since parity symmetry of the Hamiltonian cannot be used to restrict the *scattering* ansatz to even or odd wave functions, we have shifted the interaction region by $a = L/2$ as compared to figure 2.3. This slightly simplifies some of the intermediate expressions, but of course does not change any of the observables like R and T. Using formulas (2.72) for the currents we find

$$R = \frac{|B|^2}{|A|^2} \qquad \text{and} \qquad T = \frac{|C|^2}{|A|^2}, \tag{2.77}$$

where we used that the velocities $v = \hbar k/m$ are equal $v_{III}/v_I = 1$ on both sides of the barrier. In situations where the potentials $V(\pm\infty)$ of the electron source and of the detector differ the ratio of the velocities has to be taken into account. For $E > V_0$ continuity of u and u' at $x = 0$,

$$A + B = F + G, \tag{2.78}$$

$$ik(A - B) = iK(F - G), \tag{2.79}$$

and at $x = L$,

$$Fe^{iKL} + Ge^{-iKL} = Ce^{ikL}, \tag{2.80}$$

$$iK(Fe^{iKL} - Ge^{-iKL}) = ikCe^{ikL}, \tag{2.81}$$

can now be used to eliminate F and G

$$2F = A(1 + \tfrac{k}{K}) + B(1 - \tfrac{k}{K}) = e^{i(k-K)L}C(1 + \tfrac{k}{K}), \tag{2.82}$$

$$2G = A(1 - \tfrac{k}{K}) + B(1 + \tfrac{k}{K}) = e^{i(k+K)L}C(1 - \tfrac{k}{K}). \tag{2.83}$$

From these equations we can eliminate either C or B,

$$e^{2iKL}\Big(A(K^2 - k^2) + B(K - k)^2\Big) = A(K^2 - k^2) + B(K + k)^2 \tag{2.84}$$

$$A\Big((K + k)^2 - (K - k)^2\Big) = 4kKA = Ce^{ikL}\Big(e^{-iKL}(K + k)^2 - e^{iKL}(K - k)^2\Big) \tag{2.85}$$

and solve for the ratios of amplitudes

$$\frac{B}{A} = \frac{(k^2 - K^2)(e^{2iKL} - 1)}{e^{2iKL}(k - K)^2 - (k + K)^2} \tag{2.86}$$

and

$$\frac{C}{A} = \frac{-4kKe^{-ikL}e^{iKL}}{e^{2iKL}(k - K)^2 - (k + K)^2}. \tag{2.87}$$

Using $(e^{2iKL}-1)(e^{-2iKL}-1) = 2 - e^{2iKL} - e^{-2iKL} = 2(1-\cos 2Kl) = 4\sin^2 KL$ we determine the reflection coefficient

$$R = \frac{|B|^2}{|A|^2} = \left[1 + \frac{4k^2K^2}{(k^2-K^2)^2 sin^2(KL)}\right]^{-1} = \left[1 + \frac{4E(E-V_0)}{V_0^2 sin^2(KL)}\right]^{-1} \tag{2.88}$$

and the transmission coefficient

$$T = \frac{|C|^2}{|A|^2} = \left[1 + \frac{(k^2-K^2)^2 sin^2(KL)}{4k^2K^2}\right]^{-1} = \left[1 + \frac{V_0^2 sin^2(KL)}{4E(E-V_0)}\right]^{-1}. \tag{2.89}$$

In general the transmission coefficient T is less than 1, in contrast to classical mechanics, where the particle would always be transmitted. There are two cases with perfect transmission $T = 1$: The first is of course when $V_0 = 0$ and the second is a resonance phenomenon that occurs when $KL = n\pi$ for $n = 1, 2, \ldots$, i.e. when $\sin KL = 0$ so that the length L of the interaction region is a half-integral multiple of the wavelength of the electrons. Conservation of probability $R + T = 1$ holds since $\frac{1}{1+X} + \frac{1}{1+1/X} = 1$.

As we mentioned above the case of a high barrier $V_0 > E$ is related to the formulas for $E > V_0$ by analytic continuation $K = i\kappa$. For the ratios B/A and C/A we hence obtain

$$\frac{B}{A} = \frac{(k^2+\kappa^2)(e^{2\kappa L}-1)}{e^{2\kappa L}(k+i\kappa)^2 - (k-i\kappa)^2}, \tag{2.90}$$

$$\frac{C}{A} = \frac{4ik\kappa e^{-ikL}e^{\kappa L}}{e^{2\kappa L}(k+i\kappa)^2 - (k-i\kappa)^2}, \tag{2.91}$$

which leads to the reflection and transmission coefficients

$$R = \frac{|B|^2}{|A|^2} = \left[1 + \frac{4k^2\kappa^2}{(k^2+\kappa^2)^2 sinh^2(\kappa L)}\right]^{-1} = \left[1 + \frac{4E(V_0-E)}{V_0^2 sinh^2(\kappa L)}\right]^{-1}, \tag{2.92}$$

$$T = \frac{|C|^2}{|A|^2} = \left[1 + \frac{(k^2+\kappa^2)^2 sinh^2(\kappa L)}{4k^2\kappa^2}\right]^{-1} = \left[1 + \frac{V_0^2 sinh^2(\kappa L)}{4E(V_0-E)}\right]^{-1}. \tag{2.93}$$

For $E < V_0$ neither perfect transmission nor perfect reflection is possible. For large L the transmission probability falls off exponentially

$$T \quad \longrightarrow \quad \frac{16E(V_0-E)}{V_0^2}\, e^{-2\kappa L} \qquad \text{for} \qquad L \gg 1/\kappa. \tag{2.94}$$

The phenomenon that a particle has a positive probability to penetrate a classically forbidden potential barrier is called *tunneling effect*.

2.2.4 Transfer matrix and scattering matrix

The wave functions $u_i(x) = A_i e^{ik_i x} + B_i e^{-ik_i x}$ in domains of constant potential are parametrized by the two amplitudes A_i and B_i. The effect of an interaction region can therefore be regarded

as a linear map expressing the amplitudes on one side in terms of the amplitudes on the other side. This map is called *transfer matrix*. For the potential in figure 2.5 and with our ansatz

$$u_I = Ae^{ikx} + Be^{-ikx}, \qquad u_{II} = \begin{cases} Fe^{-\kappa x} + Ge^{+\kappa x} \\ Fe^{iKx} + Ge^{-iKx} \end{cases}, \qquad u_{III} = Ce^{ikx} + De^{-ikx} \tag{2.95}$$

with

$$k = \tfrac{1}{\hbar}\sqrt{2mE}, \qquad \kappa = \tfrac{1}{\hbar}\sqrt{2m(V_0 - E)}, \qquad K = \tfrac{1}{\hbar}\sqrt{2m(E - V_0)} = i\kappa \tag{2.96}$$

the matching conditions

$$A + B = F + G \qquad ik(A - B) = iK(F - G) \tag{2.97}$$

can be solved for A and B,

$$\begin{pmatrix} A \\ B \end{pmatrix} = P \begin{pmatrix} F \\ G \end{pmatrix} \qquad \text{with} \qquad P = \frac{1}{2}\begin{pmatrix} 1 + \frac{K}{k} & 1 - \frac{K}{k} \\ 1 - \frac{K}{k} & 1 + \frac{K}{k} \end{pmatrix}. \tag{2.98}$$

At $x = L$ we find

$$\begin{pmatrix} F \\ G \end{pmatrix} = Q \begin{pmatrix} C \\ D \end{pmatrix} \qquad \text{with} \qquad Q = \frac{1}{2}\begin{pmatrix} (1 + \frac{k}{K})e^{i(k-K)L} & (1 - \frac{k}{K})e^{-i(k+K)L} \\ (1 - \frac{k}{K})e^{i(k+K)L} & (1 + \frac{k}{K})e^{-i(k-K)L} \end{pmatrix}. \tag{2.99}$$

The transfer matrix $M = PQ$ now relates the amplitudes for $x \to \pm\infty$ as

$$\begin{pmatrix} A \\ B \end{pmatrix} = M \begin{pmatrix} C \\ D \end{pmatrix}, \tag{2.100}$$

where A and D are the amplitudes for incoming rays while B and C are the amplitudes for outgoing particles. Because of the causal structure it appears natural to express the outgoing amplitudes in terms of the incoming ones,

$$\begin{pmatrix} B \\ C \end{pmatrix} = S \begin{pmatrix} A \\ D \end{pmatrix}. \tag{2.101}$$

This equation defines the *scattering matrix*, or S-matrix, which can be obtained from the transfer matrix by solving the linear equations $A = M_{11}C + M_{12}D$ and $B = M_{21}C + M_{22}D$ for $B(A, D)$ and $C(A, D)$. We thus find

$$C = \tfrac{1}{M_{11}}A - \tfrac{M_{12}}{M_{11}}D, \qquad B = \tfrac{M_{21}}{M_{11}}A + (M_{22} - \tfrac{M_{12}M_{21}}{M_{11}})D \tag{2.102}$$

or

$$\begin{pmatrix} S_{11} = \frac{M_{21}}{M_{11}} & S_{12} = M_{22} - \frac{M_{12}M_{21}}{M_{11}} \\ S_{21} = \frac{1}{M_{11}} & S_{22} = -\frac{M_{12}}{M_{11}} \end{pmatrix} \tag{2.103}$$

For $D = 0$ we recover the transmission and reflection coefficients as

$$T = \left|\frac{C}{A}\right|^2 = |S_{21}|^2 = \frac{1}{|M_{11}|^2}, \qquad R = \left|\frac{B}{A}\right|^2 = |S_{11}|^2 = \frac{|M_{21}|^2}{|M_{11}|^2} \tag{2.104}$$

(we can think of the index "1" as left and of "2" as right; hence $T = S_{21}$ describes scattering from left to right and $R = S_{11}$ describes backward scattering to the left).

Conservation of the probability current implies $|B|^2 + |C|^2 = |A|^2 + |D|^2$, i.e. the outgoing current of particles is equal to the incoming current. This can be written as

$$(A^*\ D^*)\begin{pmatrix}A\\D\end{pmatrix} = |A|^2 + |D|^2 = |B|^2 + |C|^2 = (B^*\ C^*)\begin{pmatrix}B\\C\end{pmatrix} = (A^*\ D^*)\,S^\dagger\,S\begin{pmatrix}A\\D\end{pmatrix}, \quad (2.105)$$

where $S^\dagger = (S^*)^T$ is the Hermitian conjugate matrix of S. Since this equality has to hold for arbitrary *complex* numbers A and D we conclude that the S-matrix has to be unitary $S^\dagger S = 1$ or $S^\dagger = S^{-1}$. We thus recover our previous result $R + T = 1$ as the 11-component of the unitarity condition $(S^\dagger S)_{11} = S_{11}^* S_{11} + S_{21}^* S_{21} = 1$.

2.3 The harmonic oscillator

A very important and also interesting potential is the harmonic oscillator potential

$$V(x) = \frac{m\omega_0^2}{2}x^2, \quad (2.106)$$

which is the potential of a particle with mass m which is attracted to a fixed center by a force proportional to the displacement from that center. The harmonic oscillator is therefore the prototype for systems in which there exist small vibrations about a point of stable equilibrium. We will only solve the one-dimensional problem, but the generalization for three dimensions is trivial because $|\vec{x}|^2 = x_1^2 + x_2^2 + x_3^2$ so that $H = H_x + H_y + H_z$. Thus we can make a separation ansatz $u(x, y, z) = u_1(x)u_2(y)u_3(z)$ and solve every equation separately in one dimension. The time independent Schrödinger equation we want to solve is

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{m\omega_0^2}{2}x^2\right]u(x) = Eu(x). \quad (2.107)$$

For convenience we introduce the dimensionless variables

$$\xi = \sqrt{\frac{m\omega_0}{\hbar}}\ x, \quad (2.108)$$

$$\lambda = \frac{2E}{\hbar\omega_0}. \quad (2.109)$$

Then the Schrödinger equation reads

$$\left[\frac{\partial^2}{\partial\xi^2} - \xi^2 + \lambda\right]u(\xi) = 0 \quad (2.110)$$

Since $\partial_\xi^2 e^{\pm\frac{1}{2}\xi^2} = (\xi^2 \pm 1)e^{\pm\frac{1}{2}\xi^2}$ the asymptotic behavior of the solution for $\xi \to \pm\infty$ is

$$u(\xi) \backsimeq e^{-\frac{1}{2}\xi^2}, \quad (2.111)$$

where we discarded the case of exponentential growth since we need normalizability. We hence make the ansatz

$$u(\xi) = v(\xi)e^{-\frac{1}{2}\xi^2} \tag{2.112}$$

Inserting into equation (2.110) gives the confluent hypergeometric differential equation:

$$\left[\frac{\partial^2}{\partial\xi^2} - 2\xi\frac{\partial}{\partial\xi} + \lambda - 1\right] v(\xi) = 0 \tag{2.113}$$

This is the *Hermite equation* which can be solved by using the power series ansatz

$$v(\xi) = \sum_{\nu=0}^{\infty} a_\nu \xi^\nu. \tag{2.114}$$

The harmonic oscillator potential is symmetric, therefore the eigenfunctions $u(\xi)$ of the Schrödinger equation must have a definite parity. We can therefore consider separately the even and the odd solutions.

For the even solutions we have $u(-\xi) = u(\xi)$ and therefore $v(-\xi) = v(\xi)$. Hence our power series ansatz becomes

$$v(\xi) = \sum_{\nu=0}^{\infty} c_\nu \xi^{2\nu}, \qquad c_\nu = a_{2\nu}. \tag{2.115}$$

Substituting (2.115) into the Hermite equation (2.113), we find that

$$\sum_{\nu=0}^{\infty}[2(\nu+1)(2\nu+1)c_{\nu+1} + (\lambda - 1 - 4\nu)c_\nu]\xi^{2\nu} = 0. \tag{2.116}$$

This equation will be satisfied provided the coefficient of each power of ξ separately vanishes, so that we obtain the *recursion relation*

$$c_{\nu+1} = \frac{4\nu + 1 - \lambda}{2(\nu+1)(2\nu+1)}c_\nu. \tag{2.117}$$

Thus, given $c_0 \neq 0$, all the coefficients c_ν can be determined successively by using the above equation. We thus have obtained a series representation of the even solution (2.115) of the Hermite equation. If this series does not terminate, we see from (2.117) that for large ν

$$\frac{c_{\nu+1}}{c_\nu} \sim \frac{1}{\nu}. \tag{2.118}$$

This ratio is the same as that of the taylor expansion of $\exp(\xi^2)$, which implies that the wave function $u(\xi)$ has an asymptotic behavior of the form

$$u(\xi) \sim \xi^{2p} e^{\xi^2/2} \qquad \text{for} \qquad |\xi| \to \infty, \tag{2.119}$$

which would spoil normalizability. The only way to avoid this divergence is to require that the series terminates, which means that $v(\xi)$ must be a polynomial in the variable ξ^2. Using the relation (2.117) we see, that the series only terminates, when λ takes on the discrete values

$$\lambda = 4N + 1, \qquad N = 0, 1, 2, \ldots . \tag{2.120}$$

To each value N = 0,1,2,..., of N will then correspond an even function $v(\xi)$ which is a polynomial of order 2N in ξ, and an even, physically acceptable, wave function $u(\xi)$ which is given by (2.115). In a similar way we obtain the odd states by using the power series

$$u(\xi) = \sum_{\nu=0}^{\infty} b_\nu \xi^{2\nu+1}, \tag{2.121}$$

which contains only odd powers of ξ. We again substitute the ansatz into the Hermite equation and obtain a recursion relation for the coefficients b_ν. We now see, that the series terminates for the discrete values

$$\lambda = 4N + 3, \qquad N = 0, 1, 2, \ldots . \tag{2.122}$$

To each value N = 0,1,2,..., of N will then correspond an odd function $v(\xi)$ which is a polynomial of order 2N+1 in ξ, and an odd, physically acceptable wave function $u(\xi)$ given by (2.114). Putting together the results we see that the eigenvalue λ must take on one of the discrete values

$$\lambda = 2n + 1, \qquad n = 0, 1, 2, \ldots \tag{2.123}$$

where the *quantum number* n is a positive integer or zero. Inserting in (2.109) we therefore find that the energy spectrum of the linear harmonic oscillator is given by

$$\boxed{E_n = \left(n + \frac{1}{2}\right)\hbar\omega_0, \qquad n = 0, 1, 2, \ldots} \tag{2.124}$$

We see that, in contrast to classical mechanics, the quantum mechanical energy spectrum of the linear harmonic oscillator consists of an infinite sequence of discrete levels. The eigenvalues are non-degenerate since for each value of the quantum number n there exists only one eigenfunction (apart from an arbitrary multiplicative constant) and the energy of the lowest state, the **zero-point-energy** $E_0 = \hbar\omega/2$, is positive. This can be understood as a consequence of the uncertainty relation $\Delta X\, \Delta P \geq \hbar/2$ because $\langle P\rangle = \langle X\rangle = 0$ so that

$$E = \tfrac{1}{2m}(\Delta P)^2 + \tfrac{m\omega_0^2}{2}(\Delta X)^2, \tag{2.125}$$

which has the positive minimum E_0. Since the wave function factors $v_n(\xi)$ are solutions of the Hermite equation and polynomials of the order n, they are proportional to the *Hermite polynomials* $H_n(\xi)$, which can be definded as

$$H_n(\xi) = (-1)^n e^{\xi^2} \frac{d^n e^{-\xi^2}}{d\xi^n} \tag{2.126}$$

$$= e^{\xi^2/2}\left(\xi - \frac{d}{d\xi}\right)^n e^{-\xi^2/2}. \tag{2.127}$$

This leads to the explicit formula

$$H_{2n} = \sum_{k=0}^{n} (-1)^{n-k} \frac{(2n)!(2\xi)^{2k}}{(n-k)!(2k)!}, \qquad H_{2n+1} = 2\xi \sum_{k=0}^{n} (-1)^k \frac{(2n+1)!(2\xi)^{2n-2k}}{k!(2n-2k+1)!} \tag{2.128}$$

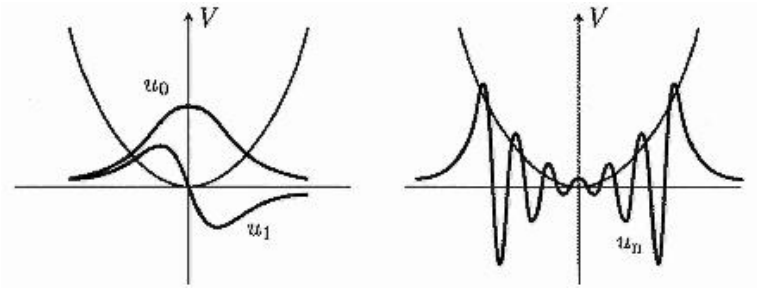


Figure 2.6: Hermite polynomials

for $n \geq 0$. The first few polynomials are

$$H_0(\xi) = 1, \tag{2.129}$$
$$H_1(\xi) = 2\xi, \tag{2.130}$$
$$H_2(\xi) = 4\xi^2 - 2, \tag{2.131}$$
$$H_3(\xi) = 8\xi^3 - 12\xi, \tag{2.132}$$
$$H_4(\xi) = 16\xi^4 - 48\xi^2 + 12, \tag{2.133}$$

as shown in figure 2.6.

The *generating function* $G(\xi, s)$ of the Hermite polynomials is

$$G(\xi, s) = \sum_{n=0}^{\infty} \frac{H_n(\xi)}{n!} s^n = e^{-s^2+2s\xi}. \tag{2.134}$$

These relations mean that if the function $exp(-s^2 + 2s\xi)$ is expanded in a power series in s, the coefficients of successive powers of s are just $1/n!$ times the Hermite polynomials $H_n(\xi)$. Putting everything together we find the eigenfunctions

$$u_n(x) = N_n e^{-\alpha^2 x^2/2} H_n(\alpha x), \qquad \alpha = \left(\frac{m\omega}{\hbar}\right)^{1/2} \tag{2.135}$$

with normalization constants N_n, which have to be determined by requiring that the wave function $u(x)$ be normalized to unity. That is

$$\int |u_n(x)|^2 dx = \frac{|N_n|^2}{\alpha} \int e^{-\xi^2} H_n^2(\xi) d\xi = 1. \tag{2.136}$$

To calculate the normalization constant we use the generating function (2.134) twice,

$$G(\xi, s) = \sum_{n=0}^{\infty} \frac{H_n(\xi)}{n!} s^n = e^{-s^2+2s\xi} \tag{2.137}$$

and

$$G(\xi, t) = \sum_{m=0}^{\infty} \frac{H_m(\xi)}{m!} t^m = e^{-t^2+2t\xi}. \tag{2.138}$$

With these two expressions we compute

$$\int e^{-\xi^2} G(\xi, s) G(\xi, t) d\xi = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{s^n t^m}{n!m!} \int e^{-\xi^2} H_n(\xi) H_m(\xi) d\xi \tag{2.139}$$

Using the fact that

$$\int e^{-x^2} dx = \sqrt{\pi} \tag{2.140}$$

We can calculate the left-hand side of (2.139) to

$$\begin{aligned} \int e^{-\xi^2} e^{-s^2+2s\xi} e^{-t^2+2t\xi} d\xi &= e^{2st} \int e^{-(\xi-s-t)^2} d(\xi - s - t) \\ &= \sqrt{\pi} e^{2st} \\ &= \sqrt{\pi} \sum_{n=0}^{\infty} \frac{(2st)^n}{n!} \end{aligned} \tag{2.141}$$

Equating the coefficients of equal powers of s and t on the right hand sides of (2.134) and (2.141), we find that

$$\int e^{-\xi^2} H_n^2(\xi) d\xi = \sqrt{\pi} 2^n n! \tag{2.142}$$

and

$$\int e^{-\xi^2} H_n(\xi) H_m(\xi) d\xi = 0 \qquad \text{for} \qquad n \neq m \tag{2.143}$$

From (2.136) and (2.142) we see that apart from an arbitrary complex multiplicative factor of modulus one the normalization constant N_n is given by

$$N_n = \left(\frac{\alpha}{\sqrt{\pi} 2^n n!} \right)^{1/2}. \tag{2.144}$$

and hence the normalized linear harmonic oscillator eigenfunctions are given by

$$\boxed{u_n(x) = \left(\frac{\alpha}{\sqrt{\pi} 2^n n!} \right)^{1/2} e^{-\alpha^2 x^2/2} H_n(\alpha x).} \tag{2.145}$$

These eigenfunctions are orthogonal and form an orthonormal set of functions

$$\int u_n^*(x) u_m(x) dx = \delta_{nm}, \tag{2.146}$$

which is a general property of eigenfunctions of self-adjoint operators as we will learn in the next chapter.

Chapter 3

Formalism and interpretation

Gott würfelt nicht mit dem Universum!
Albert Einstein

Ich denke nicht, dass es unsere Aufgabe ist dem Herrgott Vorschriften zu machen ...
Niels Bohr

The theory of quantum electrodynamics describes Nature as absurd from the point of view of common sense. And it agrees fully with the experiment. So I hope you can accept Nature as She is - absurd.
Richard P. Feynman, "QED"

The mathematical formalism of quantum theory, which we want to develop in this chapter, is based on the fact that the solutions of the Schrödinger equation form a Hilbert space, i.e. a vector space that is complete with respect to the norm defined by an inner product. All equations of the theory can be interpreted in terms of operators, i.e. linear maps on this space. This point of view is useful for theoretical as well as for practical reasons. As an example, we will solve the Schrödinger equation for the harmonic oscillator purely algebraically by introducing creation and annihilation operators. Along the way we will discuss the axioms and the interpretation of quantum mechanics, derive the general uncertainty relation, and develop new concepts and computational tools like the Heisenberg picture and density matrices.

3.1 Linear algebra and Dirac notation

The Schrödinger equation is a linear homogeneous differential equation. Its set of solutions therefore forms a vector space $\mathcal{H}$ over the complex numbers, because linear combinations of solutions with complex coefficients are again solutions. But this vector space is, in general, infinite dimensional. We should hence also admit infinite linear combinations so that convergence properties of such infinite sums have to be considered. The notion of convergence is based on a measure $||v||$ for the length of a vector, where a sequence is called convergent if the distance between its members and its limit vector goes to 0. The length $||v||$ has to be positive and is required to satisfy the triangle inequality $||v+w|| \leq ||v|| + ||w||$. It is called a norm on a vector space if it scales linearly according to $||\alpha v|| = |\alpha|\,||v||$, where $|\alpha|$ is the modulus of the complex number $\alpha \in \mathbb{C}$. A vector space with such a norm, a normed space, is called Banach space if it is complete (i.e. if it contains the limits for all Cauchy sequences, where a Cauchy sequence is a sequence for which the distances between its elements converge to 0).

Observables in quantum mechanics, like momentum or energy, are given by linear operators, i.e. by linear maps, which are the analogues of matrices in finite-dimensional spaces. Many of the concepts and tools of linear algebra can be extended to infinite-dimensional linear spaces. This is the subject of the mathematical discipline of functional analysis [Reed, Kreyszig].

Hilber spaces: In quantum mechanics there is a natural norm, namely the square root of the integral of the probability density of a wave function $\psi(x)$ at some given time t,

$$||\psi|| = \sqrt{Q} \qquad \text{with} \qquad Q = \int_{\mathbb{R}^3} d^3x \; |\psi(x)|^2 \tag{3.1}$$

(as we have shown it is time–independent for solutions of the Schrödinger equation). This norm has the additional property that it can be defined in terms of an inner product (φ, ψ) by

$$||\psi|| = \sqrt{(\psi, \psi)} \qquad \text{with} \qquad (\varphi, \psi) = \int_{\mathbb{R}^3} d^3x \; \varphi^*(x)\psi(x). \tag{3.2}$$

An inner product (φ, ψ) is semi-bilinear and symmetric up to complex conjugation,

$$(\varphi, \alpha\psi_1 + \beta\psi_2) = \alpha(\varphi, \psi_1) + \beta(\varphi, \psi_2), \qquad (\varphi, \psi) = (\psi, \varphi)^*, \tag{3.3}$$

where semi-bilinear means linear in the second entry and anti-linear in the first,

$$(\alpha\varphi_1 + \beta\varphi_2, \psi) = \alpha^*(\varphi_1, \psi) + \beta^*(\varphi_2, \psi). \tag{3.4}$$

as implied by eq. (3.3).

Note that anti-linearity (i.e. the complex conjugation of scalar coefficients) for the first entry is necessary because strict bilinearity would be inconsistent with positivity of the norm

$||\psi||^2 = (\psi, \psi) \geq 0$. To see this compare $||(i\psi)||^2 = (i\psi, i\psi)$ with $||(\psi)||^2$. A Banach space whose norm is defined by (3.2) in terms of a positive definite inner product (φ, ψ) is called a Hilbert space. The standard Hilbert space $\mathcal{H}$ of quantum mechanics is the space of complex-valued square-integrable functions $\psi(x) \in \mathcal{H} = L^2(\mathbb{R}^3)$, where the letter L stands for *Lebesgue* integration (which has to be used to make the space complete).[1] This is an ∞-dimensional vector space.

Let us pretend for a while that our Hilbert space is a finite-dimensional complex vector space. We will introduce a number of concepts like commutators and exponentiation of linear operators. The definitions will be straightforward for (finite-dimensional) matrices, but the same calculus can then be used for linear operators in Hilbert spaces. Refinements that are needed for the infinite-dimensional situation will then be discussed later on.

In linear algebra each vector space V automatically provides us with another linear space, called the *dual space* V^{dual}, which consists of the linear maps $w \in V^{dual}$ from vectors $v \in V$ to numbers $w(v) \in \mathbb{C}$. The numbers $w(v)$ are real for real and complex for complex vector spaces, respectively. We can think of vectors $v \in V$ as column vectors and of dual vectors $w \in V^{dual}$ as line vectors, so that their product, the duality bracket $\langle w, v\rangle \equiv w(v)$ is a number. If we introduce a basis e_i of V we can write each vector v as a unique linear combination $v = v^i e_i$ and each co-vector $w = w_j e^j$ is a sum of the elements of the dual basis e^j, which has upper indices and is defined by $\langle e^j, e_i\rangle = \delta^j_i$. Evaluation of w on v by linearity thus implies the formula

$$\langle w, v\rangle \equiv w(v) = w_j \langle e^j, e_i\rangle v^i = w_j v^j, \qquad \text{with} \quad w = w_j e^j, \quad v = v^i e_i, \quad \langle e^j, e_i\rangle = \delta^j_i. \tag{3.5}$$

If we now make a change of basis $\hat{e}_i = G_i{}^j e_j$ then the components of vectors transform with the inverse transposed matrix, and the same is true for the dual basis vectors $\hat{e}^j$:

$$v = v^i e_i = \hat{v}^i \hat{e}_i, \qquad \hat{e}_i = G_i{}^j e_j \qquad \Rightarrow \qquad \hat{v}^i = v^k (G^{-1})_k{}^i = (G^{-1T})^i{}_k v^k, \tag{3.6}$$

$$\delta^j_i = \langle e^j, e_i\rangle = \langle \hat{e}^j, \hat{e}_i\rangle, \quad \hat{e}_i = G_i{}^j e_j \qquad \Rightarrow \qquad \hat{e}^j = e^l (G^{-1})_l{}^j = (G^{-1T})^j{}_l e^l. \tag{3.7}$$

Co-vectors $w \in V^{dual}$, on the other hand, transform in the same way as the elements e_i of the basis, $\hat{w}_j = G_j{}^l w_l$, and also have the same index position. They are therefore called *covariant* vectors. It might be tempting to identify *contravariant* vectors $v \in V$ (column vectors, with upper indices, transforming like the dual basis $\hat{e}^j$) and covariant vectors $w \in V^{dual}$ (line vectors, with lower indices, transforming like the original basis) by transposition. Indeed this is possible in Euclidean space if we restrict ourselves to use orthonormal bases, because then the matrix G for the change of basis has to be orthogonal $G = G^{-1T}$ so that upper and lower indices transform in the same way. In other situations, like in the Minkowski space of special relativity

[1] Riemann integration would define a *pre-Hilbert space* or *inner product space*, whose Cauchy sequences need not converge. Such spaces can always be completed to Hilber spaces similarly to the completion of $\mathbb{Q}$ to $\mathbb{R}$.

(where the metric is not positive definite) or in quantum mechanics, where the inner product is *semi*-bilinear, it is important to distinguish between the two kinds of vectors.[2]

Dirac notation: Dirac introduced a very elegant and efficient notation for the use of linear algebra in quantum mechanics that is also called bra-ket notation because products are written by a bracket $\langle \ldots \rangle$ as in eq. (3.8). We introduce bra-vectors $\langle w|$ and ket-vectors $|v\rangle$, which are just the co- and contravariant vectors $\langle w| \equiv w \in V^{dual}$ and $|v\rangle \equiv v \in V$, respectively. Their duality pairing can be written as a *bra-ket* product,

$$\langle w, v\rangle = w_i v^i = \langle w| \cdot |v\rangle \equiv \langle w|v\rangle. \tag{3.8}$$

The Dirac notation is basis independent. Instead of using vector components v^i with respect to some predefined basis e_i we will rather identify a state vector by specifying its physical properties, i.e. by the quantum numbers of the state of the physical system which it describes. For the energy eigenfunctions of the harmonic oscillator we can write, for example,

$$u_n(x) \;\equiv\; \left|E = \hbar\omega_0(n + \tfrac{1}{2})\right\rangle \;\equiv\; |E_n\rangle \;\equiv\; |n\rangle, \tag{3.9}$$

where it is sufficient to characterize the state by the number $n = 0, 1, \ldots$ if it is clear from the context what quantum number we are referring to. The bra-ket notation is sufficiently flexible to allow us to write as much (or as little) information as we need. It is important, however, to keep in mind that even a complete set of quantum numbers, which by definition uniquely defines the *physical* state of the quantum system, fixes the wave function only up to an overall phase. Bra- and ket-vectors, accordingly, are determined by the quantum numbers only up to a phase $|n\rangle' = e^{i\rho}|n\rangle$ and $\langle n|' = e^{-i\rho}\langle n|$. It is important not to change the implicit choice of that phase during the course of a calculation! Observable quantities will then be independent of such choices.

The inner product allows us to define a natural map from V to its dual by inserting an element v into the first position of the inner product. For $|v\rangle \in V$ the Hermitian conjugate vector $\langle v| \in V^{dual}$ is defined by

$$|v\rangle^\dagger \equiv \langle v| \;\in\; V^{dual} \qquad \text{such that} \qquad \langle v|u\rangle = (v, u) \qquad \text{for all } u \in V. \tag{3.10}$$

Since the inner product is positive definite this conjugation is a bijective map from V to V^{dual} (this is also true for infinite-dimensional Hilbert spaces), but it is an "anti-isomorphism" and not an isomorphism because it is "anti-linear"

$$\boxed{\big(\alpha|v\rangle + \beta|w\rangle\big)^\dagger = \alpha^*\langle v| + \beta^*\langle w|.} \tag{3.11}$$

[2] In solid state physics the same distinction has to be made between the lattice Λ of atoms in a crystal and the reciprocal lattice Λ^{dual} of wave vectors; if Λ becomes finer then the reciprocal lattice becomes coarser.

Linearity can be achieved by an additional complex conjugation so that V^{dual} is isomorphic to the complex conjugate space V^*, while V can be identified with its Hermitian conjugate $V \cong V^\dagger \equiv (V^{dual})^*$. We will henceforth use these identifications and the antilinear map $|v\rangle \to |v\rangle^\dagger = \langle v| \in V^{dual}$, which corresponds to the equation $\langle v, u\rangle \equiv \langle v|u\rangle = (v, u)$. For column vectors $|v\rangle$ the Hermitian conjugate is the line vector $\langle v|$ with complex conjugate entries. For wave functions $|\psi\rangle = \psi(x,t)$ it is the complex conjugate function $\langle\psi| = |\psi\rangle^\dagger = \psi^*(x,t)$.

3.2 Operator calculus

The components v^i of a vector v in an arbitrary basis can be obtained by evaluation of the dual basis $v^i = e^i(v) = \langle e^i, v\rangle$ because $e^i(v) = e^i(v^j e_j) = v^j e^i(e_j) = v^j \delta^i_j = v^i$. For an orthonormal basis $(e_i, e_j) = \delta_{ij}$ we observe that $|e_i\rangle^\dagger = \langle e_i| = \langle e^i|$, i.e. the Hermitian conjugate vector $|e_i\rangle^\dagger$ coincide with the dual basis $\langle e^i|$ and

$$|v\rangle = v^i e_i = |e_i\rangle\langle e^i|v\rangle = \sum_i |e_i\rangle\langle e_i|v\rangle, \tag{3.12}$$

where we have chosen, for later convenience, to use Einstein's summation convention only for contractions of upper and lower indices. Since the identity (3.12) holds for all v we get a representation of the unit matrix, or identity operator

$$\mathbb{1} = \sum_i |e_i\rangle\langle e_i| = \sum_i P_i \qquad \text{with} \qquad P_i = |e_i\rangle\langle e_i|. \tag{3.13}$$

Orthonormal bases are thus characterized by the two equations

$$\langle e_i|e_j\rangle = \delta_{ij} \qquad \text{orthonormality,} \tag{3.14}$$

$$\sum_i |e_i\rangle\langle e_i| = \mathbb{1} \qquad \text{completeness.} \tag{3.15}$$

P_i is the (orthogonal) **projector** onto the direction of the basis vector $|e_i\rangle$. As an example we consider the standard basis of $\mathbb{C}^3$,

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{3.16}$$

The orthogonality relation reads

$$\langle e_1|e_1\rangle = (1,0,0)\cdot\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 1, \qquad \langle e_1|e_2\rangle = (1,0,0)\cdot\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 0, \qquad \ldots \tag{3.17}$$

and the projectors

$$|e_1\rangle\langle e_1| = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\cdot(1,0,0) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad |e_2\rangle\langle e_2| = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad |e_3\rangle\langle e_3| = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.18}$$

add up to the completeness relation

$$\sum_i |e_i\rangle\langle e_i| = \mathbb{1}. \tag{3.19}$$

While the product $\langle v|w\rangle$ of a covector and a vector yields a complex number, the tensor product $|w\rangle\langle v|$ is a matrix of rank 1 that is sometimes called *dyadic product*.

For a linear transformation $v \to Av$ the components $A^i{}_j$ of the matrix representation $v^i \to A^i{}_j v^j$ can be obtained by sandwiching the operator A between basis elements. For an orthonormal basis $\langle e_i|e_j\rangle = \delta_{ij}$ we can use the Kronecker–δ to pull all indices down so that the entries (elements) of the matrix $A^i{}_j = e^i(Ae_j)$ in Dirac notation become

$$A_{ij} = \langle e_i|A|e_j\rangle. \tag{3.20}$$

In quantum mechanics the numbers $\langle v|A|w\rangle$ are hence called **matrix elements** even for arbitrary bra- and ket-vectors v and w. The normalized diagonal term

$$\langle A\rangle_v = \frac{\langle v|A|v\rangle}{\langle v|v\rangle} \tag{3.21}$$

is called expectation value of the operator A in the state $|v\rangle$, where the denominator can obviously be omitted if and only if $|v\rangle$ is normalized $\langle v|v\rangle = 1$.

Hermitian conjugation. If we apply a linear transformation $v \to Av$ to a vector v and evaluate a covector w, i.e. multiply with w from the left, the resulting number is

$$\langle w, Av\rangle = w_i\, A^i{}_j\, v^j = \langle w| \cdot A|v\rangle. \tag{3.22}$$

But we might just as well first perform the sum over i in $w_i A^i{}_j$ and then multiply the resulting bra-vector $\langle w|A$, with the ket-vector $|v\rangle$ from the right. In the language of linear algebra this defines the *transposed* map A^T on the dual space V^{dual}, which can be written as a matrix multiplication $w_j \to (A^T)_j{}^i w_i$ with the transposed matrix A^T. Using the non-degenerate inner product we can define the *Hermitian conjugate* $A^\dagger$ of the linear operator A by

$$\boxed{(A^\dagger v, w) \equiv (v, Aw) \qquad \forall\ v, w \in V.} \tag{3.23}$$

Using $(\varphi, \psi) = (\psi, \varphi)^*$ we obtain the matrix elements

$$\langle v|A|w\rangle = \langle A^\dagger v|w\rangle = (\langle w|A^\dagger v\rangle)^* \qquad \Rightarrow \qquad \langle w|A^\dagger|v\rangle = \langle v|A|w\rangle^*. \tag{3.24}$$

For an orthonormal basis $|e_i\rangle$ the compoments become

$$\boxed{(A^\dagger)_{ij} = \langle e_i|A^\dagger|e_j\rangle = \langle e_j|A|e_i\rangle^* = A^*_{ji},} \tag{3.25}$$

so that Hermitian conjugation is transposition combined with complex conjugation of the matrix elements. Like transposition, Hermitian conjugation reverses the order of a product of matrices $(AB)^\dagger = B^\dagger A^\dagger$ and

$$(\alpha\langle\varphi|A\dots B|\psi\rangle)^* = (\alpha\langle\varphi|A\dots B|\psi\rangle)^\dagger = \alpha^*\langle\psi|B^\dagger\dots A^\dagger|\varphi\rangle \tag{3.26}$$

because Hermitian conjugation of a number is just complex conjugation.

An operator is called *self-adjoint* or *symmetric* or *Hermitian*[3] if $A^\dagger = A$. Consider a normalized eigenvector $|a_i\rangle$ for the eigenvalue a_i of a **self-adjoint** operator

$$A|a_i\rangle = a_i|a_i\rangle \quad\Rightarrow\quad \langle a_i|A^\dagger = \langle a_i|a_i^*, \qquad a_i = \langle a_i|\cdot(A|a_i\rangle) = \langle a_i|\cdot(A^\dagger|a_i\rangle) = (\langle a_i|A^\dagger)\cdot|a_i\rangle = a_i^*, \tag{3.27}$$

i.e. all **eigenvalues are real**, and hence

$$0 = \langle a_i|(A^\dagger - A)|a_j\rangle = \langle a_i|A^\dagger\cdot|a_j\rangle - \langle a_i|\cdot A|a_j\rangle = (a_i - a_j)\langle a_i|a_j\rangle \tag{3.28}$$

so that **eigenvectors** for different eigenvalues $a_i \neq a_j$ **are orthogonal** $\langle a_i|a_j\rangle = 0$.

Self-adjoint operators and spectral representation. The importance of self-adjoint operators $A = A^\dagger$ in quantum mechanics comes from the fact that they are exactly the operators for which all expectation values are real,[4]

$$(\langle\varphi|A|\psi\rangle)^* = \langle\psi|A^\dagger|\varphi\rangle = \langle\psi|A|\varphi\rangle \qquad\Rightarrow\qquad \langle\psi|A|\psi\rangle \in \mathbb{R}, \tag{3.32}$$

as we require for observable quantities. Hermitian matrices can be diagonalized and have real eigenvalues. Eigenvectors for different eigenvalues are orthogonal. In case of **degenerate eigenvalues**, i.e. eigenvalues with multiplicity greater than 1, a basis of eigenvectors for the

[3] For infinite-dimensional Hilbert spaces there is a subtle difference between the definitions of symmetric and self-adjoint operators, respectively, because due to convergence issues an operator may only be defined on a dense subset of $\mathcal{H}$ (see below). *Hermitian* is used synonymical with *symmetric* by most authors.

[4] To see that Hermiticity is also necessary for real expectation values we bring A to Jordan normal form and assume that there is a non-trivial block $A = \left(\begin{smallmatrix} a & 1 \\ 0 & a \end{smallmatrix}\right)$ with basis vectors $|e_1\rangle = \binom{1}{0}$ and $|e_2\rangle = \binom{0}{1}$, i.e.

$$A|e_1\rangle = a|e_1\rangle, \qquad A|e_2\rangle = |e_1\rangle + a|e_2\rangle. \tag{3.29}$$

Reality of $(a_i, Aa_i) = a_i(a_i, a_i) = a_i||a_i||^2$ for eigenvectors $|a_i\rangle$ implies reality of all eigenvalues a_i. For $|\psi\rangle = \alpha|e_1\rangle + \beta|e_2\rangle$ we find

$$\begin{aligned}\langle\psi|A|\psi\rangle &= (\alpha^*\langle e_1| + \beta^*\langle e_2|)\,((\alpha\, a + \beta)|e_1\rangle + \beta\, a|e_2\rangle) \\ &= a\left(|\alpha|^2||e_1||^2 + |\beta|^2||e_2||^2 + \alpha^*\beta(e_1, e_2) + \alpha\beta^*(e_2, e_1)\right) + |\beta|^2(e_2, e_1) + \alpha^*\beta||e_1||^2\end{aligned} \tag{3.30}$$

which cannot be real for all α if $\beta \neq 0$. Real expectation values hence imply diagonalizability. It remains to show that eigenvectors for different eigenvalues are orthogonal. We consider $|\varphi\rangle = \alpha|a_i\rangle + \beta|a_j\rangle$ and compute

$$\langle\varphi|A|\varphi\rangle = (\alpha^*\langle a_i| + \beta^*\langle a_j|)\,(a_i\,\alpha|a_i\rangle + a_j\,\beta|a_j\rangle = \text{real} + a_j(\alpha^*\beta(a_i, a_j)) + a_i(\alpha^*\beta(a_i, a_j))^*, \tag{3.31}$$

which cannot be real for $a_i \neq a_j$ and arbitrary α, β unless $(a_i, a_j) = 0$. We conclude that a matrix A with real expectation values is diagonalizable with real eigenvalues and orthogonal eigenspaces, and hence is Hermitian.

respective eigenvalue can be **orthonormalized by** the **Gram-Schmidt** algorithm and the resulting vectors have to be distinguished by *additional "quantum numbers"* l_i in $|a_i, l_i\rangle$ with $l_i = 1, \ldots, N_i$. The l_i have to be summed over in the completeness relation. For Hermitian matrices we thus can construct an orthonormal basis of eigenvectors $A|a_i\rangle = a_i|a_i\rangle$ with $\langle a_i|a_j\rangle = \delta_{ij}$, or, more precisely,

$$A|a_i, l_i\rangle = a_i|a_i, l_i\rangle \quad \text{with} \quad \langle a_i, l_i|a_j, k_j\rangle = \delta_{ij}\delta_{l_i k_j} \tag{3.33}$$

in the degenerate case. Using the completeness relation this implies the *spectral representation*

$$A = \sum_{i,l_i} A\,|a_i, l_i\rangle\langle a_i, l_i| = \sum_{i,l_i} a_i\,|a_i, l_i\rangle\langle a_i, l_i| = \sum_i a_i\,P_i, \tag{3.34}$$

where

$$P_i = \sum_{l_i=1}^{N_i} |a_i, l_i\rangle\langle a_i, l_i| \tag{3.35}$$

is the orthogonal projector onto the eigenspace for the eigenvalue a_i.

Unitary, traces and projection operators. We have seen that Hermitian matrices provide us with orthonormal bases of eigenvectors. A matrix U is called unitary if $U^\dagger U = UU^\dagger = \mathbb{1}$ or $U^\dagger = U^{-1}$. Different orthonormal bases $\{|a_i\rangle\}$ and $\{|b_j\rangle\}$ are related by a unitary transformation $U_{ij} = \langle a_i|b_j\rangle$ because

$$|b_j\rangle = (\sum_i |a_i\rangle\langle a_i|)\,|b_j\rangle = \sum_i |a_i\rangle U_{ij}, \quad \Rightarrow \quad U_{ij}(U^\dagger)_{jk} = \sum_j \langle a_i|b_j\rangle \cdot \langle b_j|a_k\rangle = \langle a_i|\mathbb{1}|a_k\rangle = \delta_{ik}, \tag{3.36}$$

where we used the completeness relation. In other words, the inverse change of basis is given by $\langle b_j|a_k\rangle = \langle a_k|b_j\rangle^* = (U_{kj})^* = (U^\dagger)_{jk} = U^{-1}_{jk}$.

Projection operators in quantum mechanics are always meant to be othogonal projections and they are characterized by the two conditions

$$P = P^\dagger \quad \text{and} \quad P^2 = P. \tag{3.37}$$

It follows from our previous considerations that projectors satisfy these equations. In turn, Hermiticity $P = P^\dagger$ implies the existence of a spectral representation $P_i = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i|$ and $P^2 = P$ implies $\lambda_i^2 = \lambda_i$ so that all eigenvalues are either 0 or 1. Hence $P_i = \sum' |\lambda_i\rangle\langle\lambda_i|$ is a sum of projectors $|\lambda_i\rangle\langle\lambda_i|$ onto one-dimensional subspaces spanned by $|\lambda_i\rangle$ where the sum $\sum'$ extends over the subset of basis vectors with eigenvalue 1. While the eigenvectors for a degenerate eigenvalue a_i of a matrix A in the spectral representation (3.34) are only defined up to a unitary change of basis of the respective eigenspace, the projector $P_i = \sum_{l_i} |a_i, l_i\rangle\langle a_i, l_i|$ onto such an eigenspace is independent of the choice of the orthonormal eigenvectors $|a_i, l_i\rangle$.

The axioms of quantum mechanics imply that measurements of the observable corresponding to a self-adjoint operator A yield the eigenvalue a_i with probability $\mathcal{P}(a_i) = |\langle\psi|a_i\rangle|^2$ if the state

of the system is described by the normalized vector $|\psi\rangle \in \mathcal{H}$. The resuling expectation value, i.e. the mean value $\langle A\rangle = \sum a_i \mathcal{P}(a_i)$ of the measured values weighted by their probabilities, is in accord with the definition (3.21) because the spectral representation of A implies

$$\langle\psi|A|\psi\rangle = \langle\psi| \sum_i a_i |a_i\rangle\langle a_i| \, |\psi\rangle = \sum a_i \, |\langle\psi|a_i\rangle|^2. \tag{3.38}$$

The *trace* of a matrix is the sum of its diagonal elements and can be written as

$$\mathrm{tr}\, A = \sum_i A_{ii} = \sum_i \langle e_i|A|e_i\rangle \tag{3.39}$$

for any orthonormal basis $|e_i\rangle$. An important property of traces is their invariance under cyclic permutations,

$$\mathrm{tr}(AB) = \mathrm{tr}(BA) \qquad \Rightarrow \qquad \mathrm{tr}(A_1\, A_2 \ldots A_{r-1}\, A_r) = \mathrm{tr}(A_r\, A_1\, A_2 \ldots A_{r-1}). \tag{3.40}$$

Probabilities and expectation values can be written in terms of traces and projection operators, which often simplifies calculations. Insertion of the definition $P_i = |a_i\rangle\langle a_i|$ shows that

$$\langle a_i|A|a_i\rangle = \mathrm{tr}(P_i A) = \mathrm{tr}(A P_i), \quad \Rightarrow \quad \mathcal{P}(a_i) = \langle\psi|P_i|\psi\rangle = \mathrm{tr}(P_i P_\psi), \tag{3.41}$$

where $P_\psi = |\psi\rangle\langle\psi|$ projects onto the one-dimensional space spanned by the normalized state vector $|\psi\rangle$. The second formula $\mathcal{P}(a_i) = \mathrm{tr}\; P_i P_\psi$ also holds for the probability of the measurement of the degenerate eigenvalue a_i if we use the projector $P_i = \sum_{l_i} |a_i, l_i\rangle\langle a_i, l_i|$ onto the complete eigenspace.

Commutators and anti-commutators. The commutator $[A, B]$ of two operators is defined as the difference between the two compositions $AB \equiv A \circ B$ and $BA \equiv B \circ A$,

$$\boxed{[A, B] = A\, B - B\, A \qquad \Rightarrow \qquad [A, B] = -[B, A].} \tag{3.42}$$

In the finite dimensional case it is just the difference between the matrix products AB and BA. We will often be in the situation that we know the commutators among a basic set $A, B, \ldots$ of operators, like the position operator $X_i = x_i$ and the momentum operator $P_i = \frac{\hbar}{i}\frac{\partial}{\partial x_i}$

$$\boxed{[X_i, P_j] = i\hbar\delta_{ij}.} \tag{3.43}$$

This can be verified by application to an arbitrary wave function

$$[X_i, P_j]|\psi\rangle = (X_i P_j - P_j X_i)\psi(x) = \tfrac{\hbar}{i}(x_i\partial_j\psi(x) - \partial_j(x_i\psi(x)) = -\tfrac{\hbar}{i}(\partial_j x_i)\psi(x) = i\hbar\delta_{ij}|\psi\rangle. \tag{3.44}$$

If we want to compute commutators for composite operators like the Hamilton operator $H = \frac{1}{2m}P^2 + \ldots$ one should then always use the identities

$$\boxed{[A, BC] = [A, B]C + B[A, C], \qquad [AB, C] = [A, C]B + A[B, C]} \tag{3.45}$$

rather than inserting and evaluating all the terms on a wave function and trying to recombine the result to an operator expression. (3.45) is easily verified by expanding the definitions

$$[A, BC] = ABC - BCA, \qquad [A, B]C + B[A, C] = (AB - BA)C + B(AC - CA) \tag{3.46}$$

and similarly for $[AB, C]$. These identities can be memorized as the Leibniz rule for the action of $[A, *]$ on a product BC and a similar product rule for the action of $[*, C]$ on the product AB "from the right". This "Leibniz rule" also holds for the action of $[A, *]$ on a commutator $[B, C]$ and for the action of $[*, C]$ on $[A, B]$

$$[A, [B, C]] = [[A, B], C] + [B, [A, C]], \qquad [[A, B], C] = [[A, C], B] + [A, [B, C]]. \tag{3.47}$$

Each of these equations is equivalent to the *Jacobi identity*

$$\boxed{[A, [B, C]] + [C, [A, B]] + [B, [C, A]] = 0,} \tag{3.48}$$

which states the sum over the cyclic permutations of ABC in a double commutator is zero. This is again easily verified by expanding all terms

$$A(BC-CB)-(BC-CB)A+B(CA-AC)-(CA-AC)B+C(AB-BA)-(AB-BA)C = 0. \tag{3.49}$$

The equivalence of the "product rule" (3.47) with the Jacobi identity follows from the antisymmetry of the commutator $[A, B] = -[B, A]$.

Similarly to the commutator we can define the anti-commutator

$$\boxed{\{A, B\} = AB + BA \qquad \Rightarrow \{A, B\} = \{B, A\}.} \tag{3.50}$$

For two Hermitian operators $A = A^\dagger$ and $B = B^\dagger$ the commutator is anti-Hermitian and the anti-commutator is Hermitian,

$$[A, B]^\dagger = (AB - BA)^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger = BA - AB = -[A, B], \tag{3.51}$$

$$\{A, B\}^\dagger = (AB + BA)^\dagger = B^\dagger A^\dagger + A^\dagger B^\dagger = BA + AB = \{A, B\}. \tag{3.52}$$

Since iC is Hermitian if C is anti-Hermitian the decomposition

$$AB = \tfrac{1}{2}(AB + BA) + \tfrac{1}{2}(AB - BA) = \tfrac{1}{2}\{A, B\} + \tfrac{1}{2}[A, B] \tag{3.53}$$

of an operator product AB as a sum of a commutator and an anti-commutator corresponds to a decomposition into real and imaginary part for products of Hermitian operators.

Complete systems of commuting operators. We show that two self-adjoint operators A and B commute $AB = BA$ if and only if they can be diagonalized simultaneously. Since

diagonal matrices commute, it is clear that $[A, B] = 0$ if there exists a basis such that both operators are diagonal. In order to proof the "only if" we assume that $[A, B] = 0$ and that A has been diagonalized. Then B must be block-diagonal because

$$0 = \langle a_i|[A, B]|a_j\rangle = \langle a_i|AB - BA|a_j\rangle = (a_i - a_j)\langle a_i|B|a_j\rangle \tag{3.54}$$

so that all matrix elements of B between states with different eigenvalues of A vanish. B can now be diagonalized within each block, by a change of basis that does not mix eigenstates for different eigenvalues of A and hence does not spoil the diagonalization of A. It is clear from the proof that the proposition extends to an arbitrary number of mutually commuting operators. Moreover, we see that any set of mutually commuting operators can be extended to a complete set in the sense that the simultaneous diagonalization uniquely fixes the common normalized eigenvectors up to a phase (just add an operator that lifts the remaining degeneracies within the common eigenspaces of the original set). The set of all eigenvalues $(a_i, b_j, c_k, \ldots)$ of such a complete system $A, B, C, \ldots$ thus completely characterizes the state $|a_i, b_j, c_k, \ldots\rangle$ of a quantum system.

Functions of operators. If we consider the position vector $\vec{x}$ of a particle as a vector of operators $\vec{X}$ then the potential $V(X) = V(x)$ can be a complicated function of operators X_i. If such a function is analytic $f(x) = \sum_{n=0}^{\infty} c_n x^n$ then the corresponding function of operators can be defined by the power series expansion

$$f(x) = \sum c_n x^n \qquad \Rightarrow \qquad f(A) = \sum c_n A^n. \tag{3.55}$$

For matrices the series always converges if the radius r of convergence of the Taylor series is infinite. If $0 < r < \infty$ then $f(\mathcal{O})$ can be defined by analytic continuation of its matrix elements. Of particular importance is the exponential function

$$\boxed{e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n = \lim_{n\to\infty} (1 + \frac{1}{n} A)^n,} \tag{3.56}$$

which usually appears if we are interested in the finite form of infinitesimal transformations. For example, the infinitesimal time evolution of the wave function is given by the Schrödinger equation

$$\partial_t|\psi(x,t)\rangle = \frac{1}{i\hbar} H|\psi(x,t)\rangle \quad \Rightarrow \quad |\psi(x, t_0 + \delta t)\rangle = (1 + \frac{\delta t}{i\hbar} H + O(\delta t^2))|\psi(x, t_0)\rangle \tag{3.57}$$

For a *time-independent Hamiltonian* H we obtain, after n infinitesimal time steps $\delta t = (t - t_0)/n$ with $n \to \infty$,

$$\boxed{|\psi(x,t)\rangle = U(t - t_0)|\psi(x, t_0)\rangle, \qquad U(t - t_0) = e^{-\frac{i}{\hbar}(t-t_0)H}.} \tag{3.58}$$

$U(t)$ is called **time evolution operator**. It is, actually, a one-parameter family of operators satisfying $U(t_1)U(t_2) = U(t_1 + t_2)$ and $\partial_t U(t) = -\frac{i}{\hbar} H U(t)$ with $U(0) = \mathbb{1}$.

For operators A, B the product of exponentials is not the exponential of the sum if the operators do not commute. The correction terms are expressed by the **Baker–Campbell–Hausdorff** formula

$$\boxed{e^A\, e^B = e^{A+B+\frac{1}{2}[A,B]+\frac{1}{12}([A,[A,B]]-[B,[A,B]])+\text{ multiple commutators}}} \tag{3.59}$$

(for a proof consider example (1.21) in [Grau]). In many applications the double commutators $[A,[A,B]]$ and $[B,[A,B]]$ vanish or are proportional to $\mathbb{1}$ so that the series terminates after a few terms. In particular, since A and $-A$ commute, the exponential of an anti-Hermitian operator iA is unitary,

$$A = A^\dagger \qquad \Rightarrow \qquad (e^{iA})^\dagger = e^{-iA} = (e^{iA})^{-1}. \tag{3.60}$$

The Hamilton operator of a quantum system has to be self-adjoint because it corresponds to the energy, which is an observable.[5] Time evolution is hence described by a unitary transformation $U(t) = U^\dagger(-t)$. We have already checked this in chapter 2 by showing that $\langle\psi|\psi\rangle$ is preserved under time evolution for a nonrelativistic electron in an electromagnetic field. But the present discussion is more general. Another important formula

$$\boxed{e^{\lambda A} B e^{-\lambda A} = B + \sum_{n=1}^{\infty} \frac{\lambda^n}{n!} [A,B]_{(n)} = B + \lambda[A,B] + \frac{\lambda^2}{2}[A,[A,B]] + \ldots} \tag{3.61}$$

with $[A,B]_1 = [A,B]$ and $[A,B]_{(n+1)} = [A,[A,B]_{(n)}]$ desribes the "conjugation" UBU^{-1} of an operator B by the exponential $U = e^{\lambda A}$ of λA.[6]

Arbitrary functions of *Hermitian* operators can be defined via their spectral representation,

$$\boxed{A = A^\dagger = \sum a_i\, |a_i\rangle\langle a_i| \qquad \Rightarrow \qquad f(A) = \sum f(a_i)\, |a_i\rangle\langle a_i|} \tag{3.62}$$

For analytic functions f this coincides with the power series (3.55), as is easily checked in a basis where A is diagonal. The definition (3.55) only makes sense for analytic functions, but it has the advantage that it does not require diagonalizability. With (3.62), on the other hand, even the Heaviside step function $\theta(A)$ of an operator A makes sense.

Tensor products: If we have a quantum system that is composed of two subsystems, whose states are described by $|i\rangle \in V_1$ with $i = 1, \ldots, I$ and $|m\rangle \in V_2$ with $1 \le m \le M$, then

[5] In certain contexts, like the description of particle decay, it may nevertheless be useful to consider Hamilton operators with an imaginary part.

[6] Conjugation of operators corresponds to a change of orthonormal bases $|e\rangle \to U|e\rangle$, for which the dual basis transforms as $\langle e| \to U^\dagger\langle e| = U^{-1}\langle e|$.

the states of the composite systems are superpositions of arbitrary combinations

$$|i,m\rangle \;\equiv\; |i\rangle \otimes |m\rangle, \qquad 1 \le i \le I, \qquad 1 \le m \le M \tag{3.63}$$

of the independent states in the subsystems. The vector space $V_1 \otimes V_2$ describing the composite system is called tensor product and it consists of linear combinations

$$|w\rangle = \sum_{i=1}^{I} \sum_{m=1}^{M} w_{im}\, |i,m\rangle \;\in\; V_1 \otimes V_2 \tag{3.64}$$

with an arbitray matrix w_{im} of coefficients. Its dimension $\dim(V_1 \otimes V_2) = I \cdot M$ is the product of the dimensions of the factors V_1 and V_2. The Dirac notation is particularly useful for such composite systems because we just combine the respective quantum numbers into a longer *ket*-vector. It is a simple fact of linear algebra that generic vectors in a tensor product cannot be written as a product

$$|w\rangle = \sum w_{im}\, |i,m\rangle \neq |v_1\rangle \otimes |v_2\rangle \tag{3.65}$$

for any $|v_1\rangle = \sum c_i |i\rangle$ and $|v_2\rangle = \sum d_m |m\rangle$ because this is only possible if the coefficient matrix factorizes as $w_{im} = c_i d_m$ and hence has rank 1. In quantum mechanics non-product states like (3.65) are often called *entangled states*. They play an important role in discussions about the interpretation of quantum mechanics like in the EPR paradoxon (see below).

The inner product on the tensor product space is defined by

$$\langle i,m|j,n\rangle = \langle i|j\rangle \cdot \langle m|n\rangle \tag{3.66}$$

for product states and extended by semi-bilinearity to $V_1 \otimes V_2$. In the product basis $|i,m\rangle$ operators on a tensor product space also have double-indices

$$|i,m\rangle \quad\to\quad \mathcal{O}_{i,m;j,n}\, |j,n\rangle \tag{3.67}$$

Such operators will often correspond to the combined action of some operator $\mathcal{O}^{(1)}$ on V_1 and $\mathcal{O}^{(2)}$ on V_2, like for example the rotation of the position vector of the first particle and the simultaneous rotation of the position vector of the second particle for rotating the complete system. In that situation the trace of the product operator factorizes into a product of traces

$$\mathcal{O}_{i,m;j,n} = \mathcal{O}^{(1)}_{ij} \otimes \mathcal{O}^{(2)}_{mn} \quad\Rightarrow\quad \operatorname{tr} \mathcal{O}_{i,m;j,n} = \sum_{im} \mathcal{O}_{i,m;i,m} = \sum_{i} \mathcal{O}^{(1)}_{ii} \sum_{m} \mathcal{O}^{(2)}_{mm} = \operatorname{tr} \mathcal{O}^{(1)} \cdot \operatorname{tr} \mathcal{O}^{(2)}. \tag{3.68}$$

As an example consider $\mathcal{O}^{(1)} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $\mathcal{O}^{(2)} = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$ for $V_1 = V_2 = \mathbb{C}^2$. In the basis $e_1 = |11\rangle$, $e_2 = |12\rangle$, $e_3 = |21\rangle$ and $e_4 = |22\rangle$ of the product space the product operator corresponds to the insertion of the second matrix into the first,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a\mathcal{O}^{(2)} & b\mathcal{O}^{(2)} \\ c\mathcal{O}^{(2)} & d\mathcal{O}^{(2)} \end{pmatrix} = \begin{pmatrix} ae & af & be & bf \\ ag & ah & bg & bh \\ ce & cf & de & df \\ cg & ch & dg & dh \end{pmatrix}. \tag{3.69}$$

The Dirac notation is obviously more transparent than this. It is easy to verify (3.68) for (3.69).

3.3 Operators and Hilbert spaces

Recall that the normalizable solutions $\psi(x,t)$ of the Schrödinger equation form an inner product space, i.e. a vector space with a positive definite semi-bilinear product

$$\langle\psi|\varphi\rangle = \int d^3x \psi^*(x,t)\varphi(x,t). \tag{3.70}$$

Inner product spaces are also called pre-Hilbert spaces. Such a space is called Hilbert space if it is complete with respect to the norm

$$||\psi|| = \sqrt{\langle\psi|\psi\rangle}, \tag{3.71}$$

i.e. if every Cauchy sequence converges. Cauchy sequences are sequences ψ_n with the property that for every positive number ε there exists an integer $N(\varepsilon)$ with

$$||\psi_m - \psi_n|| < \varepsilon \qquad \forall m,n > N(\varepsilon). \tag{3.72}$$

Pre-Hilbert spaces can be turned into Hilbert spaces by a standard procedure called completion, which amounts to adding the missing limits. The standard Hilbert space of quantum mechanics is the space of square integrable functions called

$$L^2(\mathbb{R}^3). \tag{3.73}$$

The letter L stands for *Lesbeques integration*, which has to be used because Riemann's definition of integration only works for a restricted class of square-integrable functions $\int|\psi^2| < \infty$ that is not complete and the Lesbeques integral can be regarded as the result of the completion.[7] A Hilbert space basis is a set of vectors $|e_i\rangle$ with some (possibly not countable) index set I such that every vector $|\psi\rangle \in \mathcal{H}$ can be written as a convergent infinite sum

$$|\psi\rangle = \sum_{n=1}^{\infty} c_n |e_{i_n}\rangle \tag{3.74}$$

for a sequence c_n of coefficients and a sequence i_n of indices $i \in I$ and hence of basis vectors $|e_{i_n}\rangle$ taken from the complete set of basis elements $|e_i\rangle$. A Hilbert space is called separable if there exists a countable basis, i.e. if we can take the index set to be $I = \mathbb{N}$. All Hilbert spaces that we need in this lecture will be separable.

[7] For example, the integral of the function $I_{\mathbb{Q}}(x)$ that is 1 for rational numbers and 0 for irrational numbers x is 0 for Lesbeques integration, because $\mathbb{Q}$ is countable so that $I_{\mathbb{Q}}$ is the limit of a Cauchy sequence of function with only finitely many values different from 0. But the Riemann integral does not exist.

3.3.1 Inequalities

In this section we derive three inequalities that hold in any Hilbert space. Let us denote the vectors as $f, g, h, \ldots \in \mathcal{H}$. The orthogonal projection of f onto g is the vector $|g\rangle\frac{\langle g|f\rangle}{\langle g|g\rangle}$ with the projection vector

$$|h\rangle = |f\rangle - |g\rangle\frac{\langle g|f\rangle}{\langle g|g\rangle} \tag{3.75}$$

orthogonal to $|g\rangle$ since $\langle g|h\rangle = 0$. Now we use the definition equation of $|h\rangle$ to obtain the Pythagorean theorem

$$\|f\|^2 = \langle f|f\rangle = \left(\langle h| + \langle g|\frac{\langle g|f\rangle}{\langle g|g\rangle}\right)\left(|h\rangle + |g\rangle\frac{\langle g|f\rangle}{\langle g|g\rangle}\right) = \|h\|^2 + \frac{|\langle g|f\rangle|^2}{\langle g|g\rangle} \tag{3.76}$$

Since $\|h\|^2 \geq 0$ we see that:

$$\|f\|^2 \geq \frac{|\langle g|f\rangle|^2}{\|g\|^2} \tag{3.77}$$

and we obtain the **Schwartz inequality**

$$\|f\|\|g\| \geq |\langle g|f\rangle| \tag{3.78}$$

which will later be used in the derivation of Heisenberg's uncertainty relation.

More generally, we can consider a set $g_1, \ldots, g_n$ of orthonormal vectors $\langle g_i, g_j\rangle = \delta_{ij}$ and write f as a sum of orthogonal projections $|g_i\rangle\langle g_i|f\rangle$ and the difference vector

$$|h\rangle = |f\rangle - \sum_{i=1}^{n} |g_i\rangle\langle g_i|f\rangle, \tag{3.79}$$

which is orthogonal to the linear subspace spanned by the $|g_i\rangle$. The Pytagorean theorem thus becomes

$$||f||^2 = \sum_{i=1}^{n} |\langle g_i|f\rangle|^2 + ||h||^2 \tag{3.80}$$

and the **Bessel inequality**

$$||f||^2 \geq \sum_{i=1}^{n} |\langle g_i|f\rangle|^2 \tag{3.81}$$

follows from positivity of $||h||^2$. For a Hilbert space basis g_i, $i \in \mathbb{N}$ the norm of h thus has to converge to 0 monotonously from above for $n \to \infty$.

The norm of $|f\rangle + |g\rangle$ is

$$\|f + g\|^2 = \langle f + g|f + g\rangle = \|f\|^2 + \|g\|^2 + \langle f|g\rangle + \langle g|f\rangle \tag{3.82}$$

Since we can write the last two terms as

$$\langle f|g\rangle + \langle g|f\rangle = \langle f|g\rangle + (\langle f|g\rangle)^* = 2Re\langle f|g\rangle \leq 2|Re\langle f|g\rangle| \leq 2|\langle f|g\rangle| \tag{3.83}$$

we can use the Schwartz inequality in this relation and obtain

$$\|f+g\|^2 \leq \|f\|^2 + \|g\|^2 + 2||f||\,||g|| \tag{3.84}$$

whose square root yields the **triangle inequality**

$$\|f+g\| \leq \|f\| + \|g\|, \tag{3.85}$$

which shows that the definition (3.71) of the norm in inner product spaces makes sense.

3.3.2 Position and momentum representations

As compared to matrices in finite-dimensional vector spaces we will encounter two kinds of complications for operators in Hilbert spaces. Consider, for example, the Hamilton operator for the potential well. For negative energies we obtained a discrete spectrum of bound states. But for free electrons there are no normalizable energy eigenstates and normalizable wave packets are superpositions of states with a continuum of energy values. Hence, the spectrum of self-adjoint operators will, in general, consist of a discrete part and a continuum without *normalizable* eigenstates. Moreover, the eigenvalues may not even be bounded, which leads to additional complications.

As an example we first consider the momentum operator $P = \frac{\hbar}{i}\frac{\partial}{\partial x}$. Working, for simplicity, in one dimension we define

$$|p_x\rangle = \frac{1}{\sqrt{2\pi\hbar}}\, e^{\frac{i}{\hbar}px}, \qquad P|p_x\rangle = p|p_x\rangle, \tag{3.86}$$

where the argument x of the wave function is indicated as a subscript of the eigenvalue p if necessary. The normalization of the momentum eigenstates $|p\rangle$ have been chosen such that

$$\langle p'|p\rangle = \frac{1}{2\pi\hbar}\int dx\, e^{\frac{i}{\hbar}(p-p')x} = \delta(p'-p), \tag{3.87}$$

where we used $\int dx\, e^{ikx} = 2\pi\delta(k)$. In three dimensions $|\vec{p}\rangle = |p_1\rangle \otimes |p_2\rangle \otimes |p_3\rangle$ so that

$$|\vec{p}_{\vec{x}}\rangle = \frac{1}{(2\pi\hbar)^{3/2}}\, e^{\frac{i}{\hbar}\vec{p}\vec{x}} \qquad \text{and} \qquad \langle\vec{p}\,'|\vec{p}\rangle = \delta^3(\vec{p}\,'-\vec{p}). \tag{3.88}$$

The product

$$\langle p|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}}\int dx\, e^{-\frac{i}{\hbar}px}\psi(x) = \tilde{\psi}(p) \tag{3.89}$$

yields the Fourier transform[8] of the wave function and the validity of the formula for the Fourier representation

$$\int dp\, |p_x\rangle\langle p|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}}\int dp\, e^{+\frac{i}{\hbar}px}\langle p|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}}\int dp\, e^{+\frac{i}{\hbar}px}\tilde{\psi}(p) = \psi(x), \tag{3.90}$$

[8] The extra factor $1/\sqrt{\hbar}$ in the normalization, as compared to the conventions in section 2, is due the rescaled argument $p = \hbar k$ of the Fourier transform.

which holds for all $\psi \in L^2(\mathbb{R})$, implies the spectral representation

$$\int dp\, |p_x\rangle\langle p_{x'}| = \mathbb{1}_{x,x'} = \delta(x - x'), \tag{3.91}$$

but now with the sum over eigenvalues with normalizable eigenstates replaced by an integral over the continuum of eigenvalues with non-normalizable eigenfunctions. The spectral representation thus becomes

$$P = P\mathbb{1} = \int dp\, P|p\rangle\langle p| = \int dp\, p\, |p\rangle\langle p|. \tag{3.92}$$

For more general self-adjoint operators like the Hamilton operator of the potential well we hence anticipate a spectral representation that will combine a sum over bound state energies with an integral over continuum states.

Similarly to the momentum operator we can now introduce a basis of eigenstates for the position operator X, where we would like to have

$$X|x\rangle = x|x\rangle \qquad \text{with} \qquad \int dx\, |x\rangle\langle x| = \mathbb{1} \qquad \text{and} \qquad \langle x|x'\rangle = \delta(x - x'). \tag{3.93}$$

But what are the wave functions $\psi_x(x')$ corresonding to these states? Since $X|x\rangle = x|x\rangle$ the wave function $\psi_x(x')$ of the state $|x_{x'}\rangle$ should vanish for $x' \neq x$ and hence be proportional to $\delta(x' - x)$, i.e. $\psi_x(x') = c\delta(x' - x)$. This ansatz satisfies (3.93) if we choose the prefactor $c = 1$ so that $\psi_x(x') = \langle x|x'\rangle$. This should not come as a surprise if we recall from section 3.1 that we can obtain the components v^i of a vector $v = v^i e_i$ by evaluation of the dual basis vectors $v^i = e^i(v)$ and that the bra-vectors obtained by Hermitian conjugation of an orthnormal basis provide the dual basis. Hence, for an arbitrary state $|\psi\rangle \in \mathcal{H}$ the products

$$\psi(x) = \langle x|\psi\rangle, \qquad \psi(p) = \langle p|\psi\rangle \tag{3.94}$$

are the **wave functions** $\psi(x)$ in **position space** and $\psi(p)$ in **momentum space**, respectively.

We hence **regard $|\psi\rangle \in \mathcal{H}$ as an abstract vector in Hilbert space** independently of any choice of basis and write $\langle x|\psi\rangle = \psi(x)$ for the position space wave function and $\langle p|\psi\rangle = \psi(p)$ for the wave function in the momentum space basis $|p\rangle$. The "unitary matrix" $\langle x|p\rangle$ for the change of basis from position space to momentum space $|p\rangle = \int dx\, |x\rangle\langle x|p\rangle$ and its inverse $\langle p|x\rangle$ are

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}\, e^{\frac{ipx}{\hbar}}, \qquad \langle p|x\rangle = \frac{1}{\sqrt{2\pi\hbar}}\, e^{-\frac{ipx}{\hbar}}. \tag{3.95}$$

Since the spectra of eigenvalues of P and X and the corresponding "matrix indices" p and x are continuous, matrix multiplication amounts to integration and the change of basis becomes

$$\psi(p) = \langle p|\psi\rangle = \langle p|\mathbb{1}|\psi\rangle = \langle p|\left(\int dx\, |x\rangle\langle x|\right)|\psi\rangle = \int dx \langle p|x\rangle\, \psi(x), \tag{3.96}$$

which is nothing but a Fourier transformation.

The basis independence of the integrated probability density $||\psi||^2 = \langle\psi|\psi\rangle$

$$\begin{aligned}\int dx\,|\psi(x)|^2 &= \int dx\,\langle\psi|x\rangle\langle x|\psi\rangle = \int dx\,\langle\psi|\int dp\,|p\rangle\langle p|\,|x\rangle\langle x|\int dp'\,|p'\rangle\langle p'|\,|\psi\rangle \\ &= \int dp\int dp'\,\langle\psi|p\rangle\int dx\,\langle p|\,x\rangle\langle x|p'\rangle\,\langle p'|\psi\rangle = \int dp\,\langle\psi|p\rangle\langle p|\psi\rangle = \int dp\,|\psi(p)|^2 \end{aligned} \tag{3.97}$$

expresses the unitarity $\int dp\,\langle p|x\rangle\langle x|p'\rangle = \delta(p-p')$ of the matrix $\langle x|p\rangle$ of the change of basis. In Fourier analysis (3.97) is called Parseval's equation.

The matrix elements of X and P are now easily evaluated in position space

$$\langle x'|X|x\rangle = x\;\delta(x-x'), \qquad\qquad \langle x'|P|x\rangle = \frac{\hbar}{i}\frac{\partial}{\partial x}\;\delta(x-x'), \tag{3.98}$$

and in momentum space

$$\langle p'|P|p\rangle = p\;\delta(p-p'), \qquad\qquad \langle p'|X|p\rangle = -\frac{\hbar}{i}\frac{\partial}{\partial p}\;\delta(p-p'), \tag{3.99}$$

which shows that $X = -\frac{\hbar}{i}\frac{\partial}{\partial p}$ and $P = p$ in momentum space. The generalizations of these formulas to three dimensions are obvious.

3.3.3 Convergence, norms and spectra of Hilbert space operators

Having gained some intuition about spectra and eigenbases of Hilbert space operators we are now turning to general definitions and results. Already for the case of a discrete spectrum, like in the Harmonic oscillator for which eletrons are always bound, it is clear that the spectral decompositon of the identity

$$\hat{\mathbb{1}} = \lim_{n\to\infty}\;\sum_{i=1}^{n}|e_i\rangle\langle e_i| \tag{3.100}$$

requires some notion of convergence for sequences of operators in order to be able to define infinite sums as limits of finite sums.

Already for sequences of Hilbert space *vectors* there are two different notions of convergence: The obvious one, which we used for the definition of completeness, is called *strong convergence*:

$$|\psi_n\rangle \longrightarrow |\psi\rangle \qquad \text{if} \qquad \lim_{n\to\infty}||\psi_n - \psi|| = 0 \qquad\qquad \text{strong limit.} \tag{3.101}$$

A second notion of convergence, which is called *weak* because it is always implied by strong convergence (see section 4.8 of [Kreyszig]), only requires that all products with bra-vectors converge:

$$|\psi_n\rangle \xrightarrow{\text{weak}} |\psi\rangle \qquad \text{if} \qquad \lim_{n\to\infty}\langle\varphi|\psi_n\rangle = \langle\varphi|\psi\rangle \quad \forall\;\langle\varphi|\;\in\;\mathcal{H}^{dual} \qquad \text{weak limit.} \tag{3.102}$$

An example of a sequence that weakly converges to 0 but that is divergent in the strong sense is the sequence $\{e_n\}$ of basis vectors of a Hilbert space basis: A sequence pointing into the infinitely many directions of a Hilbert space with constant length 1 does not converge (in the norm) because the distance $||e_n - e_m||$ between any two elements of such a sequence is always $\sqrt{2}$. But the scalar products $\langle\varphi|e_n\rangle$, which are the expansion coefficients of $\langle\varphi|$ in the basis $\{\langle e_n|\}$, have to converge to 0 because of Bessels inequality.

Convergence of operators: For us, Hilbert space operators are always meant to be linear

$$\mathcal{O}(\alpha|\varphi\rangle + \beta|\psi\rangle) = \alpha\mathcal{O}|\varphi\rangle + \beta\mathcal{O}|\psi\rangle \in \mathcal{H}. \tag{3.103}$$

These operators are important in quantum mechanics because they correspond to observables. We now have two options: Every real measurement has a bounded set of possible results. For example, we can never measure the position of a particle, say, behind the Andromeda nebula, because our particle detector has a finite size. We could hence simply restrict the concept of an observable to bounded operators, which are quite well-behaved. But, like for wave packets and plane waves, it is much more convenient to work with unbounded operators like the position operator X rather than with more realistic approximations of this operators.

We hence first define the concept of the norm of an operator, which we can think of as the modulus $|\lambda|$ of the largest eigenvalue λ:

$$\|\mathcal{O}\| = \sup_{\psi\neq 0} \left(\frac{\|\mathcal{O}\psi\|}{\|\psi\|}\right). \tag{3.104}$$

In this definition we have to use the suppremum instead of the maximum because in the infinite-dimensional case the maximum may not exist and the suppremum (which is the smallest upper bound) may be infinite $0 \leq \|\mathcal{O}\| \leq \infty$. Considering a sequence ψ_n of localized waves packets for electrons whose distance from the earth increases, say, linearly with n, it is clear that X is unbounded $||X|| = \infty$, and similarly one can show that the momentum P is also unbounded.

An operator is called bounded if $||\mathcal{O}|| < \infty$. Bounded operators $\mathcal{O} : V \to W$ can, in fact, be defined for any normed spaces V and W. For two such operators we can consider linear combinations defined by

$$(\alpha\mathcal{O}_1 + \beta\mathcal{O}_2)|\psi\rangle = \alpha\mathcal{O}_1|\psi\rangle + \beta\mathcal{O}_2|\psi\rangle \quad \in \quad W \quad \text{for} \quad |\psi\rangle \in V, \tag{3.105}$$

so that the set of all bounded operators $\mathcal{B}(V, W)$ again forms a vector space. With the operator norm defined by (3.104) the normed space $\mathcal{B}(V, W)$ is complete and hence a Banach space. In this statement we refer to the strongest notion of convergens, which is called *uniform convergence* or *convergence in the norm*. For operators there are, however, even two different weaker notions of convergence: A **sequence of operators** $O_n : V \to W$ is said to be:

- **uniformly convergent** if $(\mathcal{O}_n)$ converges in the norm of $\mathcal{B}$, i.e.

$$\lim_{n\to\infty} \|\mathcal{O}_n - \mathcal{O}\| = 0, \tag{3.106}$$

- **strongly convergent** if $(\mathcal{O}_n\psi)$ converges strongly in W for every $\psi \in V$, i.e.

$$\lim_{n\to\infty} \|\mathcal{O}_n\psi - \mathcal{O}\psi\| = 0 \qquad \forall|\psi\rangle \in V \tag{3.107}$$

- **weakly convergent** if $(\mathcal{O}_n\psi)$ converges weakly in W for every $\psi \in V$, i.e.

$$\lim_{n\to\infty} |\langle\phi|\mathcal{O}_n\psi\rangle - \langle\phi|\mathcal{O}\psi\rangle| = 0 \qquad \forall\,|\psi\rangle \in V \text{ and } \forall\,\langle\phi| \in W^{dual}, \tag{3.108}$$

where $\mathcal{O}$ denotes the limit operator $\mathcal{O} : V \to W$. The notions of strong and weak operator convergence make perfect sense also for unbounded operators, and, moreover, $\mathcal{O}_n - \mathcal{O}$ may be bounded and uniformely convergent even if the operators $\mathcal{O}_n$ and $\mathcal{O}$ are unbounded.

Spectra and resolvents of operators. Naively we think of the spectrum of an operator as the set of eigenvalues λ of the matrix A, which coincides with the values λ for which

$$A_\lambda = A - \lambda\mathbb{1} \tag{3.109}$$

is not invertible so that $\det A_\lambda = 0$. In that case there exists an eigenvector $|a_\lambda\rangle$ with

$$A_\lambda\,|a_\lambda\rangle = 0 \qquad \Leftrightarrow \qquad A\,|a_\lambda\rangle = \lambda\,|a_\lambda\rangle. \tag{3.110}$$

The generalization to infinite dimensions is based on this fact and defines the spectrum as the set of complex numbers $\lambda \in \mathbb{C}$ for which A_λ is not invertible.

We have to take into account one further complication: For unbounded operators A it may happen that they are only defined on a subset of the Hilbert space vectors. As an example consider the position operator X and the wave function $\psi(x) = \theta(x)/\sqrt{1+x^2}$ where the step function $\theta(x)$ is 1 for $x > 0$ and 0 for $x < 0$. The integral $\int|\psi|^2 = \int_0^\infty \frac{dx}{(1+x^2)} = \frac{\pi}{2}$ exists, but $\langle\psi|X|\psi\rangle = \int_0^\infty \frac{x\,dx}{1+x^2}$ diverges. Hence $x\psi(x) \notin \mathcal{H} = L^2(\mathbb{R})$ and we have to restrict the domain of definition of X to a subset $\mathcal{D}_X \subset \mathcal{H}$ if we want X to be an operator with values in $\mathcal{H}$.

We hence consider operators $A : \mathcal{D}_A \longrightarrow \mathcal{H}$ whith domain of definition $\mathcal{D}_A \subseteq \mathcal{H}$ and assume that $\mathcal{D}_A$ is dense in $\mathcal{H}$, which means that every vector $|\psi\rangle \in \mathcal{H}$ can be obtained as a limit of a sequence $\psi_n \in \mathcal{D}_A$ of vectors in the domain of definition.[9] We now define the **resolvent** R_λ, if it exists, as the inverse of $A_\lambda = A - \lambda\mathbb{1}$, i.e.

$$R_\lambda = A_\lambda^{-1} \tag{3.111}$$

[9] For the position operator X we can take the sequence $\psi_n(x) = \theta(n - |x|) \cdot \psi(x)$.

with $R_\lambda A_\lambda = \mathbb{1}$ on $\mathcal{D}_A$. The resolvent R_λ hence is a linear operator from the range of A_λ to the domain of A_λ. It does not exist if and only if there exists a vector $|\psi\rangle \in \mathcal{D}_A$ with $A_\lambda|\psi\rangle = 0$. In that case $|\psi\rangle$ is an eigenvector of A with eigenvalue λ.

A complex number $\lambda \in \mathbb{C}$ is called a **regular value** if the resolvent R_λ exists as a bounded operator and λ is called **spectral value** otherwise. The set of regular values is called **resolvent set** $\rho(A) \subset \mathbb{C}$ and its complement $\sigma(A) = \mathbb{C} - \rho(A)$ is called **spectrum** of the operator A. The spectrum $\sigma(A)$ consists of three disjoint parts:

- The **point spectrum** or **discrete spectrum** $\sigma_p(A)$ is the set of values λ such that R_λ does not exist. $\sigma_p(A)$ is the set of **eigenvalues** of A (with normalizable eigenstates; this corresponds to the bound state energies for the Hamilton operator).

- The **continuous spectrum** $\sigma_c(A)$ is the set of values λ such that R_λ exists and is defined on a set which is dense is $\mathcal{H}$, but is not bounded (for the Hamilton operator this corresponds to the energies of scattering states).

- The **residual spectrum** $\sigma_r(A)$ is the set of λ such that R_λ exists but the domain of definition is not dense in $\mathcal{H}$.

We thus obtain a decomposition of the complex plane as a disjoint union of four sets $\mathbb{C} = \rho(A) \cup \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A)$. From the definition it follows that the resolvent set is open and one can show that the resolvent R_λ is an (operator valued) holomorphic function on $\rho(A)$, so that methods of complex analysis can be used in spectral theory [Reed]. In finite dimensional cases the spectrum of a linear operator is a pure point spectrum, i.e. $\sigma_c(A) = \sigma_r(A) = \emptyset$. For self-adjoint operators it can be shown that the residual spectrum is empty $\sigma_r(A) = \emptyset$.

3.3.4 Self-adjoint operators and spectral representation

A densely defined Hilbert space operator A is called **symmetric** (or **Hermitian**) if its domain of definition is contained in the domain of definition of the adjoint operator[10]

$$\mathcal{D}_A \subseteq \mathcal{D}_{A^\dagger} \qquad \text{and} \qquad \langle A\varphi|\psi\rangle = \langle\varphi|A\psi\rangle \qquad \forall \varphi, \psi \in \mathcal{D}_A. \tag{3.112}$$

A symmetric operator is called **self-adjoint** if $\mathcal{D}_A = \mathcal{D}_{A^\dagger}$.

The difference between symmetric and self-adjoint hence is based on the domain of definition. If the domain of definition $\mathcal{D}_{A^\dagger}$ of the adjoint operator $A^\dagger$, which is defined by $\langle A^\dagger\varphi|\psi\rangle = \langle\varphi|A\psi\rangle$, is smaller than $\mathcal{D}_A$, then we first have to restict the definition of A to

[10] It can be shown that $\mathcal{D}_{A^\dagger}$ consists of all vectors $\varphi \in \mathcal{H}$ for which $(|\langle A\psi|\varphi\rangle|) \,/\, (||\psi||)$ is (uniformly) bounded for all $\psi \in \mathcal{D}_A$ with $|\psi\rangle \neq 0$; see e.g. section VIII.1 of [Reed].

a subset of $\mathcal{D}_A$, which will at the same time increase $\mathcal{D}_{A^\dagger}$. If A thus becomes (or already is) a symmetric operator then we can ask the question whether it is possible to extend $\mathcal{D}_A$ such that A becomes self-adjoint. This question has been answered by a theorem first stated (for second order differential operators) by Weyl in 1910, and generalized by John von Neumann in 1929:

Self-adjoint extension of operators: The existence of a self-adjoint extension depends on the so-called deficiency indices $n_\pm$ of A, which are the dimensions of the eigenspaces $N_\pm$ of $A^\dagger$ for some fixed positive and negativ imaginary eigenvalues $\pm i\varepsilon$, respectively,

$$A^\dagger\psi = \pm i\varepsilon\psi, \qquad \varepsilon > 0, \tag{3.113}$$

where one may set, for example, $\varepsilon = 1$. Depending on these indices there are three cases:

- If $n_+ = n_- = 0$ then the operator A is already self-adjoint.
- If $n_+ = n_- \geq 1$ then A is not self-adjoint but admits infinitely many self-adjoint extensions.
- If $n_+ \neq n_-$ then a self-adjoint extension of A does not even exist.

A detailed discussion of simple examples for these situations can be found in a paper by Bonneau, Faraut and Valent.[11]

Spectral theorem. The content of the spectral theorem is that self-adjoint operators A are essentially multiplication operators in an appropriate eigenbasis, i.e. there exists a decomposition of unity as a sum of projection operators with the direction of the projections aligned with the eigenspaces of A,

$$\mathbb{1} = \sum_i P_i, \qquad A = \sum_i a_i\, P_i. \tag{3.114}$$

In the infinite-dimensional case of self-adjoint operators in Hilbert spaces the first complication is that the spectrum may consist of discrete and continuous parts, so that the sum has to be generalized to an operation that can at the same time describe sums and integrals. This is achieved by the Riemann-Stilties integral, which allows to assign different weights to different parts of the integration intervall. Assume that $\mu(x)$ is a monotonously increasing function with only isolated discontinuities. Then we think of the mass density given by the derivative $d\mu = \mu'\, dx$ which has (positive) δ-function like concentrations at the discontinuities of μ and the **Riemann–Stiltjes integral** for smooth functions can be written as

$$\int_a^b f(x)\, d\mu(x) = \int_a^b f(x)\, \mu'(x)\, dx \tag{3.115}$$

[11] G. Bonneau, J. Faraut, G. Valent, *Self-adjoint extensions of operators and the teaching of quantum mechanics,* Am.J.Phys. 69 (2001) 322 [http://arxiv.org/abs/quant-ph/0103153]

where we include, by convention, the contribution of δ-functions located at the upper integration limit

$$\int_a^b f(x)\,d\mu(x) = \lim_{\varepsilon\to 0+}\int_a^{b+\varepsilon} f(x)\,d\mu(x) \tag{3.116}$$

and accordingly drop point-like contributions at the lower limit to make the whole integral additive for intervals. The extension of the definition of the integral for non-smooth integrands $f(x)$ then proceeds like for the case of Riemann integration by taking limits of upper and lower bounds. Using the methods of measure theory this can be generalized to the (Lesbeques–) Stiltjes integral, allowing general *measurable functions* to be integrated.

The application of Stiltjes integration to spectral theory introduces the concept of a spectral family $\{\mathbb{E}_\lambda\}$ associated with an operator A, which is the one-parameter family of sums/integrals of the projectors for all spectral values up to a certain number $\lambda \in \mathbb{R}$. At first we assume that A is bounded, so that its spectrum is contained in an interval $\lambda \in [a, b]$. $\mathbb{E}_\lambda$ grows monotonically in λ and is a family of projectors that is continuous from above, i.e. one can show [Kreyszig]

$$\forall \nu \geq \lambda \quad : \quad \mathbb{E}_\nu \geq \mathbb{E}_\lambda, \tag{3.117}$$

$$\forall \lambda < a \quad : \quad \mathbb{E}_\lambda = 0, \tag{3.118}$$

$$\forall \lambda > b \quad : \quad \mathbb{E}_\lambda = \mathbb{1}, \tag{3.119}$$

$$\lim_{\nu\to\lambda+} \mathbb{E}_\nu \quad = \quad \mathbb{E}_\lambda. \tag{3.120}$$

The **theorem of Stone** then asserts that a **bounded self-adjoint operator** A has the spectral representation

$$A = \int_{a-}^b \lambda\, d\mathbb{E}_\lambda \tag{3.121}$$

for a spectral family $\mathbb{E}_\lambda$, where the Riemann–Stiltjes integral is uniformly convergent (with respect to the operator norm). The lower limit $a-$ indicates that we have to include the δ-function contribution at $\lambda = a$ if a is part of the discrete spectrum.

Unbounded and unitary operators.[12] The spectral theorem can now be extended to unbounded operators using the Cayley transformation to a unitary operator

$$U = (A - i\mathbb{1})(A + i\mathbb{1})^{-1} \tag{3.122}$$

where the resolvent $(A + i\mathbb{1})^{-1}$ exists as a bounded operator because the spectrum of A is real.

The spectral decomposition for unitary operators follows from the fact that we can decompose them into a commuting set of self-adjoint operators $V = \frac{1}{2}(U + U^\dagger) = V^\dagger$ and

[12] The most general family of operators for which a spectral decomposition exists are the normal operators, defined by the equation $NN^\dagger = N^\dagger N$, i.e. N commutes with its adjoint, which obviously covers both the self-adjoint and the unitary case. Normal operators can also be characterized by the fact that they are unitarily diagonalizable.

$W = \frac{1}{2i}(U - U^\dagger) = W^\dagger$ which commute because

$$U = V + iW,\ UU^\dagger = U^\dagger U \quad \Rightarrow \quad V^2 - i(VW - WV) + W^2 = V^2 + i(VW - WV) + W^2 \tag{3.123}$$

so that $VW - WV = 0$. Hence they have a spectral decomposition with a common spectral family. Putting together real and imaginary part of the eigenvalues of U we find

$$U = \int_{-\pi}^{\pi} e^{i\theta} d\mathbb{E}_\theta. \tag{3.124}$$

The spectrum is located on the unit circle and the convergence of the integral is again uniform. For unbounded operators A the Cayley transform can now be inverted with the formula

$$A = i(\mathbb{1} + U)(\mathbb{1} - U)^{-1} \tag{3.125}$$

and we obtain

$$A = \int_{-\pi}^{\pi} \tan\left(\tfrac{\theta}{2}\right)\, d\mathbb{E}_\theta = \int_{-\infty}^{\infty} \lambda\, d\mathbb{F}_\lambda \tag{3.126}$$

with the appropriate change of measure density in the spectral family. Since the spectrum can be unbounded the Stiltjes integral is now defined in the sense of strong operator convergence.

3.4 Schrödinger, Heisenberg and interaction picture

We now return to the issue of time dependence in quantum mechanics, which we described so far by the time dependence of states

$$\psi(x,t) = \langle x|\psi(t)\rangle \in L^2(\mathbb{R}^3) \quad \text{for } t \geq t_0, \tag{3.127}$$

i.e. by time-dependent vectors in Hilbert space that are determined by some initial condition $\psi(x, t_0)$ at an initial time and by solving the Schrödinger equation for later times. Since the map $\psi(x, t_0) \to \psi(x, t)$ is linear it defines a linear operator $U(t, t_0)$

$$|\psi(t)\rangle = U(t, t_0)\, |\psi(t_0)\rangle \tag{3.128}$$

called **time evolution operator**. More precisely $U(t, t_0)$ is a family of operators depending on two parameters, the initial time t_0 and the final time t, where we can also consider $t < t_0$ by solving the Schrödinger equation backwards in time. If we choose some orthonormal basis $|e_i(t_0)\rangle$ at an initial time then $|e_i(t)\rangle$ also forms an orthonormal basis at later times: The normalization $\langle e_i(t)|e_i(t)\rangle = \langle e_i(t_0)|e_i(t_0)\rangle = 1$ expresses the conservation of probability, and orthogonality at later times follows from the general fact that conservation of norms implies conservation of scalar products: Since

$$\begin{aligned} ||e_1 + e_2||^2 &= \langle e_1 + e_2|e_1 + e_2\rangle = ||e_1||^2 + ||e_2||^2 + \langle e_1|e_2\rangle + \langle e_2|e_1\rangle \\ ||e_1 + ie_2||^2 &= \langle e_1 + ie_2|e_1 + ie_2\rangle = ||e_1||^2 + ||e_2||^2 + i\langle e_1|e_2\rangle - i\langle e_2|e_1\rangle. \end{aligned} \tag{3.129}$$

the scalar product $\langle e_1|e_2\rangle$ can be reconstructed from the norm by solving the equations (3.129) for $\langle e_1|e_2\rangle$ and $\langle e_2|e_1\rangle$, which become complex conjugates of one another because their sum is real and their difference imaginary. We conclude that orthonormal bases stay orthonormal (and complete) during time evolution, so that the time evolution operator $U(t, t_0)$ amounts to a change of basis and hence is a unitary operator $U^\dagger = U^{-1}$.

Inserting the definition (3.128) into the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi(t)\rangle \tag{3.130}$$

and using $\frac{\partial}{\partial t}|\psi(t_0)\rangle = 0$ we obtain

$$i\hbar\left(\frac{\partial}{\partial t}U\right)|\psi(t_0)\rangle = HU|\psi(t_0)\rangle. \tag{3.131}$$

Since this relation has to hold for all $|\psi(t_0)\rangle$ it implies the operator differential equation

$$i\hbar\frac{\partial U}{\partial t} = HU. \tag{3.132}$$

If the Hamiltonian does not explicitly depend on time this equation can be solved formally and we obtain

$$U(t, t_0) = U(t - t_0) = e^{-\frac{i}{\hbar}H(t-t_0)}, \tag{3.133}$$

which only depends on the time difference $t - t_0$.

The Heisenberg picture. With the time evolution operator we can now write expectation values of operators as

$$\begin{aligned}\langle A\rangle &= \langle\psi(t)|A|\psi(t)\rangle = \langle\psi(t_0)|U^\dagger AU|\psi(t_0)\rangle \\ &= \langle\psi(t_0)|A_H|\psi(t_0)\rangle \qquad \text{with} \qquad A_H(t) = U^\dagger(t)AU(t)\end{aligned} \tag{3.134}$$

where we assume that A does not explicitly depend on time.

So far we discussed quantum mechanics in terms of the so-called **Schrödinger picture**, in which the time dependence of the system is governed by the Schrödinger equation for a time dependent wave function and operators are time independent, at least if the apparatus is not moved and if there are no other external sources of time dependence,

$$|\psi_S(t)\rangle \equiv |\psi(t)\rangle, \qquad A_S \equiv A. \tag{3.135}$$

But since all observable quantities in quantum mechanics can be expressed in terms of expectation values, eq. (3.134) shows that we can take a different point of view and interpret the time evolution as acting on the operators according to

$$A_H(t) = U^\dagger(t, t_0)A(t_0)U(t, t_0) \tag{3.136}$$

while the states do not change

$$|\psi_H(t)\rangle = |\psi(t_0)\rangle. \tag{3.137}$$

The description of quantum mechanics in terms of $A_H(t)$ and $|\psi_H\rangle$ is called **Heisenberg picture**, whereas the descrition in terms of A_S and $|\psi_S(t)\rangle$ is called Schrödinger picture and our definitions imply

$$\langle A\rangle = \langle\psi_S(t)|A_S|\psi_S(t)\rangle = \langle\psi_H|A_H(t)|\psi_H\rangle \tag{3.138}$$

where the two pictures are related by a unitary transformation

$$|\psi_S(t)\rangle = U|\psi_H\rangle \quad , \quad |\psi_H\rangle = U^\dagger|\psi_S(t)\rangle \tag{3.139}$$

$$A_S = UA_H(t)U^\dagger \quad , \quad A_H(t) = U^\dagger A_S U \tag{3.140}$$

with $U = e^{-\frac{i}{\hbar}H(t-t_0)}$ if H is time-independent.

While the Schrödinger picture seems to be more intuitive at first glance, the Heisenberg picture shows a formal similarity with classical mechanics: Since $\partial_t U = -\frac{i}{\hbar}HU$ and $\partial_t U^\dagger = \frac{i}{\hbar}U^\dagger H$ the infinitesimal time evolution of the Heisenberg operators is

$$\frac{\partial A_H}{\partial t} = \frac{\partial U^\dagger}{\partial t} A_S\, U + U^\dagger \overbrace{\frac{\partial A_S}{\partial t}}^{\equiv 0} U + U^\dagger\, A_S \frac{\partial U}{\partial t} \tag{3.141}$$

$$= \tfrac{i}{\hbar}(U^\dagger H_S A_S U - U^\dagger A_S H_S U) \tag{3.142}$$

$$= \tfrac{i}{\hbar}\left((U^\dagger H_S U)(U^\dagger A_S U) - (U^\dagger A_S U)(U^\dagger H_S U)\right), \tag{3.143}$$

where we inserted $\mathbb{1} = UU^\dagger$. We thus obtain **Heisenberg's equation of motion**

$$\boxed{\frac{\partial A_H}{\partial t} = \frac{i}{\hbar}[H_H, A_H],} \tag{3.144}$$

which has a formal similarity to Hamilton's equations of motion $\dot{f} = \{H, f\}_{PB}$ for phase space functions $f(p,q)$, or $\frac{\partial p_i}{\partial t} = \{H, p_i\}_{PB} = (-\frac{\partial H}{\partial q_i})$ and $\frac{\partial q_i}{\partial t} = \{H, q_i\}_{PB} = \frac{\partial H}{\partial p_i}$ for coordinates and momenta; $\{\cdots\}_{PB}$ denotes the Poisson bracket. This analogy is the starting point for the general quantization rules of Hamiltonian systems.

The interaction picture (or **Dirac picture**) combines elements of the Schrödinger picture and of the Heisenberg picture so that states and operators both become time dependent. It is the starting point for approximation techniques and useful if we can write the Hamitonian as a sum of a simple (exactly solvable) time-independent part H_0 and a (small) time-dependent perturbation $V(t)$,

$$H(t) = H_0 + V(t). \tag{3.145}$$

The idea is to put the simple part of the time evolution into the time dependence of operators $A_I(t)$, thereby obtaining a modified Schrödinger equation for the time evolution of states,

which only leads to a relatively small time dependence of $|\psi_I(t)\rangle$ due to a possibly complicated but weak interaction term $V(t)$. The interaction picture is thus defined by the unitary transformation

$$|\psi_I(t)\rangle = U_0^\dagger(t)|\psi_S(t)\rangle, \qquad A_I(t) = U_0^\dagger(t)\, A_S\, U_0(t) \tag{3.146}$$

with

$$U_0(t) = e^{-\frac{i}{\hbar}(t-t_0)H_0} \tag{3.147}$$

so that

$$\langle A\rangle = \langle\psi_S(t)|A_S|\psi_S(t)\rangle = \langle\psi_I(t)|A_I(t)|\psi_I(t)\rangle \tag{3.148}$$

The Schrödinger equation in the interaction picture is now obtained by evaluating

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}|\psi_I(t)\rangle &= i\hbar\frac{\partial}{\partial t}\left(e^{\frac{i}{\hbar}H_0(t-t_0)}|\psi_S(t)\rangle\right) = -U_0^\dagger(t)H_0|\psi_S(t)\rangle + i\hbar U_0^\dagger(t)\frac{\partial}{\partial t}|\psi_S(t)\rangle & (3.149)\\ &= U_0^\dagger(t)\left(-H_0 + H_0 + V(t)\right)|\psi_S(t)\rangle = U_0^\dagger(t)V(t)U_0(t)\; U_0^\dagger(t)|\psi_S(t)\rangle, & (3.150)\end{aligned}$$

where we used the Schrödinger equation $\partial_t|\psi_S(t)\rangle = -\frac{i}{\hbar}(H_0 + V(t))|\psi_S(t)\rangle$, so that

$$\boxed{i\hbar\frac{\partial}{\partial t}|\psi_I(t)\rangle = V_I(t)|\psi_I(t)\rangle} \tag{3.151}$$

Replacing H by H_0 and U by U_0 in the derivation of Heisenberg's equation of motion we obtain the operator equation of motion

$$\boxed{\frac{\partial A_I(t)}{\partial t} = \frac{i}{\hbar}[H_{0,I}(t), A_I(t)]} \tag{3.152}$$

in the interaction picture. The time evolution operator $U_I(t) = U_0^\dagger(t)U(t)U_0(t)$ in the interaction picture desribes the time evolution of the states $|\psi_I(t)\rangle = U_I(t,t_0)|\psi_I(t_0)\rangle$ and hence satisfies the equation of motion

$$i\hbar\frac{\partial U_I}{\partial t} = V_I(t)U_I. \tag{3.153}$$

In many situations $U_I(t)$ can only be computed by approximation procedures.

3.5 Ehrenfest theorem and uncertainty relations

In this section we want to improve our understanding of the relation between quantum mechanics and classical mechanics. The content of the Ehrenfest theorem is that expectation values of observables obey classical equations of motion. Heisenberg's uncertainty relation, on the other hand, implies limitations to the validity of classical concepts.

Let us compute the time evolution

$$\frac{d}{dt}\langle A\rangle = \frac{\partial\langle\psi|}{\partial t}\, A|\psi\rangle + \langle\psi|\frac{\partial A}{\partial t}\,|\psi\rangle + \langle\psi|A\frac{\partial|\psi\rangle}{\partial t} \tag{3.154}$$

of the mean value of an observable A in the Schrödinger picture. The Schrödinger equation yields

$$\frac{d}{dt}\langle A\rangle = \langle\frac{\partial A}{\partial t}\rangle + \frac{i}{\hbar}\langle H\psi|A|\psi\rangle - \frac{i}{\hbar}\langle\psi|A|H\psi\rangle = \langle\frac{\partial A}{\partial t}\rangle + \frac{i}{\hbar}\langle\psi|H\,A\;-\;A\,H|\psi\rangle \tag{3.155}$$

so that

$$\boxed{\frac{d}{dt}\langle\psi|A|\psi\rangle = \langle\psi|\frac{\partial A}{\partial t}|\psi\rangle + \frac{i}{\hbar}\langle\psi|\,[H,A]\,|\psi\rangle} \tag{3.156}$$

The Ehrenfest theorem states that the mean values of certain quantum mechanical operators obey the classical relations.

As an example let us compute the time evolution $\frac{d}{dt}\langle X_i\rangle$ of the position operator X_i, which is time independent $\frac{\partial X}{\partial t}$ in the Schrödinger picture. We first have to compute the commutator of X with the Hamiltonian $H = \frac{1}{2m}P_jP_j + V(\vec{x})$.

$$[H,X_i] = \frac{1}{2m}[P_jP_j,X_i] + \overbrace{[V(\vec{x}),X_i]}^{\equiv 0} = \frac{1}{2m}(P_j[P_j,X_i] + [P_j,X_i]P_j) = \frac{1}{2m}\,2\frac{\hbar}{i}P_i. \tag{3.157}$$

Inserting this result into the formula (3.156) for the mean value of an operator we obtain

$$\frac{d}{dt}\langle X_i\rangle = \frac{1}{m}\langle P_i\rangle. \tag{3.158}$$

This is the quantum analogue of the classical equation $p_i = m\frac{dx_i}{dt}$. Similarly we can compute the time evolution for the momentum operator. Inserting $[H,P_i] = -\frac{\hbar}{i}\frac{\partial}{\partial x_i}V(\vec{x})$ into (3.156) we obtain

$$\frac{d}{dt}\langle P_i\rangle = -\langle\partial_i V(\vec{x})\rangle, \tag{3.159}$$

which corresponds to Newton's equation of motion $\dot{\vec{p}} = -\,\vec{\nabla}\,V(\vec{x})$.

Heisenberg's uncertainty relation. Let $\langle A\rangle$ and $\langle B\rangle$ be the expectation values of two Hermitian operators A and B in some normalized state $|\psi\rangle \in \mathcal{H}$. The uncertainty ΔA is defined by

$$(\Delta A)^2 = \big\langle A^2\big\rangle - \big\langle A\big\rangle^2 = \big\langle\,(\delta A)^2\,\big\rangle \qquad \text{with} \qquad \langle A\rangle \equiv \langle\psi|A|\psi\rangle, \qquad \delta A = A - \langle A\rangle\mathbb{1}. \tag{3.160}$$

For the operators $\delta A = A - \langle A\rangle$ and $\delta B = B - \langle B\rangle$, which describe the deviation of the oberservables A and B from their mean values, we consider the states

$$|\chi\rangle = \delta A\,|\psi\rangle \qquad \text{and} \qquad |\varphi\rangle = \delta B\,|\psi\rangle \tag{3.161}$$

whose norms are equal to the uncertainties

$$\langle\chi|\chi\rangle = \langle\psi|(\delta A)^2|\psi\rangle = (\Delta A)^2, \qquad \langle\varphi|\varphi\rangle = \langle\psi|(\delta B)^2|\psi\rangle = (\Delta B)^2, \tag{3.162}$$

and which satisfy the Schwartz inequality

$$\langle\chi|\chi\rangle\langle\varphi|\varphi\rangle \geq |\langle\chi|\varphi\rangle|^2. \tag{3.163}$$

Putting everything together we obtain the inequality

$$(\Delta A)^2(\Delta B)^2 = \langle\psi|(\delta A)^2|\psi\rangle\langle\psi|(\delta B)^2|\psi\rangle \geq |\langle\psi|(\delta A)\,(\delta B)|\psi\rangle|^2 = |\langle\delta A\delta B\rangle|^2. \tag{3.164}$$

Now we write the operator in the last term as $\delta A\ \delta B = \frac{1}{2}[\delta A, \delta B] + \frac{1}{2}\{\delta A, \delta B\}$ and consider the commutator $[\delta A, \delta B]$. Since $\langle A\rangle$ and $\langle B\rangle$ are scalars that commute with everything we observe

$$[\delta A, \delta B] = [A - \langle A\rangle\mathbb{1}\,,\ B - \langle B\rangle\mathbb{1}] = [A, B]. \tag{3.165}$$

Since a commutator of Hermitian operators is anti-Hermitian its expectation value is imaginary. Anti-commutators of Hermitian operators, on the other hand, are Hermitian and thus have real expectation values. Hence

$$\langle\delta A\delta B\rangle = \frac{1}{2}\langle\{\delta A, \delta B\}\rangle + \frac{1}{2}\langle[\delta A, \delta B]\rangle = \frac{1}{2}\langle\{\delta A, \delta B\}\rangle + \frac{1}{2}\langle[A, B]\rangle \tag{3.166}$$

decomposes the expectation value into its real and its imaginary part, so that its squared modulus becomes

$$|\langle\delta A\delta B\rangle|^2 = \frac{1}{4}|\langle[A, B]\rangle|^2 + \frac{1}{4}|\langle\{\delta A, \delta B\}\rangle|^2 \geq \frac{1}{4}|\,\langle[A, B]\rangle\,|^2. \tag{3.167}$$

Combining this estimate with the inequality (3.164) we find

$$\langle(\delta A)^2\rangle\langle(\delta B)^2\rangle \geq \frac{1}{4}|\,\langle[A, B]\rangle\,|^2 \tag{3.168}$$

and by taking positive square roots on both sides we find the general form of **Heisenberg's uncertainty relation**

$$\boxed{\Delta A\Delta B \geq \frac{1}{2}|\langle[A, B]\rangle|} \tag{3.169}$$

which establishes a lower bound on the product of uncertainties of two operators in terms of the expectation value of their commutator. The two respective observables can hence be measured simultaneously with arbitrary precission only if the operators commute, or, more precisely, if their commutator has vanishing expectation value. We should stress that this uncertainty is not simply a problem of measurement but rather an intrinsic property of quantum mechanics.

Uncertainty of position and momentum. For the most famous example of an uncertainty relation we insert the commutator between position and momentum and obtain

$$[P_j, X_i] = \frac{\hbar}{i}\delta_{ij} \qquad \Rightarrow \qquad (\Delta X_i)(\Delta P_j) \geq \frac{\hbar}{2}\delta_{ij}. \tag{3.170}$$

so that position and momentum *in the same direction* cannot be measured simultaneously with arbitrary precision.[13]

[13] For states of minimal uncertainty $\Delta X\Delta P = \frac{\hbar}{2}$ the two inequalities (3.164) and (3.167) have to be equalities, which requires that $\delta X|\psi\rangle$ and $\delta P|\psi\rangle$ are proportional and that $\delta X\delta P + \delta P\delta X$ have vanishing expectation value. It is easy to check that this can only be the case for Gaussian wave packets.

Uncertainty of time and energy. If we consider the form of a plane wave $\psi = e^{\frac{i}{\hbar}(\vec{p}\vec{x}-Et)}$ we might expect that there exists an uncertainty between energy and time analogous to the one between momentum and position. There exists, however, no time operator in quantum mechanics and hence no uncertainty relation involving t in the literal sense.

Uncertainty relations of the extected type do exist, however, if we think of time in terms of time measurements, like for example the time of the detection of a particle. Such a measurement always involves the observation of a change in time of the value of some observable A and the uncertainty of time would be the time that it takes for this change to become larger than the intrinsic uncertainty of that observable,

$$\Delta t_A = \frac{\Delta A}{|\frac{d}{dt}\langle A\rangle|}. \tag{3.171}$$

Since time evolution is generated by the Hamilton operator an uncertainty relation for Δt_A can now be obtained as an consequence of the uncertainty relation between A and H,

$$\Delta A \Delta E \geq \frac{1}{2}|\langle [H, A]\rangle|. \tag{3.172}$$

If we combine this with the equation of motion (3.156) of the expectation value for a time-independent observable A,

$$\frac{d}{dt}\langle A\rangle = \frac{i}{\hbar}\langle [H, A]\rangle, \tag{3.173}$$

we obtain $\Delta A\, \Delta E \geq \frac{\hbar}{2}\left|\frac{d}{dt}\langle A\rangle\right|$ and hence the uncertainty relation

$$\Delta t_A \cdot \Delta E \geq \frac{\hbar}{2}, \tag{3.174}$$

which is exactly of the form that we had hoped for. It is hence not possible to *simultaneously* measure the energy of a particle and time of its detection with arbitrary precision.

3.6 Harmonic oscillator and ladder operators

Using the operator calculus we now determine the energy spectrum of the harmonic oscillator by purely algebraic calculations. We begin by introducing dimensionless position and momemtum operators

$$\mathcal{X} = \sqrt{\frac{m\omega_0}{\hbar}}X, \qquad \mathcal{P} = \frac{1}{\sqrt{m\hbar\omega_0}}P \tag{3.175}$$

so that

$$H = \frac{\hbar\omega_0}{2}\,(\mathcal{P}^2 + \mathcal{X}^2). \tag{3.176}$$

Classically we can factorize $x^2 + p^2 = (x + ip)(x - ip)$ as a product of complex conjugate numbers. Analogously, we introduce the non-Hermitian **ladder operators**

$$a = \frac{1}{\sqrt{2}}(\mathcal{X} + i\mathcal{P}), \qquad a^\dagger = \frac{1}{\sqrt{2}}(\mathcal{X} - i\mathcal{P}) \tag{3.177}$$

with

$$X = \sqrt{\frac{\hbar}{2m\omega_0}}\left(a^\dagger + a\right), \qquad P = i\sqrt{\frac{m\hbar\omega_0}{2}}\left(a^\dagger - a\right). \tag{3.178}$$

Since $[\mathcal{P}, \mathcal{X}] = \frac{1}{i}$ the commutator becomes $[a, a^\dagger] = \frac{1}{2}[\mathcal{X} + i\mathcal{P}, \mathcal{X} - i\mathcal{P}] = 0 + \frac{1}{2} + \frac{1}{2} + 0 = 1$, i.e.

$$[a, a^\dagger] = -[a^\dagger, a] = 1. \tag{3.179}$$

With the quantum mechanical relation

$$\mathcal{X}^2 + \mathcal{P}^2 = \frac{1}{2}\left((a^\dagger + a)^2 - (a^\dagger - a)^2\right) = a^\dagger a + aa^\dagger = 2a^\dagger a + 1 \tag{3.180}$$

we thus obtain

$$H = \hbar\omega_0\left(a^\dagger a + \frac{1}{2}\right) = \hbar\omega_0\left(\mathcal{N} + \frac{1}{2}\right), \tag{3.181}$$

where we defined the **occupation number operator**

$$\mathcal{N} = a^\dagger a. \tag{3.182}$$

This operator is positive, i.e. all its expectation values are positive, because

$$\langle\psi|\mathcal{N}|\psi\rangle = (\langle\psi|a^\dagger)(a|\psi\rangle) = ||\,(a|\psi\rangle)\,||^2 \geq 0 \tag{3.183}$$

is the squared norm of the vector $a|\psi\rangle$. Consequently all expectation values of H, and hence all energy eigenvalues E, are bounded from below by

$$E \geq \frac{1}{2}\hbar\omega_0. \tag{3.184}$$

$\frac{1}{2}\hbar\omega_0$ is called **zero-point energy** of the harmonic oszillator.

Creation and annihilation of energy. Since $H = \hbar\omega_0(\mathcal{N} + \frac{1}{2})$ the energy spectrum can now be computed by solving the eigenvalue problem for the occupation number operator

$$\mathcal{N}|n\rangle = n|n\rangle \qquad \Rightarrow \qquad H|n\rangle = \hbar\omega_0(n + \tfrac{1}{2})|n\rangle. \tag{3.185}$$

In order so solve this equation we compute the commutators

$$[\mathcal{N}, a] = [a^\dagger, a]a = -a, \qquad [\mathcal{N}, a^\dagger] = a^\dagger[a, a^\dagger] = a^\dagger, \tag{3.186}$$

where we evaluated $[a^\dagger a, a] = a^\dagger[a, a] + [a^\dagger, a]a = [a^\dagger, a]a = a$ using the "Leibniz rule" (3.45). These commutation relations show that $a^\dagger$ (a) increases (decreases) the occupation number by one and, accordingly, the energy by $\hbar\omega_0$ because

$$\mathcal{N}|n\rangle = n|n\rangle \quad \Rightarrow \quad \begin{cases} \mathcal{N}\big(a^\dagger|n\rangle\big) = ([\mathcal{N}, a^\dagger] + a^\dagger\mathcal{N})|n\rangle = (1+n)\big(a^\dagger|n\rangle\big) \\ \mathcal{N}\left(a|n\rangle\right) = \left([\mathcal{N}, a] + a\mathcal{N}\right)|n\rangle = (-1+n)\big(a|n\rangle\big) \end{cases} \tag{3.187}$$

(where we used the identity $XY = [X, Y] + YX$). Thus $a^\dagger$ and a are called **creation** and **annihilation operator**, respectively. Their collective name is "ladder operators" because they bring us up and down the ladder of energy levels. More precisely, since (3.187) implies that $a|n\rangle$ and $a^\dagger|n\rangle$ have occupation numbers $n \pm 1$, these states must be proportional to $|n \pm 1\rangle$

$$a^\dagger|n\rangle = c_{n+}|n+1\rangle, \qquad a|n\rangle = c_{n-}|n-1\rangle. \tag{3.188}$$

Assuming that all states are normalized $\langle n|n\rangle = 1$ we can now compute the normalization factors $c_{n\pm}$. Since norms are computed by multiplication with the Hermitian conjugate states,

$$a\,|n\rangle = c_{n-}\,|n-1\rangle \quad \stackrel{conj.}{\longrightarrow} \quad \langle n|\,a^\dagger = c^*_{n-}\,\langle n-1|, \tag{3.189}$$

$$a^\dagger\,|n\rangle = c_{n+}\,|n+1\rangle \quad \stackrel{conj.}{\longrightarrow} \quad \langle n|\,a = c^*_{n+}\,\langle n+1|, \tag{3.190}$$

the eigenvalue equation $a^\dagger a|n\rangle = \mathcal{N}|n\rangle = n|n\rangle$ implies

$$\langle n+1|n+1\rangle = \frac{1}{|c_{n+}|^2}\,\langle n|a\,a^\dagger|n\rangle = \frac{1}{|c_{n+}|^2}(\,\langle n|a^\dagger a|n\rangle + 1) = \frac{n+1}{|c_{n+}|^2} = 1, \tag{3.191}$$

$$\langle n-1|n-1\rangle = \frac{1}{|c_{n-}|^2}\,\langle n|a^\dagger a|n\rangle = \frac{n}{|c_{n-}|^2} = 1, \tag{3.192}$$

so that

$$c_{n+} = \sqrt{n+1}, \qquad c_{n-} = \sqrt{n}, \tag{3.193}$$

where the phase ambiguity of the eigenvectors $|n\rangle$ has been used to choose $c_{n\pm}$ positive real.

Quantization of occupation number and energy. Now we are ready to determine the eigenvalues n. We assume that at least one eigenstate $|n\rangle$ exists for some eigenvalue $n \in \mathbb{R}$, which has to be non-negative $n \geq 0$ because of the positivity (3.183) of $\mathcal{N}$. Now we act on this state k times with the annihilation operator a and obtain

$$a|n\rangle = \sqrt{n}\,|n-1\rangle, \tag{3.194}$$

$$a^2|n\rangle = \sqrt{n(n-1)}\,|n-2\rangle, \tag{3.195}$$

$$\ldots$$

$$a^k|n\rangle = \sqrt{n(n-1)\ldots(n-k+1)}\,|n-k\rangle. \tag{3.196}$$

We thus find new energy eigenstates with occupation numbers $n-1,\ n-2,\ \ldots$ However, this procedure has to terminate because otherwise we would be able to construct energy eigenstates for arbitrary $n-k$, which turns negative for $k > n$ in contradiction to the positivity of the operator $\mathcal{N}$. Hence there must exist a positive integer K for which $a^K|n\rangle = 0$. Choosing K minimal, so that $a^{K-1}|n\rangle \neq 0$, we conclude that $a|n-K+1\rangle = 0$ and hence $\langle n-K+1|a^\dagger a|n-K+1\rangle = n-K+1 = 0$. In other words, if $a|n'\rangle = 0$ the normalization factor $c_{n'-}$ must vanish, which is the only possibility to avoid the existence of an energy eigenstate with eigenvalue

$n' - 1$. We conclude that each energy eigenvalue n must be a non-negative integer. Moreover, eq. (3.196) shows that the minimal energy state has occupation number $n = 0$, and by acting with creation operators on the ground state $|0\rangle$,

$$(a^\dagger)^n|0\rangle \;=\; \sqrt{n!}\,|n\rangle \tag{3.197}$$

we conclude that all energy eigenstates with nonnegative integer occupation number indeed exist. We thus recover the result

$$E_n = \hbar\omega(n + \frac{1}{2}) \qquad \text{with} \qquad n = 0, 1, 2, \ldots \tag{3.198}$$

of our analytical treatment of the Harmonic oscillator. Moreover, the ground state $|0\rangle$ satisfies the first order differential equation $a|0\rangle = (\mathcal{X} + i\mathcal{P})|0\rangle = 0$, which is easily solved yielding the Gaussian wave function found in section 2. The wave functions with positive occupation numbers are

$$u_n(x) = \langle x|n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n u_0(x) \tag{3.199}$$

and can be evaluated by repeated application of the differential operator $a^\dagger$.

3.6.1 Coherent states

Coherent states are, by definition, eigenstates of the annihilation operator

$$a|\lambda\rangle_{\text{coh}} = \lambda|\lambda\rangle_{\text{coh}}. \tag{3.200}$$

They exist for all complex numbers $\lambda \in \mathbb{C}$ and are unique up to normalization. This can be verified by inserting the ansatz $|\lambda\rangle = \sum_{n=0}^{\infty} c_n|n\rangle$ in terms of energy eigenstates $|n\rangle$ into the eigenvalue equation (3.200). With the choice $c_0 = 1$ the resulting recursion relation is solved by

$$|\lambda\rangle_{\text{coh}} = \sum_{n=0}^{\infty} \frac{\lambda^n}{\sqrt{n!}}|n\rangle = e^{\lambda a^\dagger}|0\rangle. \tag{3.201}$$

It will usually be sufficient to distinguish coherent states $|\lambda\rangle_{\text{coh}}$ from energy eigenstates $|n\rangle$ by the use of Greek letters for the eigenvalues of a. The eigenstate property of $e^{\lambda a^\dagger}|0\rangle$ can be verified directly,

$$a(e^{\lambda a^\dagger}|0\rangle) = e^{\lambda a^\dagger}(e^{-\lambda a^\dagger} a e^{\lambda a^\dagger})|0\rangle = e^{\lambda a^\dagger}(a + \lambda)|0\rangle = \lambda(e^{\lambda a^\dagger}|0\rangle), \tag{3.202}$$

where we used $a|0\rangle = 0$ and the formula

$$e^{\lambda A} B e^{-\lambda A} = B + \sum_{n=1}^{\infty} \frac{\lambda^n}{n!}[A, B]_n, \qquad [A, B]_{n+1} = [A, [A, B]_n], \tag{3.203}$$

with $[A,B]_1 = [A,B]$. Since $[a^\dagger, a] = -1$ all higher commutators vanish.

Scalar products among coherent states can be computed directly from the series expansion or with the Baker-Campbell-Hausdorff formula

$$\langle\lambda|\mu\rangle = \langle 0|e^{\lambda^* a} e^{\mu a^\dagger}|0\rangle = \langle 0|e^{\mu a^\dagger} e^{\lambda^* \mu [a,a^\dagger]} e^{\lambda^* a}|0\rangle = e^{\lambda^* \mu} \tag{3.204}$$

where we used that eq. (3.59) implies $e^A e^B = e^B e^A e^{[A,B]}$ if all double-commutators of A and B vanish, as is the case for $A = \lambda^* a$ and $B = \mu a^\dagger$ because $[a, a^\dagger] = 1$. We also used $e^{\lambda^* a}|0\rangle = |0\rangle$ (because only the first term of the series is nonzero) and the Hermitian conjugate formula $\langle 0|e^{\mu a^\dagger} = \langle 0|$. Eigenstates of a for different eigenvalues are not orthogonal and the eigenvalues are neihter quantized nor required to be real (which is o.k. because a is not self-adjoint). For normalized coherent states we thus find the formula $e^{-\frac{1}{2}|\lambda|^2} e^{\lambda a^\dagger}|0\rangle$.

The time evolution of coherent states is easily calculated by using the expansion in terms of energy eigenstates,

$$|\lambda\rangle(t) = e^{-\frac{i}{\hbar}Ht}\Big(\sum_{n=0}^{\infty} \frac{\lambda^n}{\sqrt{n!}}|n\rangle\Big) = \sum_{n=0}^{\infty} \frac{\lambda^n}{\sqrt{n!}} e^{-i(n+\frac{1}{2})\omega_0 t}|n\rangle = e^{-i\frac{\omega_0}{2}t}|\lambda(t)\rangle, \quad \lambda(t) = e^{-i\omega_0 t}\lambda. \tag{3.205}$$

Up to an unobservable phase factor the time evolution thus corresponds to a rotation of the eigenvalue $\lambda(t) = e^{-i\omega_0 t}\lambda$ in the complex plane. In fact, the probability density of $\langle\lambda|\lambda\rangle(t)$ is given by a Gaussian distribution with minimal uncertainty $\Delta X \Delta P = \hbar/2$ and constant shape, whose mean value oscillates with the classical frequency ω_0, explaining the name *coherent state*. This can be shown by computing the wave function in configuration space $\psi_\lambda(x) = \langle x|\lambda\rangle$, which satisfies the first order differential equation

$$(a - \lambda)\psi_\lambda(x) = \frac{1}{\sqrt{2}}\left(\alpha x + \frac{1}{\alpha}\partial_x - \sqrt{2}\,\lambda\right)\psi_\lambda(x) = 0 \qquad \text{with} \qquad \alpha = \sqrt{\frac{m\omega_0}{\hbar}}. \tag{3.206}$$

With the ansatz $\psi_\lambda(x) = e^{-Ax^2+Bx-C}$ we find $\alpha x - \frac{1}{\alpha}(2Ax - B) - \sqrt{2}\lambda = 0$ so that $A = \alpha^2/2$ and $B = \sqrt{2}\alpha\lambda$. A coherent state hence is a Gaussian wave packet of the form

$$\psi_\lambda(x) = N_\lambda\, e^{-\frac{\alpha^2}{2}\left(x - \frac{\sqrt{2}\lambda}{\alpha}\right)^2} \tag{3.207}$$

with constant width $\Delta X = \frac{1}{\sqrt{2}\,\alpha}$ whose expectation value $\langle X\rangle = \frac{\sqrt{2}}{\alpha}\,\mathrm{Re}\,\lambda(t)$, according to eq. (3.205), oscillates about the origin with the classical frequence ω_0. It is straightforward to verify that coherent states have minimal uncertainty.[14] Hence they are the quantum analogue

[14] For Gaussian wave packets of the form $u(x) = e^{-Ax^2+Bx-C}$ normalization requires $\mathrm{Re}\,C = \frac{(\mathrm{Re}\,B)^2}{4\,\mathrm{Re}\,A} - \frac{1}{4}\log\frac{2\,\mathrm{Re}\,A}{\pi}$ and the expectation values and uncertainties of X and P are

$$\langle X\rangle = \frac{1}{2}\frac{\mathrm{Re}\,B}{\mathrm{Re}\,A}, \qquad \Delta X = \frac{1}{2\sqrt{\mathrm{Re}\,A}}, \qquad \langle P\rangle = \hbar\frac{\mathrm{Im}(A^*B)}{\mathrm{Re}\,A}, \qquad \Delta P = \frac{\hbar\,|A|}{\sqrt{\mathrm{Re}\,A}}. \tag{3.208}$$

They have minimal uncertainty $\Delta X \Delta P = \hbar/2$ exactly if A is real (normalizability or course requires $\mathrm{Re}\,A > 0$).

of a classical particle oscillating in a harmonic potential, which avoids the spreading of wave packets that we observed for free particles. Like harmonic potentials in classical physics, the harmonic oscillator is ubiquitous in quantum physics. In the quantum (field) theory of many particle systems the ladder operators will create and annihilate particles. In quantum optics the particles are the photons of momentum $\hbar k$ and polarization $\vec{\varepsilon}$, created and annihilated by $a^\dagger_{\vec{\varepsilon}}(\vec{k})$ and $a_{\vec{\varepsilon}}(\vec{k})$, respectively. Coherent states are thus very useful in laser physics.

3.7 Axioms and interpretation of quantum mechanics

3.7.1 Mixed states and the density matrix

We already learned that expectation values of operators A for a system whose state is described by a vector $|\psi\rangle \in \mathcal{H}$ can be computed by traces

$$\mathrm{tr}(P_\psi A) = \mathrm{tr}\big(|\psi\rangle\langle\psi|A\big) = \textstyle\sum_i \langle a_i|\ (|\psi\rangle\langle\psi|A)\ |a_i\rangle = \sum_i a_i\,|\langle a_i|\psi\rangle|^2 = \sum_i p_{a_i}\,a_i = \langle A\rangle_\psi \quad (3.209)$$

where $p_{a_i} = |\langle a_i|\psi\rangle|^2$ is the probability to measure the eigenvalue a_i and $P_\psi = |\psi\rangle\langle\psi|$ is the projector onto the state $|\psi\rangle$. In practice we may only have incomplete information about the state of a system. If we consider, for example, an unpolarized or partially polarized electron beam then we have a reasonably well-defined velocity, but for the spin polarization we only have a *classical* probability distribution. Such systems are said to be in a **mixed state**: Let

$$\{p_i\} \qquad \text{with} \qquad \sum_i p_i = 1 \quad (3.210)$$

describe **classical probabilities** p_i for a system to be in the quantum states $|\psi_i\rangle$. Then expectation values have to be computed as quantum mechanical expectations weighted by classical probabilities,

$$\langle A\rangle = \textstyle\sum_i p_i\langle\psi_i|A|\psi_i\rangle = \sum_i p_i\,\mathrm{tr}\,(P_{\psi_i}A) = \mathrm{tr}\,(\sum_i p_i P_{\psi_i}A) \quad (3.211)$$

which motivates the definition of the **density matrix** or **density operator** as

$$\rho = \sum_i\ p_i\,P_{\psi_i} = \sum_i\ |\psi_i\rangle p_i\langle\psi_i| \qquad \Rightarrow \qquad \langle A\rangle_\rho = \mathrm{tr}(\rho A). \quad (3.212)$$

Like projectors, density matrices are self-adjoint, but in general $\rho^2 \neq \rho$. Density matrices are instead characterized by positivity and unit trace: Since classical probabilities are nonnegative

$$p_i \in \mathbb{R}_{\geq 0} \quad \Rightarrow \quad \rho = \rho^\dagger \geq 0, \qquad \mathrm{tr}\,\rho = \sum p_i = 1. \quad (3.213)$$

Every quantum mechanical system can hence be described by a density matrix. The system is in a **pure state** if $\rho = P_\psi$ is the projector onto a Hilbert space vector $|\psi\rangle \in \mathcal{H}$ because then all

eigenvalues and hence all probabilities are equal to 0 or 1 so that all remaining uncertainties have a quantum mechanical origin. This leads to the following criterion

$$\begin{array}{lcl} \rho^2 = \rho & \Leftrightarrow & \text{pure state} \\ \rho^2 \neq \rho & \Leftrightarrow & \text{mixed state.} \end{array} \tag{3.214}$$

The spectral representation implies that every matrix obeying $\rho = \rho^\dagger \geq 0$ and $\operatorname{tr}\rho = 1$ is of the form $\rho = \sum_i p_i P_{e_i}$ for some orthonormal basis $\{e_i\}$ and can hence be interpreted as the density matrix for some (pure or) mixed state of the quantum mechanical system under consideration.

Using the Schrödinger equation $\partial_t|\psi\rangle = \frac{1}{i\hbar}H|\psi\rangle$ for $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|$ we find the time evolution equation

$$\partial_t\rho = -\frac{i}{\hbar}[H,\rho] \tag{3.215}$$

for the density operator $\rho = \rho_S$ in the Schrödinger picture. This looks similar to Heisenberg's equation of motion (3.144), but mind the opposite sign! Like states $|\psi\rangle_H$, density matrices ρ_H are time-independent in the Heisenberg representation so that expectation values of time-independent operators evolve according to

$$\partial_t\langle A\rangle_\rho = \operatorname{tr}\dot\rho_S A_S = \operatorname{tr}\rho_H \dot A_H, \tag{3.216}$$

where the second equality can be checked using $\operatorname{tr}([H,\rho]A) = \operatorname{tr}(H\rho A - \rho HA) = \operatorname{tr}(\rho AH - \rho HA) = -\operatorname{tr}(\rho[H,A])$, which follows from cyclicity of the trace.

Density matrices are particularly useful for quantum statistics because, for example, a **Boltzmann distribution** can be described by the operator

$$\rho_T = e^{-\frac{H}{kT}}/Z(T), \qquad Z(T) = \operatorname{tr}(e^{-\frac{H}{kT}}) \tag{3.217}$$

with **partition function** $Z(T)$, which is very handy for formal calculations.

3.7.2 Measurements and interpretation

In the canonical formulation of classical mechanics the state of a particle is specified at any time t by a pair of dynamical variables, the canonical momentum $\vec p(t)$ and the generalized coordinate $\vec q(t)$. The time evolution is governed by Hamilton's equations of motion (which are related to the Euler-Lagrange equations of the Lagrange formalism by a Legendre transformation). In contrast, quantum mechanics is defined by the following five axioms, which we already mentioned in chapter 2, but which we now discuss in more detail (in a slightly modified version).

- Postulate 1: **State of a system**
  A (pure) state of a quantum system is *completely* specified at any time t by a vector $|\psi(t)\rangle$ in a Hilbert space $\mathcal{H}$.

- Postulate 2: **Observables and operators**
  To every measurable quantity, called observable or dynamical variable, there corresponds a self-adjoint linear operator A, whose eigenvectors form a complete basis. Operators B_k and C_l that correspond to canonically conjugate variables, like the positions X_i and the canonical momenta P_j, obey the canonical commutation relations

$$[B_k, C_l] = \tfrac{\hbar}{i}\delta_{kl}\mathbb{1}. \tag{3.218}$$

  The operator algebra defined by this equation is called *Heisenberg algebra*.

- Postulate 3: **Measurements and eigenvalues of operators**
  The measurement of an observable is related to the action of the corresponding operator A on a state vector $|\psi(t)\rangle$ as follows. The only possible result of a measurement is given by one of the eigenvalues a_n of the operator A. If the result of the measurement of A is a_n then the state of the system *immediately after* the measurement is given by the eigenstate $|a_n\rangle$; this is often called the *collapse of the wave function*. If the eigenvalue a_n is degenerate, the new state of the system is proportional to the projection of the state $|\psi\rangle$ onto the eigenspace of the eigenvalue a_n,

$$|\psi\rangle_{after} = c_n\, P_{a_n}\, |\psi(t)\rangle \qquad \text{with} \qquad P_{a_n} = \textstyle\sum_i |a_{ni}\rangle\langle a_{ni}| \tag{3.219}$$

  and normalization factor $c_n = 1/\sqrt{|\operatorname{tr}(P_\psi P_{a_n})|}$, where $|a_{ni}\rangle$ is an orthonormal basis of the eigenspace with eigenvalue a_n. If the system has been in a pure state before the measurement it will continue to be so after the measurement. If, on the other hand, the system originally is in a mixed state, appropriate measurements can be performed to remove all classical uncertainties and to *prepare* a pure state. If the eigenvalue a_n is nondegenerate then a single measurement of a_n is sufficient for this purpose.

- Postulate 4: **Probabilistic outcome of measurements**
  When measuring an observable A of a system in a state vector $|\psi\rangle$, the probability of obtaining one of the nondegenerate eigenvalues a_n of the corresponding operator A is given by

$$p(a_n) = \frac{|\langle a_n|\psi\rangle|^2}{\langle\psi|\psi\rangle}. \tag{3.220}$$

  In the case of m-fold degenerate eigenvalues a_n the formula has to be generalized to

$$p(a_n) = \sum_{j=1}^{m} \frac{|\langle a_{nj}|\psi\rangle|^2}{\langle\psi|\psi\rangle} = \operatorname{tr} P_{a_n} P_\psi. \tag{3.221}$$

  If the system is already in an eigenstate of A then a measurement of A yields the corresponding eigenvalue with probability $p(a_n) = 1$. For continuous parts of the spectrum

probabilities have to be replaced by probability densities with obvious modifications of the corresponding formulas. For the position operator X this implies, in particular, Born's probabilistic interpretation of the wave function.

- Postulate 5: **Time evolution of a system**
  The time evolution of a quantum mechanical system is determined by the Schrödinger equation
  $$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi(t)\rangle, \tag{3.222}$$
  where H is the Hamiltonian operator corresponding to the total energy of the system.

3.7.3 Schrödinger's cat and the Einstein-Podolsky-Rosen argument

The probabilistic interpretation of Schrödinger's wave function by Max Born spawned a long and controversial discussion about the proper interpretation of quantum mechanics, which was most vigorous in the 1930s but is still going on. The probabilistic *Copenhagen interpretation* was named after the affiliation of its most prominent proponent Niels Bohr, who emphasized the role of an "intelligent" or "conscious" observer inducing the collapse of the wave function by his or her measurement activities. This somewhat extreme point of view was rediculed by Einstein, who asked whether the moon would still be there when he does not look at it, and by the famous story of Schrödingers cat, sitting in a closed box with a radioactive devices that triggers the killing of the cat on the random event of a nuclear decay. The wave function of the cat would hence be a coherent superposition

$$\psi_{cat} = c_a(t)\psi_{alive} + c_d(t)\psi_{dead} \tag{3.223}$$

possibly long after the cat was actually killed (in the original version of the story by poisoning). The collapse of the wave function would only occur when the box is opened by a human being. In more recent years the role of the observer has been replaced by the concept of decoherence, which amounts to a progressive loss of quantum mechanical interference patterns due to many small interactions of a particle with its environment like, for example, with a system in thermal equilibrium. A *decoherence theorem* was proven by Hepp, Lieb, et al. in 1982. In particular, decoherence is not certain itself so that there only exists a certain probability for this effect, which gets very close to one in macroscopic systems. In 1986 Asher Perez showed that the interaction of a quantum mechanical system with a chaotic system may also trigger the collapsing of the wave function. A very recommendably discussion of decoherence and of interpretations of quantum mechanics like Everett's *many worlds* can be found in the article *100 Years of the Quantum* by Tegmark and Wheeler.[15]

[15] Max Tegmark and John Archibald Wheeler: http://arxiv.org/abs/quant-ph/0101077.

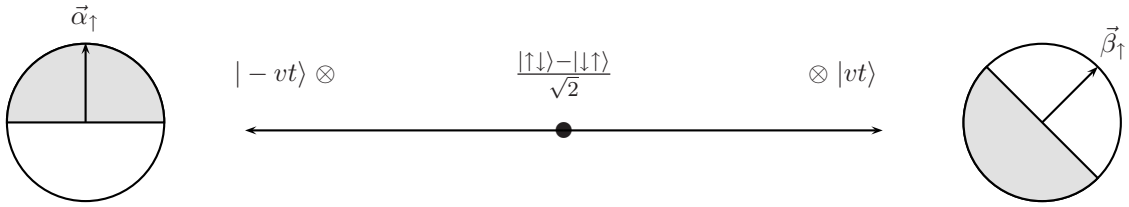


Figure 3.1: Bohm's version of the EPR experiment with the decay of a singlet state and spin measurements in directions $\vec{\alpha}$ and $\vec{\beta}$.

EPR-paradox and Bell's inequalities. In their famous 1935 article Einstein, Podolsky and Rosen tried to argue that quantum mechanics must be incomplete in the sense that there exist *hidden variables* that have to be supplemented to the quantum mechanical information of the wave function and that would, after all, remove any uncertainties except for classical probabilities due to inclomplete information about the state of a system. This paradox was the pinnacle of a long discussions over quantum theory between Albert Einstein and Niels Bohr, and it became a standard setup on the basis of which questions about the interpretation of quantum mechanics can be translated into experimentally testable predictions. Actually, what we will discuss is a simplified version of EPR due to David Bohm, who avoided a technically complicated discussion of position and momentum measurements by considering, instead, discrete spin degrees of freedom.

In Bohm's version of EPR we consider a system consisting of two spin-$\frac{1}{2}$ particles in a singlet state (i.e. the total angular momentum is zero), for which the spin degrees of freedom are described by the wave function

$$|\chi\rangle = \frac{1}{\sqrt{2}}\Big(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\Big), \tag{3.224}$$

as we will learn in detail in chapter 5. If the two particles break up in a decay process, as shown in figure 3.1, the spin degrees of freedom continue to be described by the non-product wave function (3.224) until a measurement is carried out. This phenomenon is called entanglement. The spin measurement in direction $\vec{\alpha}$ for the left-moving particle will always result in either spin-up or spin-down, both with a probability of $\frac{1}{2}$. The paradox situation, which EPR pointed out, is that conservation of angular momentum implies that the result of a spin-measurement for the second particle will immediately be influenced by the result of the first measurement. If the first particle shows spin-up then we know that, when measured with respect to the same direction $\vec{\alpha}$, the second particle will always have spin-down (and vice versa).

According to the Copenhagen interpretation of quantum mechanics the result of one measurement is governed by "objective" randomness. But this means that the result of the first measurement has to affect the second one, immediately and regardless of the distance. This, so the conclusion of EPR, would be in contradiction to causality in special relativity where

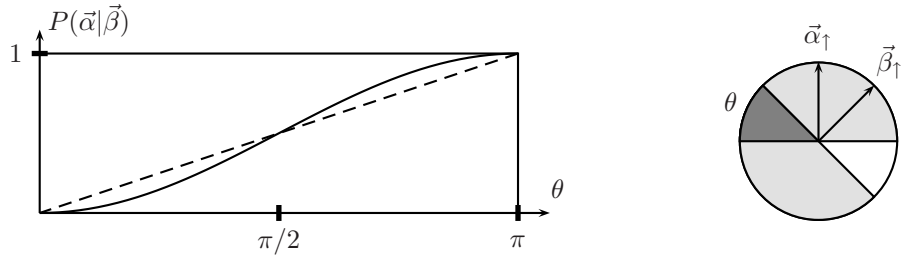


Figure 3.2: Classical (dashed line) and quantum mechanical conditional probabilities in EPR.

information can propagate only at the speed of light. The only way out seemed to be the existence of "hidden variables" which supplement the information contained in the wave function and which determine the results of future measurement. The particles would then know already right after the decay where the spin should point when measured and they would carry along this information until it is detected, thus removing the probabilistic aspects of quantum mechanics. This type of hypothetical hidden information is called **local hidden variables**. However, as Bohr pointed out the argumentation of EPR is not conclusive. Special relativity does not forbid all kinds of velocities larger than c, and while the outcome of the spin measurment of the second particle is instantly influenced by the result of the first, this cannot be used to transmit information with a velocity $v > c$ and hence does not contradict special relativity. Nevertheless, the phenomenon is astounding and hence called *spooky interaction at a distance*.

In 1932 John von Neumann gave a mathematical proof that hidden variables could not exist. However his assumptions were criticised as being too restrictive. Decisive progress only came with John Bell in 1964, who generalized the setup of the EPR paradox by measuring the probabilities for spin up or spin down in diffent directions $\vec{\alpha}$ and $\vec{\beta}$ for the two decay products, respectively. Bell showed that any *classical* probabilities due to incomplete knowledge of the values of *local* hidden variables obey certain constraints, known as **Bell's inequalities**. Essentially, the probability to find spin-up in direction $\vec{\alpha}$ for particle 1 under the condition of finding spin-direction $\vec{\beta}$ for particle 2 whould have to be linear in the angle enclosed by the vectors $\vec{\alpha}$ and $\vec{\beta}$,

$$P_{CL}(\vec{\alpha}|\vec{\beta}) = \theta/\pi \qquad \text{where} \qquad \cos\theta = \vec{\alpha}\vec{\beta}, \tag{3.225}$$

as illustrated in figure 3.2. This is clearly distinct from the quantum mechanical correlation

$$P_{QM}(\vec{\alpha}|\vec{\beta}) = \frac{1}{2}(1 - \vec{\alpha}\vec{\beta}), \tag{3.226}$$

which will be computed in chapter 5. Quantum correlations can hence be significantly stronger than allowed by local hidden variables. For the EPR setup the maximal violation of Bell's inequalities occurs for the angle $\theta = \frac{3}{4}\pi$, as can be seen in figure 3.2. Experimental results clearly confirm the predictions of quantum mechanics.

Chapter 4

Orbital angular momentum and the hydrogen atom

If I have understood correctly your point of view then you would gladly sacrifice the simplicity [of quantum mechanics] to the principle of causality. Perhaps we could comfort ourselves that the dear Lord could go beyond [quantum mechanics] and maintain causality.
-Werner Heisenberg responds to Einstein

In quantum mechanics degenerations of energy eigenvalues typically are due to symmetries. The symmetries, in turn, can be used to simplify the Schrödinger equation, for example, by a separation ansatz in appropriate coordinates. In the present chapter we will study rotationally symmetric potentials und use the angular momentum operator to compute the energy spectrum of hydrogen-like atoms.

4.1 The orbital angular momentum

According to Emmy Noether's first theorem continuous symmetries of dynamical systems imply conservation laws. In turn, the conserved quantities (called charges in general, or energy and momentum for time and space translations, respectively) can be shown to generate the respective symmetry transformations via the Poisson brackets. These properties are inherited by quantum mechanics, where Poisson brackets of phase space functions are replaced by commutators. According to the Schrödinger equation $i\hbar\partial_t\psi = H\psi$, for example, time evolution is generated by the Hamiltonian. Similarly, the momentum operator $\vec{P} = \frac{\hbar}{i}\vec{\nabla}$ generates (spatial) translations. A (Hermitian) charge operator Q is conserved if it commutes with the Hamiltonian $[H, Q] = 0$. This equation can also be interpreted as invariance of the Hamiltonian

$H = U_\lambda^{-1} H U_\lambda$ under the unitary 1-parameter transformation group of finite transformations $U_\lambda = \exp(i\lambda Q)$ that is generated by the infinitesimal transformation Q.

The constant of motion of classical mechanics that corresponds to rotations about the origin is the (orbital) angular momentum $\vec{L} = \vec{x} \times \vec{p}$. The corresponding operator $\vec{\mathcal{L}}$ in quantum mechanics is

$$\vec{\mathcal{L}} = \vec{X} \times \vec{P} = \frac{\hbar}{i}(\vec{x} \times \vec{\nabla}) = \frac{\hbar}{i} \begin{pmatrix} y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y} \\ z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z} \\ x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} \end{pmatrix}, \tag{4.1}$$

or

$$\mathcal{L}_i = \epsilon_{ijk} X_j P_k = \frac{\hbar}{i} \epsilon_{ijk} x_j \frac{\partial}{\partial x_k}. \tag{4.2}$$

There is no ordering ambiguity because X_j and P_k commute for $j \neq k$. In addition to the *orbital angular momentum* $\vec{\mathcal{L}}$, which is familiar from classical mechanics, quantum mechanical point particles can have an intrinsic angular momentum, the *spin* $\vec{\mathcal{S}}$, which will be the subject of the next chapter. The sum of all spins and orbital angular momenta of a system will be called the *total angular momentum* $\vec{\mathcal{J}}$.

4.1.1 Commutation relations

The canonical commutation relation $[X_i, P_j] = i\hbar\delta_{ij}$ implies

$$[\mathcal{L}_i, X_j] = \epsilon_{ikl}[X_k P_l, X_j] = i\hbar\epsilon_{ijk} X_k \tag{4.3}$$

and

$$[\mathcal{L}_i, P_j] = \epsilon_{ikl}[X_k P_l, P_j]. = i\hbar\epsilon_{ijk} P_k \tag{4.4}$$

The form of these results suggests that all vector operators V_j (i.e. operators with a vector index) should transform in the same way. Indeed, we will find that the (axial) vector $\mathcal{L}_j$ transforms as

$$[\mathcal{L}_i, \mathcal{L}_j] = i\hbar\epsilon_{ijk}\mathcal{L}_k. \tag{4.5}$$

To show this we use the identity

$$\varepsilon_{ikl}\varepsilon_{imn} = \delta_{km}\delta_{ln} - \delta_{kn}\delta_{lm}, \tag{4.6}$$

i.e. the sum over a common index i of a product of ε tensors is ± 1 if the free index pairs kl and mn take the same values, with the sign depending on the cyclic ordering, and vanishes otherwise. We thus find

$$[\mathcal{L}_i, \mathcal{L}_j] = \varepsilon_{jlm}[\mathcal{L}_i, X_l P_m] = i\hbar\varepsilon_{jlm}(\varepsilon_{ilk} X_k P_m + \varepsilon_{imk} X_l P_k) \tag{4.7}$$

$$= i\hbar\left((\delta_{ji}\delta_{mk} - \delta_{jk}\delta_{im}) X_k P_m + (\delta_{jk}\delta_{li} - \delta_{ji}\delta_{lk}) X_l P_k\right). \tag{4.8}$$

Since the terms with δ_{ji} cancel this agrees with

$$i\hbar\varepsilon_{ijk}\mathcal{L}_k = i\hbar\varepsilon_{ijk}\varepsilon_{klm}X_lP_m = i\hbar(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})X_lP_m, \tag{4.9}$$

which completes the proof of (4.5).

Since angular momenta in different directions do not commute they cannot be diagonalized simultaneously. If we choose $\mathcal{L}_z$ as our first observable we expect that the combination $\mathcal{L}_x^2+\mathcal{L}_y^2$, which is classically invariant under rotations about the z-axis, commutes with $\mathcal{L}_z$. This is indeed true, but it is more useful to use the completely rotation invariant $\mathcal{L}^2 = \mathcal{L}_x^2+\mathcal{L}_y^2+\mathcal{L}_z^2$ as the second generator of a maximal set of commuting operators. $\mathcal{L}^2$ is obviously Hermitian and

$$[\mathcal{L}_i, \mathcal{L}^2] = [\mathcal{L}_i, \mathcal{L}_k]\mathcal{L}_k + \mathcal{L}_k[\mathcal{L}_i, \mathcal{L}_k] = i\hbar\epsilon_{ikr}(\mathcal{L}_r\mathcal{L}_k + \mathcal{L}_k\mathcal{L}_r) = 0. \tag{4.10}$$

A similar calculation shows that $[\mathcal{L}_i, P^2] = [\mathcal{L}_i, X^2] = 0$, so that the kinetic energy commutes with $\mathcal{L}_i$ (and hence also with $\mathcal{L}^2$).

Angular momentum conservation $[\mathcal{L}_i, H]$ for rotationally symmetric Hamiltonians $H = \frac{P^2}{2m} + V(r)$ with $r = \sqrt{x^2+y^2+z^2}$ now already follows from the commutation of $\mathcal{L}_i$ with any function of X^2, but let us check this explicitly in configuration space,

$$[\mathcal{L}_j, H] = [\mathcal{L}_j, V(r)] = \frac{\hbar}{i}\varepsilon_{jkl}x_k\frac{\partial}{\partial x_l}V(r) = \frac{\hbar}{i}\varepsilon_{jkl}x_k\frac{x_l}{r}\frac{\partial}{\partial r}V(r) = 0, \tag{4.11}$$

where we used the chain rule for $V(r(x))$ with $\frac{\partial r}{\partial x_l} = \frac{x_l}{r}$ and the operator rule

$$[\frac{\partial}{\partial x_l}, A(x)]\ \psi(x) = \frac{\partial}{\partial x_l}\Big(A(x)\psi(x)\Big) - A\frac{\partial}{\partial x_l}\psi(x) = \left(\frac{\partial}{\partial x_l}A(x)\right)\psi(x), \tag{4.12}$$

or $[\partial_{x_i}, A(x)] = \partial_{x_i}A(x) + A(x)\partial_{x_i} - A(x)\partial_{x_i} = \partial_{x_i}A(x)$, i.e. commutation of ∂_{x_i} with an operator yields the partial derivative of that operator.

4.1.2 Angular momentum and spherical harmonics

We will first derive the relation

$$\Delta = \frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{1}{r^2}\frac{\mathcal{L}^2}{\hbar^2} \tag{4.13}$$

between $\mathcal{L}^2$ and the Laplacian, which will help us reduce the Schrödinger equation to an ordinary radial differential equation after separation of the angular coordinates. Hence we first evaluate $\mathcal{L}^2$ in configuration space,

$$\begin{aligned}
\mathcal{L}^2 = \mathcal{L}_i\mathcal{L}_i \ &= -\hbar^2\epsilon_{ijk}x_j\partial_k\epsilon_{ilm}x_l\partial_m = \\
&= -\hbar^2(\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl})x_j\partial_k x_l\partial_m = \\
&= -\hbar^2(x_j\partial_k x_j\partial_k - x_j\partial_k x_k\partial_j) = \\
&= -\hbar^2(x_j\partial_j + x_jx_j\partial_k\partial_k - 3x_j\partial_j - x_jx_k\partial_k\partial_j) = \\
&= -\hbar^2(x_jx_j\partial_k\partial_k - 2x_j\partial_j - x_jx_k\partial_k\partial_j).
\end{aligned} \tag{4.14}$$

Next we transform to spherical coordinates $\vec{x} = (r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta)$, hence

$$r = \sqrt{x^2+y^2+z^2}, \quad \theta = \arctan\frac{\sqrt{x^2+y^2}}{z}, \quad \varphi = \arctan\frac{y}{x}. \tag{4.15}$$

In particular

$$\frac{x_i}{r} = e_i^{(r)}, \quad x_j x_j = r^2, \quad x_j\partial_j = r\frac{\partial}{\partial r}, \tag{4.16}$$

so that we obtain

$$\mathcal{L}^2 = -\hbar^2(r^2\Delta - 2r\frac{\partial}{\partial r} - r^2\frac{\partial^2}{\partial r^2}), \tag{4.17}$$

or

$$\Delta = \partial_r^2 + \frac{2}{r}\partial_r - \frac{1}{r^2}\frac{\mathcal{L}^2}{\hbar^2} = \frac{1}{r}\partial_r^2\, r - \frac{1}{r^2}\frac{\mathcal{L}^2}{\hbar^2}, \tag{4.18}$$

which establishes (4.13).

Recalling the formula for the Laplace operator in spherical coordinates

$$\Delta = \frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{1}{r^2}\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\frac{\partial}{\partial\theta}) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right) \tag{4.19}$$

from *Mathematical Methods in Theoretical Physics* [Dirschmid,Kummer,Schweda] we conclude

$$\mathcal{L}^2 = -\hbar^2\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\frac{\partial}{\partial\theta}) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right). \tag{4.20}$$

Using the chain rule

$$\partial_i = (\partial_i r)\frac{\partial}{\partial r} + (\partial_i\theta)\frac{\partial}{\partial\theta} + (\partial_i\varphi)\frac{\partial}{\partial\varphi} \tag{4.21}$$

for $\mathcal{L}_z = \frac{\hbar}{i}(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial_x})$ one can check that

$$\mathcal{L}_z = \frac{\hbar}{i}\frac{\partial}{\partial\varphi}. \tag{4.22}$$

The common eigenfunctions for the angle-dependent part $-\mathcal{L}^2/\hbar^2$ of the Laplace operator and for $i\mathcal{L}_z/\hbar = \partial/\partial\varphi$ are again known from the *Mathematical Methods in Theoretical Physics*. They are the *spherical harmonics* [German: *Kugelflächenfunktionen*]

$$Y_{lm}(\theta,\varphi) = (-1)^m\sqrt{\frac{2l+1}{4\pi}\,\frac{(l-m)!}{(l+m)!}}\,P_l^{(m)}(\cos\theta)\,e^{im\varphi} = (-1)^m Y^*_{l,-m}(\theta,\varphi) \tag{4.23}$$

with the associated Legendre functions

$$P_l^{(m)}(\xi) = \frac{1}{2^l l!}\,(1-\xi^2)^{\frac{m}{2}}\,\frac{d^{l+m}}{d\xi^{l+m}}\,(\xi^2-1)^l = (-1)^m\frac{(l+m)!}{(l-m)!}P_l^{(-m)}(\xi) \tag{4.24}$$

For $m = 0$ they reduce to the Legendre polynomials $P_l(\xi) = P_l^{(0)}(\xi)$. These results are obtained by a the separation ansatz $Y_{lm} = \Theta(\theta)\Phi(\varphi)$. The eigenvalues for the angular momenta become

$$\mathcal{L}^2\,Y_{lm} = \hbar^2 l(l+1)\,Y_{lm}, \qquad \mathcal{L}_3\,Y_{lm} = \hbar m\,Y_{lm}, \tag{4.25}$$

with l a nonnegative integer and $m \in \mathbb{Z}$ obeying $-l \leq m \leq l$. The quantization conditions and the ranges of the eigenvalues follow from termination conditions for the power series ansatz and from single-valuedness at $\xi = \cos\theta = \pm 1$.

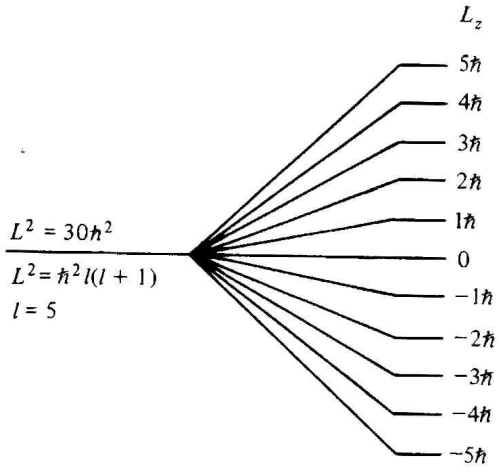


Figure 4.1: The eigenvalue $\hbar^2 l(l+1)$ of $\mathcal{L}^2$ is $(2l+1)$–fold degenerate

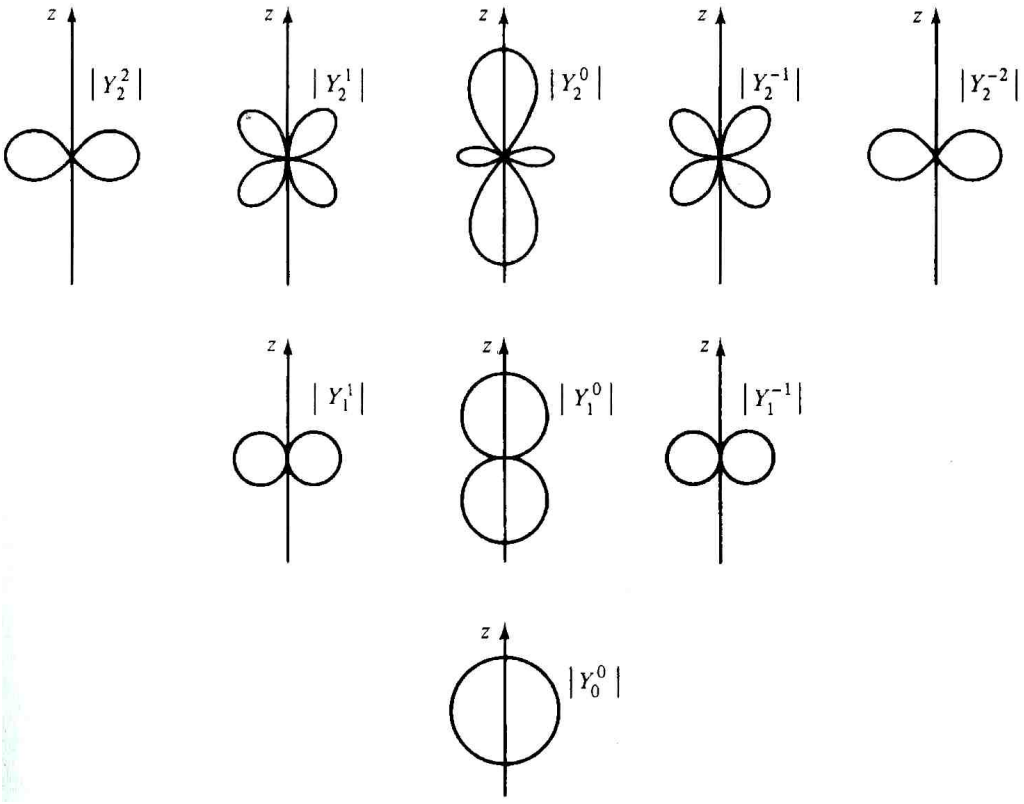


Figure 4.2: Polar plots of $|Y_l^m|$ versus Θ in any plane through the z-axis for $l = 0, 1, 2$. Note the equality $|Y_l^m| = |Y_l^{-m}|$, which follows from $Y_{l,m}^* = (-1)^m Y_{l,-m}$.

The completeness of the spherical harmonics enables us to solve the stationary Schrödinger equation for rotation invariant potentials by a separation ansatz $u(\vec{x}) = R_\lambda(r)Y_{lm}(\theta,\varphi)$ with

$$H_l R_\lambda(r) = \lambda\, R_\lambda(r), \qquad \text{with} \quad H_l = -\frac{\hbar^2}{2m}\frac{1}{r}\,\partial_r^2\, r + \frac{\hbar^2 l(l+1)}{2mr^2} + V(r). \tag{4.26}$$

The energy eigenvalues λ are (2l+1)-fold degenerate due to the *magnetic* quantum number m. Note that we need the *two* observables $\mathcal{L}^2$ and $\mathcal{L}_z$ to characterize the wave function dependence on the *two* angle coordinates θ and φ.

4.2 The hydrogen atom

4.2.1 The two particle problem

Consider a system of two particles of the masses m_1 and m_2 and positions $\vec{x}_1$ and $\vec{x}_2$, respectively. If there are no forces from outside translation invariance implies that the potential energy $V(\vec{x}_1,\vec{x}_2)$ only depends on the difference vector $\vec{x} = \vec{x}_1 - \vec{x}_2$. In classical mechanics this system is hence described by the Lagrangian

$$L(\vec{x}_1,\dot{\vec{x}}_1;\vec{x}_2,\dot{\vec{x}}_2) = T - V = \frac{1}{2}(m_1\dot{\vec{x}}_1^2 + m_2\dot{\vec{x}}_2^2) - V(\vec{x}_1 - \vec{x}_2). \tag{4.27}$$

The description can be simplified by using the relative coordinates

$$\vec{x} = \vec{x}_1 - \vec{x}_2 \tag{4.28}$$

and the center of mass coordinates

$$\vec{x}_g = \frac{m_1\vec{x}_1 + m_2\vec{x}_2}{m_1 + m_2} \tag{4.29}$$

as new variables, so that $\vec{x}_1 = \vec{x}_g + \frac{m_2}{m_1+m_2}\vec{x}$ and $\vec{x}_2 = \vec{x}_g - \frac{m_1}{m_1+m_2}\vec{x}$. In terms of the total mass M and the reduced mass μ,

$$M = m_1 + m_2, \qquad \mu = \frac{m_1 m_2}{m_1 + m_2}, \tag{4.30}$$

the total momentum $\vec{p}_g$ and the relative momentum $\vec{p}$ are

$$\vec{p}_g = \quad M\dot{\vec{x}}_g = m_1\dot{\vec{x}}_1 + m_2\dot{\vec{x}}_2 = \vec{p}_1 + \vec{p}_2 \tag{4.31}$$

$$\vec{p} = \quad \mu\dot{\vec{x}} = \frac{m_2\vec{p}_1 - m_1\vec{p}_2}{m_1 + m_2} \tag{4.32}$$

and the Hamiltonian becomes

$$H(\vec{x}_g,\vec{p}_g;\vec{x},\vec{p}) = H_g(\vec{x}_g,\vec{p}_g) + H_r(\vec{x},\vec{p}) = \frac{\vec{p}_g^{\,2}}{2M} + \frac{\vec{p}^{\,2}}{2\mu} + V(\vec{x}). \tag{4.33}$$

$H_g = \frac{\vec{p}_g^2}{2M}$ describes the uniform free motion $\dot{\vec{p}}_g = M\ddot{\vec{x}}_g = 0$ of the center of mass, while the reduced Hamiltonian

$$H_r = \frac{\vec{p}^2}{2\mu} + V(\vec{x}) \tag{4.34}$$

described the dynamics $\dot{\vec{p}} = \mu\ddot{\vec{x}} = -\vec{\nabla}V(\vec{x})$.

In quantum mechanics the canonical commutation relations $[X_i^{(1)}, P_j^{(1)}] = [X_i^{(2)}, P_j^{(2)}] = i\hbar\delta_{ij}$ and $[X_i^{(1)}, P_j^{(2)}] = [X_i^{(2)}, P_j^{(1)}] = 0$ are, as expected, equivalent to

$$[X_i, P_j] = [X_i^{(g)}, P_j^{(g)}] = i\hbar\delta_{ij}, \qquad [X_i^{(g)}, P_j] = [X_i, P_j^{(g)}] = 0 \tag{4.35}$$

(i.e. the change of variables amounts to a canonical transformation). Hence H_g and H_r commute and can be diagonalized simultaneously with a separation ansatz $u(\vec{x}_1, \vec{x}_2) = u_g(\vec{x}_g)u_r(\vec{x})$ and the total energy becomes $E = E_g + E_r$. After the separation of the center of mass motion the dynamics is hence described by a one-particle problem with effective mass $\mu = \frac{m_1 m_2}{m_1+m_2}$ and potential $V(\vec{x})$.

4.2.2 The hydrogen atom

In this section we consider a simplified hydrogen-like atom (or ion) with a nucleus of atomic number Z and a single electron, where we neglect the spin and relativistic correction terms in the Hamiltonian, as well as the structure of the nucleus whose role is restricted to a massive point-like source for the Coulomb potential. It consists of protons with the mass m_p and elementary charge q,

$$m_p \ = 1,7 \cdot 10^{-27}\,\text{kg}, \qquad q \ = 1,6 \cdot 10^{-19}\,\text{Coulomb}, \tag{4.36}$$

and a number of neutrons, and the electron has charge $-q$ and mass

$$m_e = 0,91 \cdot 10^{-30} kg. \tag{4.37}$$

The electrostatic interaction potential between the electron and the point-like nucleus thus is

$$V(r) = -\frac{q^2}{4\pi\epsilon_0}\frac{Z}{r} = -\frac{Ze^2}{r}, \tag{4.38}$$

where $r = \sqrt{(\vec{x}_e - \vec{x}_{nucleus})^2}$ denotes the distance between the electron and the nucleus and

$$e^2 = \frac{q^2}{4\pi\epsilon_0}. \tag{4.39}$$

For the hydrogen atom $Z = 1$, while $Z = 2, 3, \ldots$ for the ions He^+, Li^{++} ....

The quantum mechanics of this system is described by the Hamiltonian

$$H(\vec{x},\vec{p}) = \frac{\vec{p}^2}{2\mu} + V(r) = \frac{\vec{p}^2}{2\mu} - \frac{Ze^2}{r}, \tag{4.40}$$

where the reduced mass

$$\mu = \frac{m_e m_{nucleus}}{m_e + m_{nucleus}} \approx m_e \left(1 - \frac{m_e}{m_{nucleus}}\right) \tag{4.41}$$

is very close to m_e since $m_{nucleus} \gg m_e$.

Now we recall the Laplace operator in spherical coordinates (4.19), which has a radial and a tangential part,

$$\Delta = \underbrace{\frac{1}{r}\frac{\partial^2}{\partial r^2} r}_{radial\ component} - \underbrace{\frac{1}{r^2}\frac{\mathcal{L}^2}{\hbar^2}}_{tangential\ component} . \tag{4.42}$$

The reduced Hamiltonian of a hydrogen-like atom thus becomes

$$H = -\frac{\hbar^2}{2\mu}\Delta - \frac{Ze^2}{r} = -\frac{\hbar^2}{2\mu}\left(\frac{1}{r}\frac{\partial^2}{\partial r^2} r - \frac{1}{r^2}\frac{\mathcal{L}^2}{\hbar^2}\right) - \frac{Ze^2}{r} \tag{4.43}$$

or

$$H = \frac{p_r^2}{2\mu} + \frac{\mathcal{L}^2}{2\mu r^2} + V(r), \qquad p_r = \frac{\hbar}{i}\frac{1}{r}\frac{\partial}{\partial r} r\,. \tag{4.44}$$

For bound states we expect negative energy eigenvalues $E < 0$ with

$$\left(\frac{p_r^2}{2\mu} + \frac{\mathcal{L}^2}{2\mu r^2} - \frac{Ze^2}{r}\right) u(\vec{x}) = E\, u(\vec{x}). \tag{4.45}$$

With the separation ansatz $u(\vec{x}) = R(r)Y_{lm}(\theta,\varphi)$ we obtain the radial eigenvalue equation

$$H_l\, R(r) = \left[-\frac{\hbar^2}{2\mu}\left(\frac{1}{r}\frac{\partial^2}{\partial r^2} r\right) + \frac{\hbar^2 l(l+1)}{2\mu r^2} - \frac{Ze^2}{r}\right] R(r) = E\, R(r) \tag{4.46}$$

with a Hamiltonian H_l depending on an integer parameter l. For large angular momentum l the radial Hamiltonian H_l thus has an effective repulsive contribution proportional to $1/r^2$, which is called *centrifugal barrier* (it stabilizes excited energy levels at high values of l). For fixed l (and m) we introduce a label, the *principal quantum number* n, for the different eigenvalues $E_{n,l}$ of H_l and we set

$$R(r) = \frac{1}{r} u_{n,l}(r). \tag{4.47}$$

Multiplication with $\frac{2r\mu}{\hbar^2}$ yields the differential equation

$$\left(-\frac{\partial^2}{\partial r^2} + \frac{l(l+1)}{r^2} - \frac{2\mu}{\hbar^2}\frac{Ze^2}{r} - \frac{2\mu E_{n,l}}{\hbar^2}\right) u_{n,l} = 0 \tag{4.48}$$

We first consider the asymptotics of its solutions $u_{n,l}$ for $r \to 0$ and for $r \to \infty$.

- For $r \to \infty$ this equation reduces to

$$(-\partial_r^2 + \kappa^2)\, u_{n,l} = 0 \qquad \text{with} \qquad \kappa = \frac{\sqrt{-2\mu E}}{\hbar} \tag{4.49}$$

whose solution is

$$u_{n,l} \sim Ae^{-\rho} + Be^{\rho} \qquad \text{with} \qquad \rho = \kappa r \tag{4.50}$$

$u_{n,l}$ has to vanish at infinity, hence only $e^{-\rho}$ is acceptable for $r \to \infty$.

- For $r \to 0$ the radial equation becomes

$$\left(-\partial_r^2 + \frac{l(l+1)}{r^2}\right) u_{n,l} = 0\,. \tag{4.51}$$

The ansatz $u_{n,l} \sim r^q$ yields $l(l+1) - q(q-1) = 0$, hence

$$u_{n,l} \sim Ar^{-l} + Br^{l+1}. \tag{4.52}$$

Normalizability requires $u_{n,l}$ to vanish at the origin so that $u_{n,l} \sim r^{l+1}$ for $r \to 0$.

Introducing the Bohr radius a_0

$$a_0 = \frac{\hbar^2}{\mu e^2} = 0.529 \cdot 10^{-10}\,\mathrm{m}, \tag{4.53}$$

equation (4.48) takes the form

$$\left(\partial_r^2 - \frac{l(l+1)}{r^2} + \frac{2Z}{a_0}\frac{1}{r} - \kappa^2\right) u_{n,l}(r) = 0, \tag{4.54}$$

or, in terms of dimensionless variables $\rho = \kappa r$ and n,

$$\left(\partial_\rho^2 - \frac{l(l+1)}{\rho^2} + \left(\frac{2n}{\rho} - 1\right)\right) u_{n,l} = 0, \qquad n = \frac{Z}{\kappa a_0}, \tag{4.55}$$

where the *principal quantum number* n parametrizes the energy eigenvalue $E = -\mathcal{R}\, Z^2/n^2$ with $\mathcal{R} = \hbar^2/(2\mu a_0^2) = \mu e^4/(2\hbar^2) = 2.18 \cdot 10^{-18}\mathrm{J} = 13.6\,\mathrm{eV} = 1$ Rydberg.

In order to account for the asymptotics of the solutions we write

$$u_{n,l}(\rho) = e^{-\rho}\rho^{l+1}F(\rho), \tag{4.56}$$

where $F(\rho)$ should be nonzero at the origin and should not grow faster than polynomial at ∞. Since $\partial_\rho^2 u = e^{-\rho}\rho^{l+1}(F'' + 2(\frac{l+1}{\rho} - 1)F' + (1 - 2\frac{l+1}{\rho} + \frac{l(l+1)}{\rho^2})F)$ we obtain

$$\left(\rho\frac{\partial^2}{\partial\rho^2} + 2(l+1-\rho)\frac{\partial}{\partial\rho} + 2(n-l-1)\right) F_{n,l}(\rho) = 0. \tag{4.57}$$

Expanding $F(\rho)$ into a power series $\sum_{j=0}^{\infty} a_j \rho^j$ the l.h.s. of (4.57) becomes a sum of three terms:

$$\rho F'' = \sum_{j=1}^{\infty} \rho^j \, (j(j+1)a_{j+1}), \tag{4.58}$$

$$2(l+1-\rho)F' = \sum_{j=0}^{\infty} \rho^j \, (2(l+1)(j+1)a_{j+1} - 2ja_j), \tag{4.59}$$

$$2(n-l-1)F = \sum_{j=0}^{\infty} \rho^j \, 2(n-l-1)a_j. \tag{4.60}$$

The vanishing of the coefficient of ρ^j implies the recursion relation

$$(j+1)(j+2l+2)a_{j+1} = 2(l+j+1-n)a_j. \tag{4.61}$$

For large j the ratio a_{j+1}/a_j is approximately $2/j$, which is the same as in the Taylor series of $e^{2\rho}$. The asymptotic behavior of the resulting solution would effectively invert the exponential damping in our ansatz (4.56). Normalizability therefore requires that the series terminates, which implies $l + j + 1 - n = 0$ for some nonnegative integer j, i.e. the principal quantum number n has to be a positive integer

$$n = 1, 2, 3, \ldots, \qquad 0 \leq l < n \tag{4.62}$$

and the energy eigenvalues are

$$E_n = -\frac{Z^2\mathcal{R}}{n^2} = -\frac{1}{2\mu}\left(\frac{Z\hbar}{a_0 n}\right)^2. \tag{4.63}$$

Somewhat surprisingly, the energy levels do not depend on the orbital quantum number l. For fixed principal quantum number n we therefore have an degeneracy of

$$\sum_{l=0}^{n-1}(2l+1) = n^2, \tag{4.64}$$

which is larger than what is implied by angular momentum conservation. The energy degeneracy for different values of l is a special property of the pure Coulomb interaction. It is lifted in nature by additional interaction terms that lead to the fine structure and hyperfine structure of the spectral lines. Note, however, that the degeneracy due to angular momentum conservation cannot be lifted by any corrections except in the presence of external forces, like an external magnetic field, which would break the rotation symmetry of the complete system (i.e. the atom plus its interaction with the environment). For a proper discussion of these effects we need to consider the spin of the electron. This will be the subject of the next section.

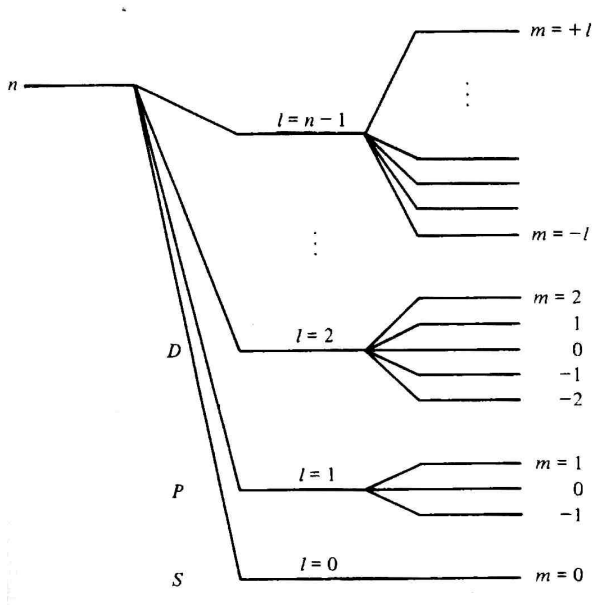


Figure 4.3: Term diagram for the hydrogen atom illustrating all n^2 degenerate states corresponding to the principal quantum number n.

The non-uniqueness of common eigenfunctions of H, L^2 and L_z for the Coulomb potential implies the existence of an independent conserved quantity that lifts this degeneracy. An appropriate observable can be constructed in terms of the Runge–Lenz vector,

$$\vec{M} = \frac{1}{2m}(\vec{P} \times \vec{\mathcal{L}} - \vec{\mathcal{L}} \times \vec{P}) - \alpha \frac{\vec{x}}{r} \qquad \text{for} \qquad V(r) = -\frac{\alpha}{r}, \tag{4.65}$$

which is well-known in the classical mechanics of planetary motion. It is straightforward to check its conservation $[H, \vec{M}] = 0$. Evaluation of the classical version $\vec{p} \times \vec{L}/m - \alpha \vec{x}/r$ at the perihelion shows that this vector points along the direction of the principal axis of the Kepler ellipse, which is a constant of non-relativistic motion in a pure $1/r$ potential. The quantum operator (4.65) is obtained by Weyl symmetrization, which is necessary for self-adjointness.[1]

[1] With $[\vec{P}, \frac{1}{r}] = i\hbar \frac{\vec{x}}{r^3}$, $[P_i, \frac{\vec{x}_j}{r}] = \frac{\hbar}{ir}(\delta_{ij} - \frac{x_i x_j}{r^2})$, $\vec{A} \times (\vec{B} \times \vec{C}) = A_j \vec{B} C_j - A_j B_j \vec{C}$ and $(\vec{x} \times \vec{\mathcal{L}})^\dagger = -\vec{\mathcal{L}} \times \vec{x}$ we can verify

$$[H, \vec{M}] = [-\tfrac{\alpha}{r}, \tfrac{\vec{P}\times\vec{\mathcal{L}} - \vec{\mathcal{L}}\times\vec{P}}{2m}] + [\tfrac{P^2}{2m}, -\alpha\tfrac{\vec{x}}{r}] = \tfrac{i\hbar\alpha}{2m}\left(\tfrac{\vec{x}}{r^3} \times \vec{\mathcal{L}} - \vec{\mathcal{L}} \times \tfrac{\vec{x}}{r^3} + \vec{P}\tfrac{1}{r} - (\vec{P}\vec{x})\tfrac{\vec{x}}{r^3} + \tfrac{1}{r}\vec{P} - \tfrac{\vec{x}}{r^3}(\vec{x}\vec{P})\right) = 0 \tag{4.66}$$

The Lenz vector, of course, does not commute with $\mathcal{L}_z$ but rather transforms as a vector, $[\mathcal{L}_i, M_j] = i\hbar\varepsilon_{ijk}M_k$. Since $\vec{M} \cdot \vec{\mathcal{L}} = \vec{\mathcal{L}} \cdot \vec{M} = 0$ and (after a tedious calculation) $M^2 = \frac{2H}{m}(\mathcal{L}^2 + \hbar^2) + \lambda^2$ is a function of H and $\mathcal{L}^2$, only M_z qualifies for the additional commuting operator that lifts the degeneracy. The algebra is completed (after further tedious calculations) by

$$[M_i, M_j] = i\hbar \frac{2H}{m} \varepsilon_{ijk} \mathcal{L}_k. \tag{4.67}$$

For fixed energy $H\psi = E\psi$ the six conserved charges $\mathcal{L}_i$ and $\mathcal{M}_j = \sqrt{\frac{-m}{2E}} M_j$ form an angular momentum algebra $SO(4)$ in 4 dimensions, or, equivalently, two independent angular momentum algebras $SO(3)$ generated by $\frac{1}{2}(\mathcal{L}_i \pm \mathcal{M}_i)$. The properties of abstract angular momentum algebras, which will be derived in the next section, can then be used for a complete algebraic computation of the energy levels of the hydrogen atom. For more details see `http://hbar.physik.uni-oldenburg.de/vlqm/VLqm/node72.html` or [Hannabuss].

4.3 Summary

- For a Hamiltonian of the form $H = \frac{P^2}{2m} - V(r)$, which is symmetric under rotations, the angular momentum $\mathcal{L} = \vec{X} \times \vec{P}$ is conserved $[H, \mathcal{L}_i] = 0$ and the algebra $[\mathcal{L}_i, \mathcal{L}_j] = i\hbar\varepsilon_{ijk}\mathcal{L}_k$ leads to the following three commuting operators

$$[H, \mathcal{L}_z] = [H, \mathcal{L}^2] = [\mathcal{L}_z, \mathcal{L}^2] = 0. \tag{4.68}$$

- The common eigenfunctions of $\mathcal{L}^2$ and $\mathcal{L}_z$ are the spherical harmonics with eigenvalues

$$\mathcal{L}^2 Y_{lm} = \hbar^2 l(l+1) Y_{lm} \tag{4.69}$$

and

$$\mathcal{L}_z Y_{lm} = \hbar m Y_{lm} \tag{4.70}$$

with l and $|m| \leq l$ integer. The eigenvalue of $\mathcal{L}^2$ is $(2l+1)$ fold degenerate.

- The Schrödinger equation for the hydrogen atom can be solved by reducing the non-relativistic two-body problem to the one-body problem with reduced mass $\mu = m_1 m_2/M$ and a free center of mass motion with total mass $M = m_1 + m_2$.

With the formula for the Laplace operator

$$\Delta = \frac{1}{r}\frac{\partial^2}{\partial r^2} r - \frac{1}{r^2}\frac{\mathcal{L}^2}{\hbar^2} \tag{4.71}$$

and a separation ansatz in spherical coordinates the energy eigenvalues

$$E_n = -\frac{Z^2 \mathcal{R}}{n^2} = -2\mu \left(\frac{Z\hbar}{a_0 n}\right)^2, \tag{4.72}$$

are determined by the termination condition of the power series solution to the radial equation (4.57), which is related to the differential equation

$$xL''(x) + (2l + 2 - x)L'(x) - (l + 1 - n)L(x) = 0 \tag{4.73}$$

for the associated Laguerre polynomials

$$L_r^s(x) = \partial_x^s L_r(x) = \partial_x^s \, e^x \, \partial_x^r \, e^{-x} \, x^r, \qquad \text{with} \qquad r = n + l, \qquad s = 2l + 1 \tag{4.74}$$

by $x = 2\rho = 2\kappa r$ with $\kappa = \sqrt{\frac{-2\mu E_n}{\hbar^2}} = \frac{Z}{n a_0}$. The normalized wave functions are

$$u_{nlm} = \sqrt{\frac{(n-l-1)!(2\kappa^3)}{2n((n+1)!)^3}} \, (2\kappa r)^l e^{-\kappa r} \, L_{n+l}^{2l+1}(2\kappa r) \, Y_{lm}(\theta, \varphi) \tag{4.75}$$

where $n \in \mathbb{N}$ is the principal quantum number, $l < n$ the orbital quantum number and m the magnetic quantum number. Due to the approximation of a pure Coulomb interaction and electrons without spin E_n is n^2-fold degenerate.

Chapter 5

Angular Momentum and Spin

I think you and Uhlenbeck have been very lucky to get your spinning electron published and talked about before Pauli heard of it. It appears that more than a year ago Kronig believed in the spinning electron and worked out something; the first person he showed it to was Pauli. Pauli rediculed the whole thing so much that the first person became also the last ...

– Thompson (in a letter to Goudsmit)

The first experiment that is often mentioned in the context of the electron's spin and magnetic moment is the Einstein–de Haas experiment. It was designed to test Ampère's idea that magnetism is caused by "molecular currents". Such circular currents, while generating a magnetic field, would also contribute to the angular momentum of a ferromagnet. Therefore a change in the direction of the magnetization induced by an external field has to lead to a small rotation of the material in order to preserve the total angular momentum.

For a quantitative understanding of the effect we consider a charged particle of mass m and charge q rotating with velocity v on a circle of radius r. Since the particle passes through its orbit $v/(2\pi r)$ times per second the resulting current $I = qv/(2\pi r)$, which encircles an area $A = r^2\pi$, generates a magnetic dipole moment $\mu = IA/c$,

$$I = \frac{qv}{2\pi r} \qquad \Rightarrow \qquad \mu = \frac{IA}{c} = \frac{qv\ r^2\pi}{2\pi r\ c} = \frac{qvr}{2c} = \frac{q}{2mc}L = \gamma L, \qquad \gamma = \frac{q}{2mc}, \tag{5.1}$$

where $\vec{L} = m\vec{r} \times \vec{v}$ is the angular momentum. Now the essential observation is that the **gyromagnetic ratio** $\gamma = \mu/L$ is independent of the radius of the motion. For an arbitrary distribution of electrons with mass m_e and elementary charge e we hence expect

$$\vec{\mu}_e = g\,\mu_B\,\frac{\vec{L}}{\hbar} \qquad \text{with} \qquad \mu_B = \frac{e\hbar}{2m_e c}, \tag{5.2}$$

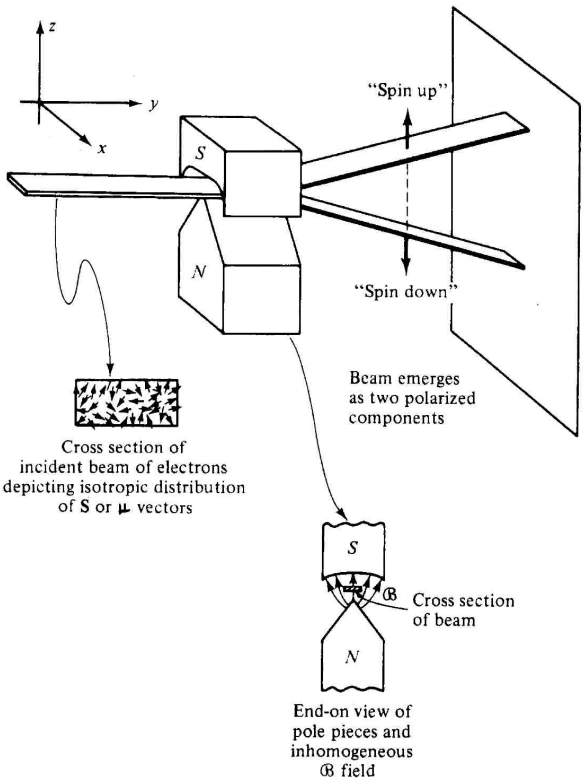


Figure 5.1: Splitting of a beam of silver atoms in an inhomogeneous magnetic field.

where the **Bohr magneton** μ_B is the expected ratio between the magnetic moment $\vec{\mu}_e$ and the dimensionless value $\vec{L}/\hbar$ of the angular momentum. The **g-factor** parametrizes deviations from the expected value $g = 1$, which could arise, for example, if the charge density distribution differs from the mass density distribution. The experimental result of Albert Einstein and Wander Johannes de Haas in 1915 seemed to be in agreement with Lorentz's theory that the rotating particles causing ferromagnetism are electrons.[1]

Classical ideas about angular momenta and magnetic moments of particles were shattered by the results of the experiment of Otto Stern and Walther Gerlach in 1922, who sent a beam of Silver atoms through an inhomogeneous magnetic field and observed a split into two beams as shown in *fig. 5.1.* The magnetic interaction energy of a dipole $\vec{\mu}$ in a magnetic field $\vec{B}$ is

$$E = \vec{\mu}\vec{B} = \gamma\vec{L}\vec{B} \tag{5.3}$$

which imposes a force $\vec{F} = -\vec{\nabla}(\vec{\mu}\vec{B})$ on the dipole. If the beam of particles with magnetic dipoles $\vec{\mu}$ passes through the central region where $B_z \gg B_x, B_y$ and $\frac{\partial B_z}{\partial z} \gg \frac{\partial B_z}{\partial x}, \frac{\partial B_z}{\partial y}$ the force

$$F_z \approx -\gamma L_z \frac{\partial B_z}{\partial z} \tag{5.4}$$

[1] The experiment was repeated by Emil Beck in 1919 who found "very precisely half of the expected value" for L/μ, which we now know is correct. At that time, however, g was still believed to be equal to 1. As a result Emil Beck only got a job as a high school teacher while de Haas continued his scientific career in Leiden.

points along the z-axis. It is proportional to the gradient of the magnetic field, which hence needs to be inhomogeneous. For an unpolarized beam the classical expectation would be a continuous spreading of deflections. Quantum mechanically, any orbital angular momentum L_z would be quantized as $L_z = m\hbar$ with an odd number $m = -l, 1-l, \ldots, l-1, l$ of split beams. But Stern and Gerlach observed, instead, *two* distinct lines as shown in *fig. 5.2*.

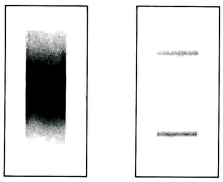


Figure 5.2: Stern and Gerlach observed two distinct beams rather than a classical continuum.

In 1924 Wolfgang Pauli postulated two-valued quantum degrees of freedom when he formulated his exclution principle, but he first opposed the idea of rotating electrons. In 1926 Samuel A. Goudsmit and George E. Uhlenbeck used that idea, however, to successfully guess formulas for the hyperfine splitting of spectral lines,[2] which involved the correct spin quantum numbers. Pauli pointet out an apparent discrepancy by a factor of two between theory and experiment, but this issue was resolved by Llewellyn Thomas. Thus Pauli dropped his objections and formalized the quantum mechanical theory of spin in 1927.

The unexpected experimental value $g = 2$ for the electron's g-factor could only be understood in 1928 when Paul A.M. Dirac found the relativistic generalization of the Schrödinger equation, which we will discuss in chapter 7. Almost 20 years later, Raby et al. discovered a deviation of the magnetic moment from Dirac theory in 1947, and at the same time Lamb et al. reported similar effects in the spectral lines of certain atomic transitions. By the end of that year Julian Schwinger had computed the leading quantum field theoretical correction a_e to the quantum mechanical value,

$$g/2 \equiv 1 + a_e = 1 + \frac{\alpha}{2\pi} + \mathcal{O}(\alpha^2) = 1 + 0.001161 + \mathcal{O}(\alpha^2) \tag{5.5}$$

and within a few years Schwinger, Feynman, Dyson, Tomonaga and others developed quantum electrodynamics (QED), the quantum field theory (QFT) of electrons and photons, to a level that allowed the consistent computation of perturbative corrections. Present theoretical calculations of the anomalous magnetic moment a_e of the electron, which include terms through order α^4, also need to take into account corrections due to strong and weak nuclear forces. The impressive agreement with the experimental result

$$a_e = \begin{cases} 0.001\,159\,652\,1884\ (43) & \text{experimental} \\ 0.001\,159\,652\,2012\ (27) & \text{theory (QFT)} \end{cases} \tag{5.6}$$

[2] The history as told by Goudsmit can be found in his very recommendable jubilee lecture, whose transcript is available at `http://www.lorentz.leidenuniv.nl/history/spin/goudsmit.html`.

shows the remarkable precision of QFT, which is the theoretical basis of elementary particle physics. Modern precision experiments measure a_e in Penning traps, which are axially symmetric combinations of a strong homogeneous magnetic field with an electric quadrupole, in which single particles or ions can be trapped, stored and worked with for several weeks.[3]

5.1 Quantization of angular momenta

Compelled by the experimental facts discussed above we now investigate general properties of angular momenta in order to find out how to describe particles with spin $\hbar/2$. In the previous chapter we found the commutation relations

$$[L_i, L_j] = i\hbar\epsilon_{ijk}L_k \qquad \Rightarrow \qquad [\vec{L}^2, L_i] = 0 \tag{5.7}$$

of the *orbital* angular momentum $\vec{L} = \vec{X} \times \vec{P}$ with eigenfunctions Y_{lm} and eigenvalues

$$\vec{L}^2 Y_{lm} = \hbar^2 l(l+1)Y_{lm}, \qquad L_z Y_{lm} = \hbar m Y_{lm} \tag{5.8}$$

of $\vec{L}^2$ and L_z so that all states with total angular momentum quantum number l come in an odd number $2l+1$ of incarnations with magnetic quantum number $m = -l, \dots l$. We thus want to understand how the electron can have an even number *two* of incarnations, as is implied by the Stern–Gerlach experiment and also by the double occupation of orbitals allowed by the Pauli principle.

If we think of the total angular momentum $\vec{J} = \vec{L} + \vec{S}$ as the sum of a (by now familiar) orbital part $\vec{L}$ and an (abstract) spin operator $\vec{S}$ then it is natural to expect that the total angular momentum $\vec{J}$ should obey the same kind of commutation relations

$$[J_i, J_j] = i\hbar\epsilon_{ijk}J_k \qquad \Rightarrow \qquad [\vec{J}^2, J_i] = 0 \tag{5.9}$$

so that, for example, the commutator with $\frac{1}{i\hbar}J_z$ rotates J_x into J_y and J_y into $-J_x$. We will now refrain, however, from a concrete interpretation of $\vec{J}$ and call any collection of three self-adjoint operators $J_i = J_i^\dagger$ obeying (5.9) an *angular momentum algebra*. Like in the case of the harmonic oscillator we will see that this algebra is sufficient to determine all eigenstates, which

[3] The high precision of almost 12 digits can be achieved because $1+a_e = \omega_s/\omega_c$ is the ratio of two frequencies, the spin flip frequency $\omega_s = g\mu_B B_z/\hbar$ and the cyclotron frequency $\omega_c = \frac{e}{m_e c}B_z$, and independent of the precise value of B_z. The cyclotron frequeny corresonds to the energy spacings between the Landau levels of electrons circling in a magnetic field: Landau invented a nice trick for the computation of the associated energy quanta: If we use the gauge $\vec{A} = B_z X \vec{e}_y$ for a magnetic field $\vec{B} = B_z\vec{e}_z$ in z-direction then the Hamiltonian becomes $H = \frac{1}{2m_e}\left(P_x^2 + (P_y - \frac{e}{c}B_z X)^2 + P_z^2\right)$. The operators P_x and $\tilde{X} = X - \frac{c}{eB_z}P_y$, which determine the dynamics in the xy-plan via the Hamiltonian $H_{xy} = \frac{1}{2m_e}P_x^2 + \frac{e^2B_z^2}{2m_e c^2}\tilde{X}^2$ of a harmonic oscillator, obviously satisfy a Heisenberg algebra $[P_x, \tilde{X}] = \frac{\hbar}{i}$. Recalling the (algebraic) solution of the harmonic oscillator we thus obtain the energy eigenvalues $E_n = (n+\frac{1}{2})\hbar\omega_c$ of the Landau levels with $\omega_c = \frac{eB_z}{m_e c} = 2\mu_B B_z/\hbar$ [Landau-Lifschitz].

we denote by $|j,\mu\rangle$. The eigenvalues of the maximal commuting set of operators J^2 and J_z can be parametrized as

$$\vec{J}^2\,|j,\mu\rangle = \hbar^2 j(j+1)\,|j,\mu\rangle \tag{5.10}$$

$$J_z\,|j,\mu\rangle = \hbar\,\mu\,|j,\mu\rangle \tag{5.11}$$

with $j \geq 0$ because $\vec{J}^2$ is non-negative. We also impose the normalization $\langle j',\mu'|j,\mu\rangle = \delta_{jj'}\delta_{\mu\mu'}$, where orthogonality for different eigenvalues is implied by $J_i^\dagger = J_i$.

Ladder Operators. As usual in quantum mechanics the general strategy is to diagonalize as many operators as possible. In order to diagonalize the action of J_z, i.e. a rotation in the xy-plane, we define the ladder operators

$$J_\pm = J_x \pm iJ_y, \tag{5.12}$$

where an analogy with the Harmonic oscillator would relate H to J_3 and $(\mathcal{X},\mathcal{P})$ to (J_x,J_y), which are transformed into one another by the commutator with H and J_3, respectively. In any case we find

$$[J_z,J_\pm] = [J_z,J_x] \pm i[J_z,J_y] = i\hbar J_y \mp i(i\hbar)J_x = \pm\hbar J_\pm \tag{5.13}$$

and

$$[J_+,J_-] = [J_x+iJ_y, J_x-iJ_y] = -i[J_x,J_y] + i[J_y,J_x] = 2\hbar J_z. \tag{5.14}$$

Since $[J_z,J_\pm] = \pm\hbar J_\pm$ the ladder operators $J_\pm$ shift the eigenvalues of J_z by $\pm\hbar$,

$$J_z\ J_\pm\,|j,\mu\rangle = (J_\pm J_z \pm \hbar J_\pm)|j,\mu\rangle = \hbar(\mu\pm 1)\ J_\pm|j,\mu\rangle, \tag{5.15}$$

so that

$$J_\pm\,|j,\mu\rangle = N_\pm\,|j,\mu\pm 1\rangle. \tag{5.16}$$

Since the eigenstates $|j,\mu\rangle$ are normalized by assumption, the normalization factors $N_\pm$ can be computed by evaluating the norms

$$||\,(J_\pm\,|j,\mu\rangle)\,||^2 = \langle j,\mu|J_\mp J_\pm|j,\mu\rangle = |N_\pm|^2\langle j,\mu\pm 1|j,\mu\pm 1\rangle = |N_\pm|^2, \tag{5.17}$$

where we used that $J_\pm^\dagger = J_\mp$. The expectation values of $J_\mp J_\pm$ are evaluated by relating these operators to J^2 and J_z. We first compute

$$J_\mp J_\pm = (J_x \mp iJ_y)(J_x \pm iJ_y) = J_x^2 + J_y^2 \pm i[J_x,J_y] = J_x^2 + J_y^2 \mp \hbar J_z \tag{5.18}$$

and since $J^2 = J_x^2 + J_y^2 + J_z^2$ we can express everything in terms of the diagonalized operators

$$J_\mp J_\pm = J^2 - J_z^2 \mp \hbar J_z \tag{5.19}$$

and obtain

$$|N_\pm|^2 = \hbar^2 \left(j(j+1) - \mu(\mu \pm 1)\right) = \hbar^2 \left(j \mp \mu)(j \pm \mu + 1)\right) \tag{5.20}$$

so that we end up with the important formula

$$\boxed{J_\pm |j, \mu\rangle = \hbar\sqrt{(j \mp \mu)(j \pm \mu + 1)}\, |j, \mu \pm 1\rangle} \tag{5.21}$$

for the ladder operators in the basis $|j, \mu\rangle$.

Quantization. The quantization condition for j can now be derived as follows. Since J_x^2 and J_y^2 are positive operators $J^2 = J_x^2 + J_y^2 + J_z^2 \geq J_z^2$, so that all eigenvalues of J_z^2 are bounded by the eigenvalue of J^2,

$$|\mu| \leq \sqrt{j(j+1)}. \tag{5.22}$$

For fixed total angular momentum quantum number j we conclude that μ is bounded from below and from above. Since the ladder operators $J_\pm$ do not change j, repeated raising and repeated lowering must both terminate,

$$J_+ |j, \mu_{max}\rangle = 0, \qquad J_- |j, \mu_{min}\rangle = 0. \tag{5.23}$$

But this implies

$$J_- J_+ |j, \mu_{max}\rangle = |N_+|^2 |j, \mu_{max}\rangle = 0, \tag{5.24}$$

$$J_+ J_- |j, \mu_{min}\rangle = |N_-|^2 |j, \mu_{min}\rangle = 0, \tag{5.25}$$

and hence

$$\mu_{min} = -j, \quad \mu_{max} = j, \qquad \mu_{max} - \mu_{min} = 2j \in \mathbb{N}_0 \tag{5.26}$$

where $2j$ must be a non-negative integer because we get from $|j, \mu_{min}\rangle$ to $|j, \mu_{max}\rangle$ with $(J_+)^k$ for $k = \mu_{max} - \mu_{min} = 2j$. We thus have shown that quantum mechanical spins are quantized in half-integral units $j \in \frac{1}{2}\mathbb{N}_0$ with μ ranging from $-j$ to j in integral steps. The magnetic quantum number hence can have $2j + 1$ different values for fixed total angular momentum. In particular, a doublet like observed in Stern–Gerlach is consistent and implies $j = 1/2$.

Naively one might expect that the eigenvalue of J^2 is the square of the maximal eigenvalue of J_z. But this is not possible because of an uncertainty relation, as can be seen from the following chain of inequalities:

$$J^2 = J_x^2 + J_y^2 + J_z^2 \geq J_z^2 + (\Delta J_x)^2 + (\Delta J_y)^2 \geq J_z^2 + 2\Delta J_x \Delta J_y \tag{5.27}$$

because $A^2 = (\Delta A)^2 + (\langle \Delta A \rangle)^2 \geq (\Delta A)^2$ and $(a - b)^2 = a^2 + b^2 - 2ab \geq 0$. Combining this with the uncertainty relation $\Delta J_x \Delta J_y \geq \frac{1}{2}|\langle [J_x, J_y] \rangle|$, where $[J_x, J_y] = i\hbar J_z$, we obtain

$$J^2 \geq J_z^2 + \hbar |J_z| = \hbar^2(\mu^2 + |\mu|). \tag{5.28}$$

This explains our parametrization of the eigenvalue of J^2 as $\hbar^2 j(j+1)$ and the above derivation of the eigenvalue spectrum shows that the inequality is saturated for $\mu_{max} = j$ and $\mu_{min} = -j$. We conclude that it does not make sense to think of the angular momentum of a particle as pointing into a particular direction: Due to the uncertainty relation between J_i and J_j for $i \neq j$ it is impossible to simultaneously measure different components of the angular moment, just like it is impossible to measure position and momentum of a particle simultaneously. Expectation values $\langle\psi|\vec{J}|\psi\rangle$, on the other hand, are usual vectors that do point into a particular direction.

5.2 Electron spin and the Pauli equation

According to general arguments of rotational invariance and angular momentum conservation we expect that the total angular momentum $\vec{J} = \vec{L} + \vec{S}$ is the sum of an intrinsic and an orbital part, which can be measured independently and hence ought to commute,

$$\vec{J} = \vec{L} + \vec{S}, \qquad [L_i, S_j] = 0. \tag{5.29}$$

Moreover, each of these angular momentum vectors obeys the same kind of algebra

$$[J_i, J_j] = i\hbar\varepsilon_{ijk}J_k, \qquad [L_i, L_j] = i\hbar\varepsilon_{ijk}L_k, \qquad [S_i, S_j] = i\hbar\varepsilon_{ijk}S_k. \tag{5.30}$$

and transforms as a vector under a rotation of the complete system (i.e. of the position *and* of the spin of a particle)

$$[J_i, L_j] = i\hbar\varepsilon_{ijk}L_k, \qquad [J_i, S_j] = i\hbar\varepsilon_{ijk}S_k, \qquad [J_i, J_j] = i\hbar\varepsilon_{ijk}J_k. \tag{5.31}$$

More generally, we can decompose the total angular momentum into a sum of (commuting) contributions of independent subsystens $\vec{J} = \sum_i \vec{L}_{(i)} + \sum_i \vec{S}_{(i)}$ for systems composed of several spinning particles with respective orbital angular momenta $\vec{L}_{(i)} = \vec{X}_{(i)} \times \vec{P}_{(i)}$ and spins $\vec{S}_{(i)}$.

We now focus on the spin degree of freedom and consider the case $s = \frac{1}{2}$ that is relevant for electrons, protons and neutrons. The basis in which S^2 and S_z are diagonal consists of two states $|\frac{1}{2}, \pm\frac{1}{2}\rangle$ which span the Hilbert space $\mathcal{H} = \mathbb{C}^2$. We can hence identify $|\frac{1}{2}, \pm\frac{1}{2}\rangle$ with the natural basis vectors $e_1 = \binom{1}{0}$ and $e_2 = \binom{0}{1}$. In order to save some writing it is useful to introduce the abbreviations $|\pm\rangle = |\frac{1}{2}, \pm\frac{1}{2}\rangle$ and

$$|\tfrac{1}{2}, +\tfrac{1}{2}\rangle = |+\rangle = |\uparrow\rangle \equiv \binom{1}{0}, \qquad |\tfrac{1}{2}, -\tfrac{1}{2}\rangle = |-\rangle = |\downarrow\rangle \equiv \binom{0}{1}. \tag{5.32}$$

According to (5.21) the action of the spin operators is

$$S_z|\uparrow\rangle = +\tfrac{\hbar}{2}|\uparrow\rangle, \qquad S_+|\uparrow\rangle = 0, \qquad S_-|\uparrow\rangle = \hbar|\downarrow\rangle \tag{5.33}$$

$$S_z|\downarrow\rangle = -\tfrac{\hbar}{2}|\downarrow\rangle, \qquad S_+|\downarrow\rangle = \hbar|\uparrow\rangle, \qquad S_-|\downarrow\rangle = 0 \tag{5.34}$$

which corresponds to the matrices

$$S_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad S_+ = \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad S_- = \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \tag{5.35}$$

in the natual basis. Solving $S_\pm = S_x \pm iS_y$ for $S_x = \frac{1}{2}(S_+ + S_-)$ and $S_y = \frac{1}{2i}(S_+ - S_-)$ we find

$$S_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad S_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad S_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.36}$$

We hence can write the spin operator as

$$\vec{S} = \frac{\hbar}{2}\vec{\sigma} \tag{5.37}$$

where $\vec{\sigma}$ are the **Pauli–matrices**

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.38}$$

The σ–matrices obey the following equivalent sets of identities,

$$(\sigma_i)^2 = \mathbb{1}, \qquad \sigma_i\sigma_j = -\sigma_j\sigma_i = i\varepsilon_{ijk}\sigma_k \quad \text{for} \quad i \neq j, \tag{5.39}$$

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij}, \qquad [\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k, \tag{5.40}$$

$$\boxed{\sigma_i\sigma_j = \delta_{ij}\mathbb{1} + i\varepsilon_{ijk}\sigma_k.} \tag{5.41}$$

They are traceless,

$$\operatorname{tr}\sigma_i = 0, \qquad \operatorname{tr}\sigma_i\sigma_j = 2\delta_{ij}, \tag{5.42}$$

and it is easily checked for Hermitan 2×2 matrices $A = A^\dagger$ that

$$A = \tfrac{1}{2}\left(\mathbb{1}\operatorname{tr}A + \vec{\sigma}\,\operatorname{tr}(A\vec{\sigma})\right), \tag{5.43}$$

where $\vec{\sigma}\operatorname{tr}A\vec{\sigma} \equiv \sum_{i=1}^3 \sigma_i \operatorname{tr}A\sigma_i$ is a linear combination of the three matrices σ_i with coefficients $\operatorname{tr}(A\sigma_i)$. The Pauli matrices hence form a basis for the 3-dimensional linear space of all traceless Hermitian 2×2 matrices.

The complete state of an electron is now specified by the position and spin degrees of freedom $|\vec{x} \in \mathbb{R}^3, s = \frac{1}{2}, \mu = \pm\frac{1}{2}\rangle$. Once we agree that electrons have spin $s = \frac{1}{2}$ we can omit this redundant information. In the S_z basis we find

$$|\psi\rangle = \sum_{\mu=\pm} |x,\mu\rangle\langle x,\mu|\psi\rangle = \psi_+(x)\,|\tfrac{1}{2},\tfrac{1}{2}\rangle + \psi_-(x)\,|\tfrac{1}{2},-\tfrac{1}{2}\rangle = \psi_+(x)|\uparrow\rangle + \psi_-(x)|\downarrow\rangle = \begin{pmatrix}\psi_+(x)\\ \psi_-(x)\end{pmatrix}, \tag{5.44}$$

i.e. the spinning electron is described by *two wave functions* $\psi_\pm(x)$.[4]

[4] More familiarly, a vector field $\vec{v}(x)$, i.e. a wave function with spin 1, is described by *three* component functions $v_i(x)$.

5.2.1 Magnetic fields: Pauli equation and spin-orbit coupling

Due to the experimental value $g = 2$ the total magnetic moment of the electron is

$$\vec{\mu}_{total} = \frac{e}{2m_e c}\left(\vec{L} + 2\vec{S}\right) = \frac{e}{2m_e c}\left(\vec{L} + \hbar\vec{\sigma}\right) \tag{5.45}$$

and the corresponding interaction energy with a magnetic field is

$$H_{int} = \vec{\mu}_{total}\vec{B} = \mu_B\left(\tfrac{1}{\hbar}\vec{L} + \vec{\sigma}\right)\vec{B} = \frac{\mu_B}{\hbar}\left(\vec{L} + 2\vec{S}\right)\vec{B}, \tag{5.46}$$

where $\mu_B = \frac{e\hbar}{2m_e c}$ is the Bohr magneton. The complete Hamiltonian thus becomes

$$H_{Pauli} = \frac{\vec{P}^{\,2}}{2m} + V(\vec{x}) + \frac{\mu_B}{\hbar}\left(\vec{L} + 2\vec{S}\right)\vec{B}. \tag{5.47}$$

The corresponding Schrödinger equation is called **Pauli equation** (without spin-orbit coupling)

$$i\hbar\frac{\partial\psi}{\partial t} = \left(-\frac{\vec{\hbar}^{\,2}}{2m}\Delta + V(\vec{x}) + \frac{\mu_B}{\hbar}\left(\vec{L} + 2\vec{S}\right)\vec{B}\right)\psi, \tag{5.48}$$

which is a system of differential equations for the two components $\psi_\pm(x)$ of the wave function **spinor** $\psi = \binom{\psi_+}{\psi_-}$ that are coupled by the magnetic interaction term $\vec{S}\vec{B}$.

Spin–Orbit coupling. When an electron moves with a velocity $\vec{v}$ in the electric field produced by a nucleus it observes, in its own frame of reference, a modified magnetic field according to the transformation rule

$$\vec{B}' = \vec{B} - \frac{1}{c}\left(\vec{v}\times\vec{E}\right), \tag{5.49}$$

where $\vec{B}$ is the magnetic field in the rest frame (of the nucleus). We can therefore try to take into account relativistic corrections to the Pauli equation by considering the interaction of this field with the magnetic moment $\vec{\mu}_e = \frac{e}{m_e c}\vec{S}$ of the electron.

If we ignore the weak magnetic field of the nucleus[5] $\vec{B} \approx 0$ then its electric field

$$\vec{E} = \frac{1}{e}\frac{dV}{dr}\frac{\vec{x}}{r} \tag{5.50}$$

induces a velocity-dependent magnetic field

$$\vec{B}' = -\frac{1}{ecr}\frac{dV}{dr}\underbrace{(\vec{v}\times\vec{x})}_{=-\frac{1}{m_e}\vec{L}} = \frac{1}{ecr}\frac{dV}{dr}\left(\frac{1}{m_e}\vec{L}\right). \tag{5.51}$$

The corresponding interaction energy $\Delta E = \vec{\mu}_e\vec{B}'$ suggests the spin-orbit correction

$$\vec{\mu}_e = \frac{e}{m_e c}\vec{S} \qquad \Rightarrow \qquad \Delta E_{\text{naiv}} = \frac{1}{m_e^2 c^2 r}\frac{dV}{dr}\left(\vec{L}\vec{S}\right). \tag{5.52}$$

[5] Note that the magnetic moment $\mu = g\frac{q}{2mc}$ is proportional to the inverse mass.

The correct **spin-orbit interaction** energy differs from this by a factor $\frac{1}{2}$ and will be derived from the fully relativistic Dirac equation in chapter 7,

$$H_{SO} = \frac{1}{2m_e^2c^2r}\frac{dV}{dr}\left(\vec{L}\vec{S}\right) = \frac{1}{2m_e^2c^2}\vec{L}\vec{S}\frac{Ze^2}{r^3} \tag{5.53}$$

where Z is the atomic number of the nucleus.

5.3 Addition of Angular Momenta

The spin-orbit interaction is an instance of the more general phenomenon that angular momenta $\vec{J}_1$ and $\vec{J}_2$ coming from different degrees of freedom interact so that only the total angular momentum $\vec{J} = \vec{J}_1 + J_2$ is conserved. In order to be able to take advantage of this conservation it is therefore necessary to reorganize the Hilbert space spanned by the $N = (2j_1+1)(2j_2+1)$ states $|j_1, m_1\rangle \otimes |j_2, m_2\rangle$ into a new basis $|j, m, \ldots\rangle$ in which

$$\vec{J}^2 = \vec{J}_1^2 + \vec{J}_2^2 + 2\vec{J}_1\vec{J}_2 \qquad \text{and} \qquad J_z = J_{1z} + J_{2z} \tag{5.54}$$

are diagonal. In order to simplify the notation we use, in the present section, the indices 1 and 2 exclusively to label the angular momenta $\vec{J}_1$ and $\vec{J}_2$, and we use x, y, z to label the different components. Since

$$[\vec{J}_1\vec{J}_2, J_{1z}] = [J_{1x}, J_{1z}]J_{2x} + [J_{1y}, J_{1z}]J_{2y} = -i\hbar J_{1y}J_{2x} + i\hbar J_{1x}J_{2y} \tag{5.55}$$

the commutator

$$[J^2, J_{1z}] = 2i\hbar(J_{1x}J_{2y} - J_{1y}J_{2x}) = -[J^2, J_{2z}] \tag{5.56}$$

is non-zero so that we can not diagonalize J_{1z} or J_{2z} simultaneously with J^2 and J_z. But $\vec{J}_1^2$ and $\vec{J}_2^2$ both commute with J^2 and J_z and we can continue to use their eigenvalues to label the states in the new basis by $|j, m, j_1, j_2\rangle$ with

$$J^2\ |j, m, j_1, j_2\rangle = \hbar^2 j(j+1)\quad |j, m, j_1, j_2\rangle, \tag{5.57}$$

$$J_z\ |j, m, j_1, j_2\rangle = \hbar\, m\qquad\quad |j, m, j_1, j_2\rangle, \tag{5.58}$$

$$J_1^2\ |j, m, j_1, j_2\rangle = \hbar^2 j_1(j_1+1)\ |j, m, j_1, j_2\rangle, \tag{5.59}$$

$$J_2^2\ |j, m, j_1, j_2\rangle = \hbar^2 j_2(j_2+1)\ |j, m, j_1, j_2\rangle. \tag{5.60}$$

During the course of our analysis we will show that j and m completely characterize the new basis so that no further independent commuting operators exist. For fixed j the angular momentum algebra implies that all magnetic quantum numbers with $-j \le m \le j$ are present and all of them can be obtained from a single one by repeated application of the ladder operators $J_\pm$. Such a **multiplet** of $2j+1$ states is called an **irreducible representation** R_j of the

$m = m_1 + m_2$	number of states
$j_1 + j_2$	1
$j_1 + j_2 - 1$	2
$j_1 + j_2 - 2$	3
...	...
$j_1 - j_2 + 1$	$2j_2$
$j_1 - j_2$	$2j_2 + 1$
$j_1 - j_2 - 1$	$2j_2 + 1$
...	...
$-j_1 + j_2 + 1$	$2j_2 + 1$
$-j_1 + j_2$	$2j_2 + 1$
$-j_1 + j_2 - 1$	$2j_2$
...	...
$-j_1 - j_2 + 2$	3
$-j_1 - j_2 + 1$	2
$-j_1 - j_2$	1

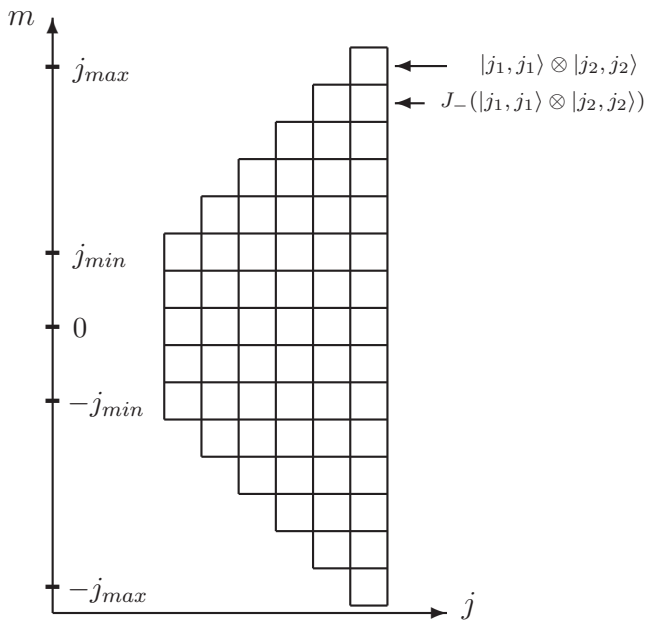


Table 5.1: Reorganization of the Hilber space of states $|j_1, m_1\rangle \otimes |j_2, m_2\rangle$.

angular momentum algebra and the change of basis that we are about to construct is called a decomposition of the tensor product $R_{j_1} \otimes R_{j_2}$ into a direct sum of irreducible representations R_j. The result of our analysis will be that $|j_1 - j_2| \leq j \leq j_1 + j_2$.

For fixed j_1 and j_2 there are $2j_1 + 1$ different values of m_1 and $2j_2 + 1$ different values of m_2. Since $J_z = J_{1z} + J_{2z}$ the eigenvalues in the old basis $|j_1 j_2 m_1 m_2\rangle = |j_1 m_1\rangle \otimes |j_2 m_2\rangle$ and in the new basis $|j\, j_1 j_2\, m\rangle$ are related by $m = m_1 + m_2$, but since $[J^2, J_{1z}] \neq 0$ a specific vector $|j_1 j_2 m_1 m_2\rangle$ may contribute to states with different total angular momentum j. In order to find the possible values of j and the linear combinations of the eigenstates of J_{1z} and J_{2z} that are eigenstates of J^2 we organize the Hilbert space according to the total magnetic quantum number m, as shown in table 5.1. For fixed m we draw as many boxes in the respective row as there are independent combinations $m_1 + m_2 = m$ with $|m_1| \leq j_1$ and $|m_2| \leq j_2$. The numbers of boxes in one row are listed for the case $j_1 \geq j_2$ (otherwise exchange J_1 and J_2).

The next step is to understand the horizontal position of the boxes along the j-axis between $j_{min} = |j_1 - j_2|$ and $j_{max} = j_1 + j_2$. For the maximal value $j_1 + j_2$ of m there is only one box, which corresponds to the state $|j_1 m_1\rangle \otimes |j_2 m_2\rangle$. This box must belong to a spin multiplet R_j with angular momentum $j = j_1 + j_2$ because $j \geq m = j_1 + j_2$ and for a larger value of j a state with a larger value of m would have to exist. Hence $j_{max} = m_{1max} + m_{2max}$ and $|j_{max}, m_{max}\rangle = |j_1, j_2, m_{1max}, m_{2\,max}\rangle$. Having identified the state $|j, m\rangle$ with $j = m = j_1 + j_2$ we can obtain all other states of $R_{j_{max}}$ by repeated application of the lowering operator $J_- = J_{1-} + J_{2-}$ with

$$(J_-)^k |j_1 + j_2, j_1 + j_2\rangle = (J_{1-} + J_{2-})^k |j_1, j_2, j_1, j_2\rangle = \sum_{l=0}^{k} \binom{k}{l} (J_{1-})^l \, |j_1 j_1\rangle \otimes (J_{2-})^{k-l} \, |j_2 j_2\rangle. \quad (5.61)$$

Iterating the formula (5.21) we find $J_-^k|j,j\rangle = \hbar^k\, k!\sqrt{\binom{2j}{k}}\ |j,j-k\rangle$. With $m = j_1+j_2-k$, $m_1 = j_1 - l$ and $m_2 = j_2 - k + l = m - m_2$ we hence obtain

$$\sqrt{\tbinom{2j_1+2j_2}{j_1+j_2-m}}\ |j_1+j_2,m\rangle = \sum_{m_1+m_2=m} \sqrt{\tbinom{2j_1}{j_1-m_1}}\ |j_1,m_1\rangle \otimes \sqrt{\tbinom{2j_2}{j_2-m_2}}\ |j_2,m_2\rangle. \tag{5.62}$$

If we now remove the $2j_1+2j_2+1$ states of the form (5.62), i.e. the last column in table 5.1, then the important observation is that we are left with a single box of maximal m, which now is $m_{max} = j' = j_1+j_2-1$. Repeating the above argument we hence conclude that there is a unique angular momentum multiplet $R_{j'}$. The state with the largest magnetic quantum number $|j',j'\rangle$ is determined by orthogonality to the state $|j_1+j_2, j_1+j_2-1\rangle$ in (5.62), which has the same magnetic quantum number $m = j'$ but a different total angular momentum. Iteration of this procedure until all states are exhausted shows that there is a unique multiplet with total angular momentum j with

$$|j_1 - j_2| \le j \le j_1 + j_2. \tag{5.63}$$

As a check we count the number of states in the new basis. Assuming $j_1 \ge j_2$ the number of colums in table 5.1 is $2j_2+1$ and

$$\textstyle\sum_{j=j_1-j_2}^{j_1+j_2}(2j+1) = (2j_2+1)\frac{(2j_1+2j_2+1)+(2j_1-2j_2+1)}{2} = (2j_1+1)(2j_2+1) \tag{5.64}$$

in accord with the dimension of the tensor product. Having determined the range of eigenvalues of J^2 and having established their non-degeneracy we now turn to the discussion of the unitary change of basis.

5.3.1 Clebsch-Gordan coefficients

The matrix elements $\langle j_1 j_2 m_1 m_2|jmj_1j_2\rangle$ of the unitary change of basis

$$|jmj_1j_2\rangle = \textstyle\sum_{m_1+m_2=m} |j_1j_2m_1m_2\rangle\langle j_1j_2m_1m_2|jmj_1j_2\rangle \tag{5.65}$$

are called **Clebsch–Gordan (CG) coefficients**, for which a number of notations is used,

$$\langle j_1j_2m_1m_2|jmj_1j_2\rangle \equiv \langle j_1j_2m_1m_2|jm\rangle \equiv C^{j_1j_2j}_{m_1m_2m} \equiv C^j_{m_1m_2}. \tag{5.66}$$

For the shorthand notation $C^j_{m_1m_2}$ the values j_1, j_2 must be known from the context; also the order of the quantum numbers and the index position may vary. Note that the elements of the CG matrix can be chosen to be real, so that the CG matrix becomes orthogonal.

In chapter 6 we will need to know the CG coefficients for the total angular momentum $\vec{J} = \vec{L} + \vec{S}$ of a spin$-\frac{1}{2}$ particle. From (5.62) we can read off the CG coefficients

$$\langle j_1j_2m_1m_2|jm\rangle = \sqrt{\tbinom{2j_1}{j_1-m_1}\tbinom{2j_2}{j_2-m_2}/\tbinom{2j}{j-m}} \qquad \text{for} \qquad j = j_1+j_2. \tag{5.67}$$

Specializing to the case $j_1 = l$ and $j_2 = \frac{1}{2}$, so that $m_s = \pm\frac{1}{2}$ and $m_l = m_j \mp \frac{1}{2}$, we observe that $\binom{2s}{s\pm m_s} = 1$ and

$$\binom{2l}{l-m_l}/\binom{2j}{j-m_j} = \frac{(2l)!}{(l-m_l)!(l+m_l)!}\frac{(j-m_j)!(j+m_j)!}{(2l+1)!}, \tag{5.68}$$

which yields the first line of the orthogonal matrix

$C^j_{m_l,m_s}$	$m_s = \frac{1}{2}$	$m_s = -\frac{1}{2}$
$j = l + \frac{1}{2}$	$\sqrt{\frac{l+m_j+1/2}{2l+1}}$	$\sqrt{\frac{l-m_j+1/2}{2l+1}}$
$j = l - \frac{1}{2}$	$-\sqrt{\frac{l-m_j+1/2}{2l+1}}$	$\sqrt{\frac{l+m_j+1/2}{2l+1}}$

. (5.69)

The second line follows from unitarity with signs chosen such that the determinant is positive.[6]

Explicit formulas for general CG coefficients have been derived by Racah and by Wigner (see [Grau] appendix A6 or [Messiah] volume 2). The coupling of three (or more) angular momenta can be analyzed by iteration. It is easy to see that now there are degeneracies, i.e. the resulting states are no longer uniquely described by $|j, m\rangle$ with $m = m_1 + m_2 + m_3$. The 8 states of the coupling of three spins $j_1 = j_2 = j_3 = \frac{1}{2}$, for example, organize themselves into a unique spin 3/2 "quartet" and two non-unique spin 1/2 "doublet" representations. The resulting basis depends on the order of the iteration $\vec{J} = \vec{J}_{12} + \vec{J}_3 = \vec{J}_1 + \vec{J}_{23}$ with $\vec{J}_{12} = \vec{J}_1 + \vec{J}_2$ and $\vec{J}_{23} = \vec{J}_2 + \vec{J}_3$. The recoupling coefficients, which are matrix elements of the corresponding unitary change of basis, can be expressed in terms of the Racah W-coefficients or in terms of the Wigher 6-j symbol, which essentially differ by sign conventions.[7] These quantities are used in atomic physics.

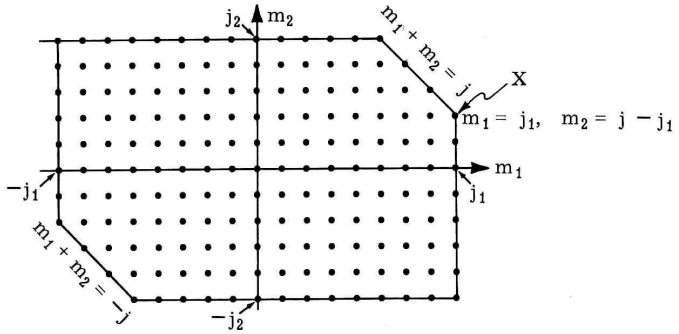
[6] Recursion relations that can be used to compute all CG coefficients follow from

$$\langle j_1, j_2, m_1, m_2|J_\pm|j, m\rangle = \langle j_1, j_2, m_1, m_2|\,(J_{1\pm} + J_{2\pm})\,|j, m\rangle \tag{5.70}$$

where $J_{i\pm}$ can be evaluated on the bra-vector and $J_\pm$ on the ket,

$$\begin{aligned}\sqrt{(j\mp m)(j\pm m+1)}\langle j_1, j_2, m_1, m_2|j, m\pm 1\rangle = &\sqrt{(j_1\pm m_1)(j_1\mp m_1+1)}\langle j_1, j_2, m_1\mp 1, m_2|j, m\rangle \\ &+\sqrt{(j_2\pm m_2)(j_2\mp m_2+1)}\langle j_1, j_2, m_1, m_2\mp 1|j, m\rangle.\end{aligned} \tag{5.71}$$

Possible values of m_1 and m_2 for fixed j_1, j_2 and j are shown in the following graphics, where the corner X



can be used as the starting point of the recursion because the linear equations (5.71) relate corners of the triangles $\begin{smallmatrix} m_1-1 & \\ & \searrow \\ m_2-1 & \end{smallmatrix}^{m_1,m_2}$ and $\begin{smallmatrix} & m_2+1 \\ \nearrow & \\ & m_1+1 \end{smallmatrix}^{m_1,m_2}$ (the overall normalization has to be determined from unitarity).

[7] The Wigner $6j$–symbols and the Racah W-coefficients, which describe the recoupling of 3 spins, are related

5.3.2 Singlet, triplet and EPR correlations

Another interesting case is the addition of two spin-1/2 operators

$$\vec{S} = \vec{S_1} + \vec{S_2} \tag{5.74}$$

The four states in the tensor product basis are $|\pm\rangle \otimes |\pm\rangle$, or $|\uparrow\uparrow\rangle$, $|\uparrow\downarrow\rangle$, $|\downarrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$. The total spin can have the values 0 and 1, and the respective multiplets, or representations, are called singlet and triplet, respectively.[8] The **triplet** consists of the states

$$|1,\ 1\rangle = |\uparrow\uparrow\rangle, \tag{5.75}$$

$$|1,\ 0\rangle = \tfrac{1}{\sqrt{2}\,\hbar} S_- |\uparrow\uparrow\rangle = \tfrac{1}{\sqrt{2}} \left(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\right), \tag{5.76}$$

$$|1,\text{–}1\rangle = \tfrac{1}{\sqrt{2}\,\hbar} S_- |1,0\rangle = |\downarrow\downarrow\rangle. \tag{5.77}$$

The **singlet** state is the superposition of the $S_z = 0$ states $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$ orthogonal to $|1,0\rangle$,

$$|0,\ 0\rangle = \tfrac{1}{\sqrt{2}} \left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right). \tag{5.78}$$

An important application for a system with two spins is the hyperfine structure of the ground state of the hydrogen atom, which is due to the magnetic coupling between the spins of the proton and of the electron. The energy difference between the ground state singlet and the triplet excitation corresponds to a signal with 1420.4 MHz, the famous 21 cm hydrogen line, which is used extensively in radio astronomy. Due to the weakness of the magnetic interaction the lifetime of the triplet state is about 10^7 years.

In order to derive a formula for the projectors to singlet and triplet states we note that $\vec{S}^2 = \vec{S_1}^{\,2} + \vec{S_2}^{\,2} + 2\vec{S_1}\vec{S_2}$ with

$$\vec{S}^2|1,m\rangle = 2\hbar^2|1,m\rangle, \qquad \vec{S}^2|0,0\rangle = 0 \qquad \text{and} \qquad \vec{S_1}^2 = \vec{S_2}^2 = \tfrac{3}{4}\hbar^2\mathbb{1}. \tag{5.79}$$

The operator $\vec{S_1}\vec{S_2} = \frac{1}{2}\vec{S}^2 - \frac{3}{4}\hbar^2\mathbb{1}$ therefore has eigenvalue $-\frac{3}{4}\hbar^2$ on the singlet and $\frac{1}{4}\hbar^2$ on triplet states. With the tensor product notation $\vec{S_1} = \vec{S} \otimes \mathbb{1}$ and $\vec{S_2} = \mathbb{1} \otimes \vec{S}$, i.e.

$$\vec{S_1}\,\vec{S_2} \equiv \vec{S} \otimes \vec{S} = \tfrac{\hbar^2}{4}\vec{\sigma} \otimes \vec{\sigma} \tag{5.80}$$

---

by

$$\left\{ \begin{matrix} j_1 & j_2 & J_{12} \\ j_3 & J & J_{23} \end{matrix} \right\} = (-1)^{j_1+j_2+j_3+J} W(j_1 j_2 J j_3; J_{12} J_{23}). \tag{5.72}$$

The Wigner $3j$–symbols and the Racah V-coefficients are related to the Clebsch Gordan coefficients by

$$\left\{ \begin{matrix} j_1 & j_2 & j \\ m_1 & m_2 & -m \end{matrix} \right\} = (-1)^{j-j_1+j_2} V(j_1 j_2 j; m_1 m_2 m) = \frac{(-1)^{m+j_1-j_2}}{\sqrt{2j+1}} \langle j_1 j_2 m_1 m_2 | j m \rangle. \tag{5.73}$$

Wigner's sign choices have the advantage of higher symmetry; see [Messiah] vol. II and, for example, `http://mathworld.wolfram.com/Wigner6j-Symbol.html` and `http://en.wikipedia.org/wiki/Racah_W-coefficient`.

[8] The German names for $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$ are Singulett, Dublett, Triplett, Quartett, ..., respectively.

we thus obtain the **projector P_0 onto** the **singlet** state

$$P_0 \;=\; |0,0\rangle\langle 0,0| \;=\; \tfrac{1}{4}\mathbb{1} - \tfrac{1}{\hbar^2}\vec{S}_1\vec{S}_2 = \tfrac{1}{4}\left(\mathbb{1} - \vec{\sigma}\otimes\vec{\sigma}\right) \tag{5.81}$$

and the **projector P_1 onto triplet states**

$$P_1 = \sum_{m=-1}^{1} |1,m\rangle\langle 1,m| = \tfrac{3}{4}\mathbb{1} + \tfrac{1}{\hbar^2}\vec{S}_1\vec{S}_2 = \tfrac{1}{4}\left(3\mathbb{1} + \vec{\sigma}\otimes\vec{\sigma}\right), \tag{5.82}$$

which we now apply to the computation of probabilities.

Correlations in the EPR experiment. We want to compute the conditional probability $P(\vec{\alpha}|\vec{\beta})$ for the measurement of spin up in the direction of a unit vector $\vec{\alpha}$ for one particle if we measure spin up in the direction of a unit vector $\vec{\beta}$ for the other decay product of a quantum mechanical system composed of two spin $\frac{1}{2}$ particles that has been prepared in the singlet state $|0,0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$, i.e. in Bohm's version of the EPR experiment.

The conditional probability $P(\vec{\alpha}|\vec{\beta}) = P(\vec{\alpha}\wedge\vec{\beta})/P(\vec{\beta})$ is the probability that we measure spin up in direction $\vec{\alpha}$ for the first particle and spin up in direction $\vec{\beta}$ for the second particle divided by the probability for the latter. It is easily seen that $\Pi_\pm = \frac{1}{2}(\mathbb{1}\pm\sigma_z)$ is the projector $|\pm\rangle\langle\pm|$ onto the S_z eigenstate $|\pm\rangle$. More generally, the spin operator $\vec{\alpha}\vec{S}$ for a spin measurement in the direction $\vec{\alpha}$ has eigenvalues $\pm\frac{\hbar}{2}$, so that $\vec{\alpha}\vec{\sigma}$ has eigenvalues ± 1. Therefore the projectors $\Pi_{\pm\vec{\alpha}}$ onto the eigenspaces of spin up and spin down in the direction of the unit vector $\vec{\alpha}$ are

$$\Pi_{\pm\vec{\alpha}} = \frac{1}{2}(\mathbb{1}\pm\vec{\alpha}\vec{\sigma}) \tag{5.83}$$

and the probability for spin up in direction $\vec{\beta}$ for the second particle in the singlet state is

$$\begin{aligned} P(\vec{\beta}) &= \operatorname{tr}(P_0\,\Pi^{(2)}_{\vec{\beta}}) = \operatorname{tr}\left(\tfrac{1}{4}(\mathbb{1}\otimes\mathbb{1} - \sigma_i\otimes\sigma_i)\circ(\mathbb{1}\otimes\Pi_{\vec{\beta}})\right) \\ &= \tfrac{1}{4}\operatorname{tr}\left(\mathbb{1}\otimes\Pi_{\vec{\beta}} - \sigma_i\otimes\sigma_i\Pi_{\vec{\beta}}\right) = \tfrac{1}{4}\left(\operatorname{tr}\mathbb{1}\cdot\operatorname{tr}\Pi_{\vec{\beta}} - \operatorname{tr}\sigma_i\cdot\operatorname{tr}(\sigma_i\Pi_{\vec{\beta}})\right) \end{aligned} \tag{5.84}$$

where we used the factorization $\operatorname{tr}(\mathcal{O}_1\otimes\mathcal{O}_2) = \operatorname{tr}\mathcal{O}_1\cdot\operatorname{tr}\mathcal{O}_2$ of traces of product operators shown in (3.68). In the two-dimensional one-particle spin spaces the traces are

$$\operatorname{tr}\mathbb{1} = 2, \qquad \operatorname{tr}\sigma_i = 0, \qquad \operatorname{tr}\Pi_{\vec{\beta}} = \tfrac{1}{2}\operatorname{tr}\mathbb{1} = 1, \qquad \operatorname{tr}(\sigma_i\Pi_{\vec{\beta}}) = \tfrac{1}{2}\operatorname{tr}(\sigma_i + \sigma_i\sigma_j\beta_j) = \beta_i \tag{5.85}$$

and we find $P(\vec{\beta}) = 1/2$, as we had to expect because of rotational symmetry. For the conditional probability we thus obtain

$$P(\alpha|\beta) = P(\alpha\wedge\beta)/P(\beta) = 2P(\alpha\wedge\beta) = 2\operatorname{tr}P_0\,\Pi^{(1)}_{\vec{\alpha}}\Pi^{(2)}_{\vec{\beta}} \tag{5.86}$$

$$= 2\operatorname{tr}\left(\tfrac{1}{4}\big(\mathbb{1}\otimes\mathbb{1} - \sigma_i\otimes\sigma_i\big)\circ\big(\Pi_{\vec{\alpha}}\otimes\Pi_{\vec{\beta}}\big)\right) \tag{5.87}$$

$$= \tfrac{1}{2}\left(\operatorname{tr}(\Pi_{\vec{\alpha}})\operatorname{tr}(\Pi_{\vec{\beta}}) - \operatorname{tr}(\sigma_i\Pi_{\vec{\alpha}})\operatorname{tr}(\sigma_i\Pi_{\vec{\beta}})\right) = \tfrac{1}{2}(1-\vec{\alpha}\vec{\beta}) = \tfrac{1}{2}\big(1-\cos(\vec{\alpha},\vec{\beta})\big), \tag{5.88}$$

which is the result we used in the discussion of EPR in chapter 3.

Chapter 6

Methods of Approximation

So far we have solved the Schrödinger equation for rather simple systems like the harmonic oscillator and the Coulomb potential. There are, however, many situations where exact solutions are not available. In the present section we will discuss a number of approximation techniques: Time independent perturbation theory (Rayleigh–Schrödinger), the variational method (Riesz) and time dependent perturbation theory. As applications we work out the fine structure of hydrogen, discuss the Zeeman and the Stark effect, compute the ground state energy of helium and derive the selection rules for the dipol approximation for absorption and emission of light.

6.1 Rayleigh–Schrödinger perturbation theory

Time independent (Rayleigh–Schrödinger) perturbation theory applies to quantum mechanical systems for which a time-independent Hamiltonian H is the sum of an exactly solvable part H_0 and a small perturbation,

$$H = H_0 + \lambda V, \tag{6.1}$$

where λ serves as a formal infinitesimal parameter that organizes the orders of the perturbative expansion. Of course only the smallness, in some appropriate sense, of the product λV will be relevant for the quality of the resulting approximation.

Non-degenerate time-independent perturbation theory. We first assume that the eigenvalue problem for H_0 has been solved

$$H_0|a0\rangle = E_{a0}|a0\rangle \tag{6.2}$$

with discrete[1] non-degenerate eigenvalues E_{a0} for some index set $\{a\}$ and corresponding eigenstates $|a0\rangle$. The label 0 refers to the unperturbed Hamiltonian H_0 and to the $0th$ approximation.

[1] If appropriate, sums have to be augmented by integrals over continuous parts of the spectra as usual.

We now make the ansatz that the exact solution to the eigenvalue problem

$$H|a\rangle = E_a|a\rangle \tag{6.3}$$

can be expanded into a power series

$$E_a = E_{a0} + \lambda E_{a1} + \lambda^2 E_{a2} + \ldots = \sum_{i=0}^{\infty} \lambda^i E_{ai}, \tag{6.4}$$

$$|a\rangle = |a0\rangle + \lambda|a1\rangle + \lambda^2|a2\rangle + \ldots = \sum_{i=0}^{\infty} \lambda^i |ai\rangle. \tag{6.5}$$

Even if E_a and $|a\rangle$ are not analytic in λ so that the expansion does not converge to the exact solution we may get excellent approximations for sufficiently small λ. In that case (6.4)–(6.5) is called an asymptotic expansion. A well-known example is the anomalous magnetic moment of the electron (5.6), for which the perturbation series is known to diverge.

The next step is to insert the ansatz into the stationary Schrödinger equation

$$(H_0 + \lambda V)\left(\sum_{i=0}^{\infty} \lambda^i |ai\rangle\right) = \left(\sum_{j=0}^{\infty} \lambda^j E_{aj}\right)\left(\sum_{k=0}^{\infty} \lambda^k |ak\rangle\right), \tag{6.6}$$

or

$$\sum_{i=0}^{\infty} \lambda^i H_0|ai\rangle + \sum_{i=0}^{\infty} \lambda^{i+1} V|ai\rangle = \sum_{i=0}^{\infty} \lambda^i \left(\sum_{l=0}^{i} E_{a,i-l}|al\rangle\right), \tag{6.7}$$

and to compare coefficients of λ^i. For λ^0 we obtain, of course, the unperturbed equation $H_0|a0\rangle = E_{a0}|a0\rangle$. For the higher order terms we bring the $|ai\rangle$–terms to the left-hand-side and collect all other terms on the other side,

- for λ^1:
$$(H_0 - E_{a0})|a1\rangle = (E_{a1} - V)|a0\rangle, \tag{6.8}$$
- for λ^2:
$$(H_0 - E_{a0})|a2\rangle = (E_{a1} - V)|a1\rangle + E_{a2}|a0\rangle, \tag{6.9}$$

  ...
- for λ^i:
$$(H_0 - E_{a0})|ai\rangle = (E_{a1} - V)|a, i-1\rangle + \ldots + E_{ai}|a0\rangle. \tag{6.10}$$

The best way to analyze the content of these equations is to compute their products with the complete set $\langle b0|$ of eigenstates of H_0.

First order corrections. When evaluating $\langle b0|$ on (6.8) we can replace $\langle b0|H_0$ in the first term by $\langle b0|E_{b0}$ and thus obtain

$$\langle b0|(E_{b0} - E_{a0})|a1\rangle = \langle b0|(E_{a1} - V)|a0\rangle = \delta_{ab}E_{a1} - \langle b0|V|a0\rangle. \tag{6.11}$$

We thus have to distinguish two cases:

- for $a = b$ the equation can be solved for

$$\boxed{E_{a1} = \langle a0|V|a0\rangle,} \tag{6.12}$$

  so that the first order energy correction is simply the expecation value of the perturbation.

- for $a \neq b$ the equation can be solved for

$$\langle b0|a1\rangle = -\frac{\langle b0|V|a0\rangle}{E_{b0} - E_{a0}}. \tag{6.13}$$

These scalar products are just the expansion coefficients of $|a1\rangle$ in the basis $|b0\rangle$ so that

$$\boxed{|a1\rangle = \sum_{b\neq a} |b0\rangle \, \frac{\langle b0|V|a0\rangle}{E_{a0} - E_{b0}},} \tag{6.14}$$

where we omitted a potential contribution of $|a0\rangle$ for reasons that we now discuss in detail.

If we consider the products of $\langle b0|$ with the equations (6.8)–(6.10) we observe that the l.h.s. vanishes for $b = a$ and the r.h.s. contains a term $E_{ai}\delta_{ab}$. The resulting equations hence determine the energy corrections E_{ai} for each order λ^i. For $b \neq a$ the assumption of non-degenerate energy levels $E_{b0} \neq E_{a0}$ implies that the l.h.s. becomes $(E_{b0} - E_{a0})\langle b0|ai\rangle$ so that these equations determine the expansion coefficients of $|ai\rangle$ in the basis $|b0\rangle$ for $b \neq a$. But $\langle a0|ai\rangle$ remains completely undetermined! The reason for this is easily understood: The Schrödinger equation is linear, and so are all equations that we derived with our perturbative ansatz. Every solution $|a\rangle = |a0\rangle + \lambda|a1\rangle + \ldots$ can hence be rescaled by an overall factor $f(\lambda) = 1 + f_1\lambda + \ldots$, which would reorganize the perturbation series such that the coefficient of $|a0\rangle$ in the expansion of $|ai\rangle$ can be changed arbitrarily without impairing the orthonormalization $\langle a0|b0\rangle = \delta_{ab}$ of the unperturbed states. We are hence free to simply choose

$$\langle a0|ai\rangle = 0 \tag{6.15}$$

as is done, for example, in [Schwabl]. This is the most convenient way to fix the ambiguous coefficients. Then $|a1\rangle$ and all higher order corrections are orthogonal to $|a0\rangle$ so that (according to Pythagoras) the norm of $|a\rangle$ differs from the norm of $|a0\rangle$ by a positive correction term of order λ^2. If we want to keep $|a\rangle$ normalized at each order in perturbation theory then we can divide $|a\rangle$ as obtained with the choice (6.15) by its norm, which will keep $\langle a0|a1\rangle = 0$ but change $\langle a0|ai\rangle$ for $i \geq 2$ such that $|a\rangle$ stays normalized. We are now ready to proceed with the

Second order energy correction. Multiplicaton of $\langle b0|$ with (6.9) yields

$$\langle b0|(E_{b0} - E_{a0})|a2\rangle = \langle b0|(E_{a1} - V)|a1\rangle + \delta_{ab}E_{a2}. \tag{6.16}$$

For $a = b$ we use $\langle a0|a1\rangle = 0$ and solve for $E_{a2} = \langle a0|V|a1\rangle$, thus obtaining the energy correction

$$\boxed{E_{a2} = \sum_{b\neq a} \frac{|\langle a0|V|b0\rangle|^2}{E_{a0} - E_{b0}}} \tag{6.17}$$

which shows that the contribution of other states to the energy correction is proportional to the squared matrix element of the perturbation, but suppressed by an energy denominator for states located at very different energy levels. It is straightforward to work out the second order corrections to the wave function[2] but since our interest is usually focused on energy spectra this would only be relevant for the computation of E_{a3}.

6.1.1 Degenerate time independent perturbation theory

In the above derivation we used that the energy levels are non-degenerate and the results (6.14) and (6.17) show that we get in trouble with vanishing energy denominators if $E_{a0} = E_{b0}$ for $a \neq b$. Indeed, equation (6.11) becomes inconsistent if there is an offdiagonal matrix element $\langle b0|V|a0\rangle$ between states with degenerate unperturbed energies. The reason is easily understood because the choice of basis is ambiguous within the eigenspace for a particular eigenvalue and when the degeneracy is lifted by λV then an arbitrarily small perturbation requires a non-infinitesimal change of basis. This is inconsistent with a perturbative ansatz. What we thus need to do is to diagonalize the matrix $\langle b0|V|a0\rangle$ within each degeneration space $E_{a0} = E_{b0}$ so that the perturbative ansatz becomes consistent. The vanishing r.h.s. of (6.11) thus implies

$$E_{\hat{a}1} = \langle \hat{a}0|V|\hat{a}0\rangle \qquad \text{with} \qquad \langle \hat{b}0|V|\hat{a}0\rangle = 0 \quad \text{for} \quad E_{\hat{a}0} = E_{\hat{b}0},\ \hat{a} \neq \hat{b} \tag{6.19}$$

so that the eigenvalues $E_{\hat{a}}$ of $\langle b0|V|a0\rangle$ yield the first order energy corrections.

6.2 The fine structure of the hydrogen atom

The fine-structure of the hydrogen atom, which partially lifts the degeneracies of the pure Coulomb interaction that we observed in chapter 4, is an important application of degenerate perturbation theory. The relevant Hamiltonian consists of the relativistic corrections

$$H = \underbrace{\frac{\vec{P}^2}{2m_e} + V(r)}_{H_0} - \underbrace{\frac{\vec{P}^4}{8m_e^3c^2}}_{H_{RK}} + \underbrace{\frac{1}{2m_e^2c^2r}\frac{dV}{dr}\vec{L}\vec{S}}_{H_{SO}} + \underbrace{\frac{\hbar^2}{8m_e^2c^2}\Delta V(r)}_{H_D} \tag{6.20}$$

[2] The solution $\langle b0|a2\rangle = \left(\langle b0|V|a1\rangle - E_{a1}\langle b0|a1\rangle\right)/(E_{a0} - E_{b0})$ of (6.16) leads to the second order wave function correction

$$|a2\rangle = \textstyle\sum_{b\neq a} |b0\rangle \left(\sum_{c\neq a} \frac{\langle b0|V|c0\rangle}{E_{a0}-E_{b0}} \frac{\langle c0|V|a0\rangle}{E_{a0}-E_{c0}} - \frac{E_{a1}\langle b0|V|a0\rangle}{(E_{a0}-E_{b0})^2}\right) - \rho|a0\rangle, \tag{6.18}$$

where the normalization of $|a\rangle$ at order λ^2 requires $\rho = \frac{1}{2}\langle a1|a1\rangle$.

with $V(r) = -\frac{Ze^2}{r}$, where we dropped the unobservable zero-point energy mc^2 of the electron[3] (as well as that of the nucleus). We now discuss the fine structure corrections

H_{RK} ... relativistic kinetic energy correction
H_{SO} ... spin orbit coupling
H_D ... Darwin term

in turn. At the end we will observe a surprising simplification of the complete result.

Relativistic kinetic energy. Since the electron is much lighter than the nucleus, its velocity is much larger and the resulting relativistic corrections to the kinetic energy E_{RK} set the scale for the size of the fine structure corrections. According to the Bohr model the velocity of the ground state orbit is

$$\frac{v}{c} = \frac{e^2}{\hbar c} = \alpha \approx \frac{1}{137} \qquad \Rightarrow \qquad \frac{E_{RK}}{E_0} \sim \frac{v^2}{c^2} \sim \alpha^2 \sim 10^{-4} \tag{6.21}$$

so that first order perturbation theory is sufficient for the derivation of very precise values for the level splittings. For small velocities the relativistic expression

$$E = \sqrt{m_e^2c^4 + P^2c^2} = m_ec^2\sqrt{1 + \left(\tfrac{P}{m_ec}\right)^2} \tag{6.22}$$

can be expanded into a power series using

$$\sqrt{1+x} = \textstyle\sum_{n=0}^{\infty} \binom{1/2}{n} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \ldots, \tag{6.23}$$

with the result

$$E = m_ec^2 + \tfrac{P^2}{2m_e} - \tfrac{1}{8}\tfrac{P^4}{m_e^3c^2} + \ldots \tag{6.24}$$

The second term is the non-relativistic kinetic energy, which is the kinetic part of the Schrödinger Hamiltonien H_0. Thus the leading relativistic correction is

$$H_{RK} = -\frac{P^4}{8m_e^3c^2}, \qquad \Delta E_{nl}^{RK} = \langle nlm|H_{RK}|nlm\rangle \tag{6.25}$$

where ΔE denotes the first order energy correction (6.12), which depends on the principal quantum number n and the orbital quantum number l. Since the momentum commutes with the spin operator, we can ignore the spin quantum numbers. Since the states $|nlm\rangle$ satisfies the unperturbed Schrödinger equation $H_0|nlm\rangle = E_n|nlm\rangle$ with energy

$$E_n = -\frac{Z^2}{n^2}\frac{m_ee^4}{2\hbar^2} = -\frac{Z^2e^2}{2n^2}\frac{1}{a_0} \tag{6.26}$$

we can use $P^2 = 2m_e(H_0 - V)$ with $V(r) = -\frac{Ze^2}{r}$ to trade expectation values of powers of P^2 for expectation values of powers of the Coulomb potential,

$$H_{RK} = -\frac{P^4}{8m_e^3c^2} = -\frac{1}{2m_ec^2}\left(\frac{P^2}{2m_e}\right)^2 = -\frac{1}{2m_ec^2}\left(H_0 + \frac{Ze^2}{r}\right)^2. \tag{6.27}$$

[3]While the replacement of the electron mass m_e by the reduced mass μ with $\frac{1}{\mu} = \frac{1}{m_p} + \frac{1}{m_e}$ leads to a larger absolute shift in the total binding energies than the fine structure, it can be ignored for the relative level splittings.

Inserting the expectation values

$$\left\langle \frac{1}{r} \right\rangle_{nl} = \frac{Z}{a_0 n^2} = -\frac{2E_n}{Ze^2}, \qquad a_0 = \frac{\hbar^2}{m_e e^2}, \tag{6.28}$$

$$\left\langle \frac{1}{r^2} \right\rangle_{nl} = \frac{Z^2}{a_0^2 n^3 (l+\frac{1}{2})} = \frac{4E_n^2}{Z^2 e^4} \frac{n}{l+\frac{1}{2}}, \tag{6.29}$$

where a_0 denotes the Bohr radius, the **relativistic kinetic energy** correction becomes

$$\Delta E_{nl}^{RK} = -\frac{m_e Z^4 e^8}{2n^4 c^2 \hbar^4} \left(\frac{n}{l+\frac{1}{2}} - \frac{3}{4} \right), \tag{6.30}$$

which is of order α^2 as expected.

Spin-orbit coupling. The Hamitonian H_{SO} of the spin-orbit coupling

$$H_{SO} = \frac{1}{2m_e^2 c^2 r} \frac{dV}{dr} \vec{L}\vec{S} = \frac{1}{2m_e^2 c^2} \vec{L}\vec{S} \frac{Ze^2}{r^3} \tag{6.31}$$

has been discussed already in chapter 5 and will be derived from the Dirac equation in chapter 7. For the evaluation of the radial part of the expectation value we need

$$\left\langle \frac{1}{r^3} \right\rangle_{nl} = \frac{Z^3 m_e^3 e^6}{\hbar^6} \frac{1}{n^3 l(l+1)(l+\frac{1}{2})}. \tag{6.32}$$

For the evaluation of the angular part we recall that $J^2 = (L+S)^2 = L^2 + S^2 + 2\vec{L}\vec{S}$ so that the spin-orbit term can be written as

$$\vec{L}\vec{S} = \tfrac{1}{2} \left(J^2 - L^2 - S^2 \right). \tag{6.33}$$

Since J^2 does not commute with L_z the spin-orbit Hamiltonian is not diagonal in the basis $|nlm_l\rangle \otimes |sm_s\rangle$ and we have to apply degenerate perturbation theory in the total angular momentum basis $|n\,jlsm_j\rangle$. The energy correction thus becomes

$$\langle n\,jlsm_j | H_{SO} | n\,jlsm_j \rangle = \frac{Ze^2}{4m_e^2 c^2} \, \hbar^2 (j(j+1) - l(l+1) - s(s+1)) \, \langle nl | \frac{1}{r^3} | nl \rangle \tag{6.34}$$

For an electron $s = \frac{1}{2}$ and therefor $j = l \pm \frac{1}{2}$. The spin-orbit coupling hence lifts, for example, the geneneracy of the 6 electron states $|21m\rangle$ in the p orbital with principal quantum number $n = 2$ and induces a fine-structure difference between the energy levels of the 4 states in the spin $\frac{3}{2}$ quartet $2p_{3/2}$ and the 2 states in the doublet $2p_{1/2}$. We hence need to distinguish the two cases $j = l \pm \frac{1}{2}$ and obtain

$$\Delta E_{njl}^{SO} = \frac{Ze^2}{4m_e c^2} \left\langle \frac{1}{r^3} \right\rangle_{nl} \begin{cases} \hbar^2 \left((l+\frac{1}{2})(l+\frac{3}{2}) - l(l+1) - \frac{1}{2}\frac{3}{2} \right) \\ \hbar^2 \left((l-\frac{1}{2})(l+\frac{1}{2}) - l(l+1) - \frac{1}{2}\frac{3}{2} \right) \end{cases} \tag{6.35}$$

$$= \frac{Ze^2}{4m_e c^2} \left\langle \frac{1}{r^3} \right\rangle_{nl} \begin{cases} \hbar^2 l \\ -\hbar^2 (l+1) \end{cases}. \tag{6.36}$$

Inserting (6.32) the final result becomes

$$\Delta E^{SO}_{njl} = \frac{m_e Z^4 e^8}{4c^2\hbar^4}\ \frac{1}{n^3 l(l+1)(l+\frac{1}{2})} \cdot \begin{cases} l & \text{for } j = l+\frac{1}{2},\ l>0, \\ (-l-1) & \text{for } j = l-\frac{1}{2},\ l>0. \end{cases} \tag{6.37}$$

For $l = 0$ the result is $\Delta E^{SO}_{nj0} = 0$ because the matrix element of $J^2 - L^2 - S^2$ vanishes.

The Darwin term. For the last contribution to the relativistic corrections we start with a heuristic argument that is based on the idea of a *Zitterbewegung*, i.e. an uncertainty in the position of an electron that is of the order of its Compton wave length

$$\lambda\!\!\!^{-}_e = \frac{\hbar}{m_e c} \approx 3.86 \times 10^{-13} m. \tag{6.38}$$

The effect of averaging over a slightly smeared out position of an electron in an electrostatic field would amount to an effective potential

$$V_{eff}(\vec{x}) = \Big\langle V(\vec{x}+\delta\vec{x})\Big\rangle_{\delta\vec{x}} = V(\vec{x}) + \Big\langle \delta x^i\Big\rangle \partial_i V(\vec{x}) + \frac{1}{2}\Big\langle \delta x^i \delta x^j\Big\rangle \partial_i\partial_j V(\vec{x}) + \dots \tag{6.39}$$

Imposing rotational symmetry of the fluctuations we expect $\langle \delta x^i\rangle = 0$ and

$$\Big\langle \delta x^i \delta x^j\Big\rangle = \delta_{ij}\Big\langle \delta x^1 \delta x^1\Big\rangle = \frac{1}{3}\delta^{ij}\Big\langle \delta\vec{x}\delta\vec{x}\Big\rangle = \frac{1}{3}\delta^{ij}(\delta|\vec{x}|)^2 = \frac{1}{3}\delta^{ij}(\delta r)^2. \tag{6.40}$$

If we set the expectation value of the fluctuation δr of the position equal to the Compton wavelength $\lambda\!\!\!^{-}_e$ we obtain

$$V_{eff}(\vec{x}) = V(\vec{x}) + \frac{\hbar^2}{6m_e^2c^2}\Delta V(\vec{x}) \tag{6.41}$$

where $\Delta = \delta^{ij}\partial_i\partial_j$ is the Laplace operator. This line of ideas leads to the correct functional form, but the correct prefactor ($\frac{1}{8}$ instead of $\frac{1}{6}$) of the Darwin term

$$H_D = \frac{\hbar^2}{8m_e^2c^2}\Delta V(\vec{x}) = -\frac{\hbar^2 Z e^2}{8m_e^2c^2}\,\Delta\frac{1}{r} \tag{6.42}$$

is obtained from the Dirac equation, as we will see in chapter 7). The Coulomb potential solves the Poisson equation with point-like source,

$$\Delta\frac{1}{r} = -4\pi\delta^3(\vec{x}). \tag{6.43}$$

Because of the δ-function only the s-waves contribute to the **Darwin term**

$$\Delta E^D_{nl} = \langle nlm|H_D|nlm\rangle = \frac{\pi\hbar^2 Z e^2}{2m_e^2c^2}|\psi_{nlm}(0)|^2 = \frac{m_e Z^4 e^8}{2n^3\hbar^4c^2}\delta_{l,0}. \tag{6.44}$$

Note that the Darwin term exactly corresponds to the formal limit $l \to 0$ of the spin-orbit correction (6.37) for the case $j = l + \frac{1}{2}$

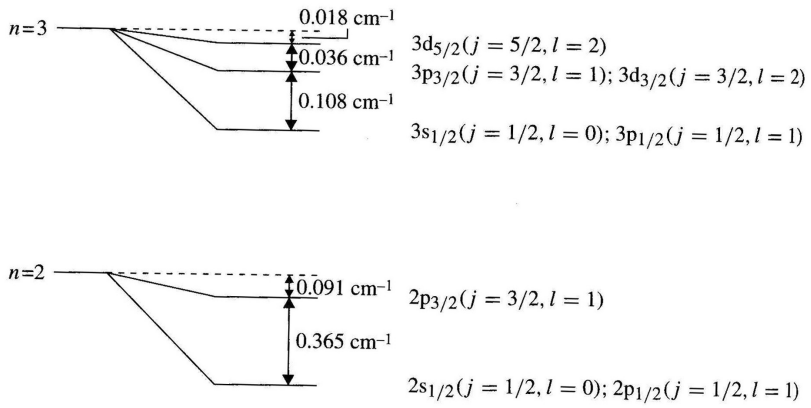


Figure 6.1: Fine structure splitting of the $n = 2$ and $n = 3$ levels of the hydrogen atom.

Fine structure. Putting everyting together we obtain the complete fine structure energy correction

$$\Delta E_{nj}^{FS} = \frac{m_e Z^4 e^8}{2\hbar^4 n^4 c^2} \left(\frac{3}{4} - \frac{n}{j + \frac{1}{2}} \right). \tag{6.45}$$

As one can observe in figure 6.1 for the orbitals with principal quantum numbers $n = 2$ and $n = 3$, the energy shift is always negative because $j + \frac{1}{2} \leq n$. It is independent of the orbital quantum number l and only depends on n and the total angular momentum j, which leads to a degeneracy of the orbitals $2s_{1/2}$ and $2p_{1/2}$ that will be important in our discussion of the linear Stark effect.

6.3 External fields: Zeeman effect and Stark effect

We now analyze the level splittings that are due to external static electromagnetic fields. Such fields reduce the symmetry, at most, to rotations about some axis and hence can lead to further liftings of degeneracies that are otherwise protected by the 3-dimensional rotation symmetry of an isolated atom.

The Zeeman Effect. Taking into account the g-factor of the electron the Hamilton operator for the interaction with an external magnetic field is

$$H_Z = \vec{B}(\vec{\mu}_L + \vec{m}_S) = \frac{e}{2m_e c}(\vec{L} + 2\vec{S})\vec{B}. \tag{6.46}$$

For a constant magnetic field along the z-axis we have

$$\vec{B} = B_z \vec{e}_z \qquad \Rightarrow \qquad H_Z = \frac{e}{2m_e c}(L_z + 2S_z)B_z = \frac{e}{2m_e c}(J_z + S_z)B_z. \tag{6.47}$$

Taking the hydrogen atom without fine structure as the starting point we note that the energy levels are degenerate for fixed n in the orbital quantum number $l < n$ and in the magnetic quantum numbers m_l, m_s. If we want to treat the spin-orbit coupling and the Zeeman Hamiltonian as perturbations, the problem is that $[H_{SO}, H_Z] \neq 0$ so that these operators cannot be diagonalized simultaneously (in the degeneration space of fixed n). We would therefore have to treat both interactions at once, thus diagonalizing much larger matrices. While this can be done (see [Schwabl] section 14.1.3), we rather consider the two limiting situations where one of the effects is dominant and the other is treated as a small perturbation on top of the larger one. Accordingly, there is a weak field and a strong field version of the Zeeman effect, where the latter is associated with the name Paschen-Back effect, as we will discuss below.

Weak field Zeeman effect. For weak external magnetic fields the spin-orbit coupling is dominant. We hence compute the matrix elements of H_Z between eigenstates $|njlsm_j\rangle$ of H_{SO} and need to diagonalize within the degenerate subspace of fixed n, j, l. This will be a good approximation as long as the matrix elements of H_Z between states with different total angular momentum are small compared to the energy denominators (6.17) caused by the fine structure splitting due to H_{SO} so that second order perturbative corrections are small as compared to the leading order. This is the precise meaning of what we call a weak magnetic field.

Since we assume that H_{SO} is dominant the first order energy correction is now computed for fixed J^2, i.e. in the basis $|njlsm_j\rangle$ and within the $2j+1$ dimensional subspace $m_j = -j, \ldots, j$. But in this subspace H_Z is already diagonal because $J_z + S_z$ commutes with J_z. Hence we only need to evaluate

$$\begin{aligned} \Delta E^Z_{m_j} &= \langle njlsm_j| \frac{eB}{2m_e c} (J_z + S_z) |njlsm_j\rangle \\ &= \frac{eB}{2m_e c} \big(\hbar m_j + \langle njlsm_j|S_z|njlsm_j\rangle\big). \end{aligned}$$

In order to evaluate S_z we expand $|njlsm_j\rangle$ in the basis $|lsm_l m_l\rangle$,

$$|njlsm_j\rangle = \sum_{m_s=\pm\frac{1}{2}} |nlsm_l m_s\rangle \underbrace{\langle lsm_l m_s|jlsm_j\rangle}_{\text{Clebsch-Gordan coeff. } C^j_{m_l m_s}}. \tag{6.48}$$

The Clebsch-Gordan coefficients $C^j_{m_l m_s} \equiv C^{lsj}_{m m_l m_s}$ for $\vec{J} = \vec{L} + \vec{S}$ were computed in (5.69),

$C^j_{m_l,m_s}$	$m_s = \frac{1}{2}$	$m_s = -\frac{1}{2}$
$j = l + \frac{1}{2}$	$\sqrt{\frac{l+m_j+1/2}{2l+1}}$	$\sqrt{\frac{l-m_j+1/2}{2l+1}}$
$j = l - \frac{1}{2}$	$-\sqrt{\frac{l-m_j+1/2}{2l+1}}$	$\sqrt{\frac{l+m_j+1/2}{2l+1}}$

(6.49)

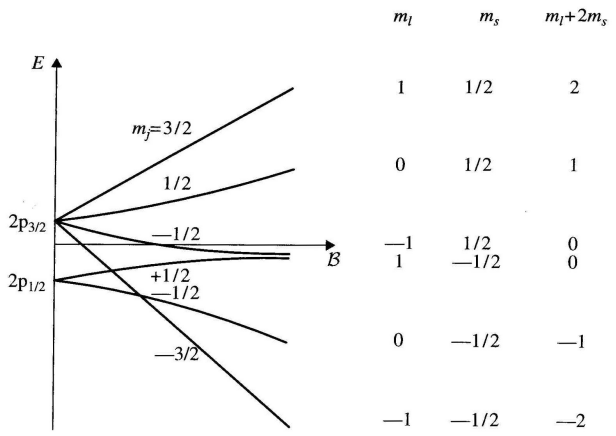


Figure 6.2: Schematic diagram for the splitting of the 2p levels of a hydrogen atom as a function of an external magnetic field B. For small B the degeneracy of the six 2p levels is completely removed, but as B becomes large the levels $j = \frac{3}{2}, m = -\frac{1}{2}$ and $j = \frac{1}{2}, m = \frac{1}{2}$ converge, so that the degeneracy is only partially removed in the limit of very large B (Paschen-Back effect).

with $m_j = m_l + m_s$. For the two cases $j = l \pm \frac{1}{2}$ the matrix elements of S_z thus become

$$\begin{aligned}\langle jlsm_j|S_z|jlsm_j\rangle &= \left(C^{l\pm\frac{1}{2}}_{m_l,+\frac{1}{2}}\left\langle jlsm_l,+\frac{1}{2}\right| + C^{l\pm\frac{1}{2}}_{m_l,-\frac{1}{2}}\left\langle jlsm_l,-\frac{1}{2}\right|\right)\times \\ &\qquad S_z\left(C^{l\pm\frac{1}{2}}_{m_l,+\frac{1}{2}}\left|jlsm_l,+\frac{1}{2}\right\rangle + C^{l\pm\frac{1}{2}}_{m_l,-\frac{1}{2}}\left|jlsm_l,-\frac{1}{2}\right\rangle\right) \qquad (6.50)\\ &= \frac{\hbar}{2}\left(\left|C^{l\pm\frac{1}{2}}_{m_l,+\frac{1}{2}}\right|^2 - \left|C^{l\pm\frac{1}{2}}_{m_l,-\frac{1}{2}}\right|^2\right) = \pm\frac{\hbar}{2}\frac{2m_j}{2l+1}. \qquad (6.51)\end{aligned}$$

The energy shift induced by a weak external magnetic field is therefore

$$\Delta E^Z_{jlm_j} = \frac{e\vec{B}}{2m_e c}\left[\hbar m_j \pm \frac{\hbar m_j}{2l+1}\right] = \frac{e\hbar B}{2m_e c}m_j \cdot \begin{cases}\frac{2l+2}{2l+1} & j = l+1/2\\ \frac{2l}{2l+1} & j = l-1/2\end{cases}. \qquad (6.52)$$

The spin-orbit coupling already removes the degeneracy in j. From equation (6.52) we see that a weak external magnetic field in addition lifts the degeneracy in m_j, thus explaining the name *magnatic quantum number*. A level with given quantum numbers n and j thus splits into $2j+1$ distinct lines. As an example consider the 2p orbitals of the hydrogen atom. The $2p_{3/2}$ level, with $j = l + 1/2$, splits into 4 levels according to $m_j = \frac{3}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{3}{2}$ with $\Delta E^Z = \frac{e\hbar B}{2m_e c}m_j \cdot \frac{4}{3}$. The $2p_{1/2}$ levels split into two with $m_j = \pm 1/2$ and $\Delta E^Z = \frac{e\hbar B}{2m_e c}m_j \cdot \frac{2}{3}$ (see figure 6.2).

Strong field and Paschen–Back effect. In the case of very strong magnetic fields the spin-orbit term becomes (almost) irrelevant and the Zeeman term H_Z forces the electrons into states that are (almost) eigenstates of $L_z + 2S_z = J_z + S_z$. The total angular momentum J^2 is hence no longer conserved, but L^2 and S^2 commute with $\hat{H}_Z$ so that we can use the original basis $|nlm_l\rangle \otimes |sm_s\rangle$ for the calculation. The fact that a strong magnetic field thus breaks up the coupling between spin and orbital angular momentum and makes them individually conserved quantities is called Paschen–Back effect.

The energy shift due to the external magnetic field B is now easily evaluated as

$$\Delta E^Z_{nlsm_lm_s} = \frac{e\hbar B}{2m_e c}\left(m_l + 2m_s\right). \tag{6.53}$$

The magnetic field B does not remove the degeneracy of the hydrogenic energy levels in l and it removes the degeneracy in m_l and m_s only partially. Considering again the 2p level as shown in figure 6.2 we insert $m_l = -1, 0, 1$ and $m_s = \pm 1/2$ into equation (6.53). Due to the g-factor $m_l + 2m_s$ can now assume all 5 integral values between ± 2, but the value $m_l + 2m_s = 0$ can be obtained in two different ways and hence corresponds to a degenerate energy level. In figure 6.2 one observes that the $m_j = -\frac{1}{2}$ line originating from $2p_{3/2}$ and the $m_j = \frac{1}{2}$ line originating from $2p_{1/2}$ converge for large B.

The Stark effect. A hydrogen atom in a uniform *electric field* $\vec{E} = E_z\vec{e}_z$ experiences a shift of the spectral lines that was first observed in 1913 by Stark. The interaction energy of the electron in the external field amounts to an external potential $V_S = -e\vec{E}\vec{x}$ and hence to an interaction Hamiltonian

$$H_S = -e\vec{E}\vec{X} = -eE_z z. \tag{6.54}$$

First of all we shall assume that E is large enough for the fine structure effects to be negligible. We hence work in the basis $|nlm\rangle$ and ignore the spin because the electric field does not couple to the magnetic moment of the electron. The matrix elements of H_S are strongly constrained by symmetry considerations. First we note that z is invariant under rotations about the z-axis so that J_z is conserved and $\langle l'm'|z|lm\rangle$ is proportional to $\delta_{m,m'}$. Moreover, under a parity transformation $\vec{X} \to -\vec{X}$ the interaction term H_S is odd, so that

$$\langle l'm'|z|lm\rangle \quad \overset{\vec{x} \mapsto -\vec{x}}{\longrightarrow} \quad (-1)^{l-l'+1}\langle l'm'|z|lm\rangle. \tag{6.55}$$

Since the integral $\int d^3x|\psi(\vec{x})|^2 z$ is invariant under the change of variables $\vec{x} \mapsto -\vec{x}$ the matrix element can be non-zero only if $l - l'$ is odd. Moreover, one can show that $|l - l'| \leq 1$ for the *electric dipole* matrix element $\langle|\vec{X}|\rangle$, as we will learn in the context of tensor operators.[4] Nonzero matrix elements therefore cannot be diagonal so that a linear Stark effect (i.e. a contribution in first order perturbation theory) can only occur if energy levels are degenerate for different orbital angular momenta. Such degeneracies only occur for excited states of the hydrogen atom (in the ground state one can only observe the *quadratic Stark effect*, i.e. a level splitting in second order perturbation theory.

The simplest situation for which we can hope for a linear Stark effect is for $l = 0, 1$ and $n = 2$, for which there may be a nonzero matrix element between the states $|200\rangle$ and $|210\rangle$,

[4] A vector operator $\vec{X}$ correponds to addition of spin 1. Adding spin one to a state of angular momentum l can only yield angular momentum l' with $|l' - l| \leq 1$ (see chapter 9). The same argument will apply to selection rules of in the dipol approximation for absorption and emission of electromagnetic radiation (see section 6.5).

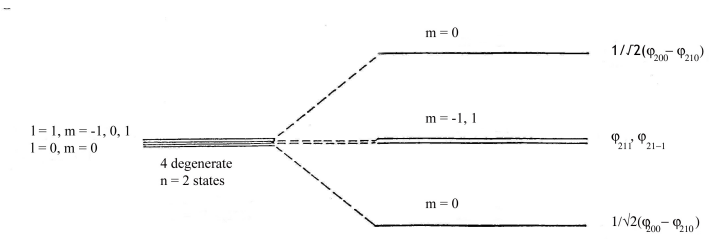


Figure 6.3: Splitting of the degenerate $n = 2$ levels of hydrogen due to the linear Stark effect.

which are degenerate and satisfy the selection rules $l - l' = 1$ and $m = m'$. Evaluation of the matrix element yields

$$\langle 210|z|200\rangle = \langle 200|z|210\rangle = -3eE_z a_0, \tag{6.56}$$

where a_0 is the Bohr radius. Since a matrix of the form $\left(\begin{smallmatrix} 0 & \lambda \\ \lambda & 0 \end{smallmatrix}\right)$ has eigenvalues $\pm\lambda$ the level shifts of the linear Stark effect in the hydrogen atom for $n = 0$, which affect the two states with magnetic quantum number $m = 0$, are

$$\Delta E^S_{n=2,m=0} = \pm 3eE_z a_0. \tag{6.57}$$

as shown in figure 6.3. Recall that the linear Stark effect can only occur if there are degenerate energy levels of different parity, which can only occur for hydrogen.

6.4 The variational method (Riesz)

The variational method is an approximation technique that is not restricted to small perturbations from solvable situations but rather requires some qualitative idea about how the ground state wave function looks like. It is based on the following fact:

Theorem: A wave function $|u\rangle$ is a solution to the stationary Schrödinger equation if and only if the energy functional

$$E(u) = \frac{\langle u|H|u\rangle}{\langle u|u\rangle} \tag{6.58}$$

is stationary, i.e.

$$H|u\rangle = E|u\rangle \qquad \Leftrightarrow \qquad \delta E = 0 \tag{6.59}$$

for arbitrary variations $u \to u+\delta u$, where we do not normalize u in order to have unconstrained variations.

For the *proof* of this theorem we compute the variation of the functional (6.58). Since variations are infinitesimal changes they obey the same rules as differentiation, including the formula $(f/g)' = f'/g - fg'/g^2$, i.e.

$$\delta E = \frac{\delta(\langle u|H|u\rangle)}{\langle u|u\rangle} - \frac{(\langle u|H|u\rangle)\,\delta(\langle u|u\rangle)}{(\langle u|u\rangle)^2} = \frac{\delta(\langle u|H|u\rangle)}{\langle u|u\rangle} - E\frac{\delta(\langle u|u\rangle)}{\langle u|u\rangle}. \tag{6.60}$$

Using the product rule $\delta\langle u|u\rangle = \langle\delta u|u\rangle + \langle u|\delta u\rangle$ stationarity of the energy implies

$$0 = ||u||^2 \cdot \delta E = \langle\delta u|H|u\rangle - E\langle\delta u|u\rangle + \langle u|H|\delta u\rangle - E\langle u|\delta u\rangle. \tag{6.61}$$

If the variations of u and u^* can be done independently then the first two terms (and the last two terms) on the r.h.s. have to cancel one another, i.e. $\langle\delta u|H|u\rangle - E\langle\delta u|u\rangle = 0$, for arbitrary variations $\langle\delta u|$ of $u^*(\vec{x})$, which is equivalent to the Schrödinger equation. To see that this is indeed the case we replace u by $v = iu$ in (6.61) so that $\delta v = i\delta u$ and $\delta v^* = -i\delta u^*$, implying

$$0 = -i\big(\langle\delta u|H|u\rangle - E\langle\delta u|u\rangle\big) + i\big(\langle u|H|\delta u\rangle - E\langle u|\delta u\rangle\big). \tag{6.62}$$

Adding i times (6.62) to (6.61) we find

$$\delta E = 0 \qquad \Rightarrow \qquad \langle\delta u|(H-E)|u\rangle = 0 \quad \forall \quad \langle\delta u|, \tag{6.63}$$

which implies the Schrödinger equation. This completes the proof since, in turn, $(H-E)|u\rangle = 0$ implies the vanishing of the variation (6.61).

If we expand $|u\rangle$ in a basis of states $|u\rangle = \sum_n c_n|e_n\rangle$ then our theorem tells us that the Schrödinger equation is equivalent to the equations $\frac{\partial E}{\partial c_n} = 0$. But for an infinite–dimensional Hilbert space we would have to solve infinitely many equations.

The variational method thus proceeds by introducing a family of trial wave functions

$$u(\alpha_1, \alpha_2, \ldots, \alpha_n) \tag{6.64}$$

parametrized by a finite number of variables α_i and extremizes the energy functional within the subset of Hilbert space functions that are of the form (6.64) for some values of the parameters α_i, i.e. we solve

$$\frac{\partial E(u(\alpha_1,\ldots))}{\partial\alpha_1} = \ldots = \frac{\partial E(u(\alpha_n,\ldots))}{\partial\alpha_n} = 0. \tag{6.65}$$

If the correct wave function $|u_0\rangle$ for the ground state happens to be contained in the family (6.64) of trial functions then the solution to the stationarity equations with the smallest value of $E(u)$ provides us with the exact solution to the Schrödinger equation. If we have, on the other hand, a badly chosen family that does not anywhere come close to $|u_0\rangle$ then our approximation to the ground state energy may be arbitrarily bad. Nevertheless, for an orthonormal energy eigenbasis $|e_n\rangle$

$$\langle u|H|u\rangle = \sum |c_n|^2\langle e_n|H|e_n\rangle = \sum |c_n|^2 E_n \geq ||u||^2 E_{min} \quad \Rightarrow \quad E \geq E_{min} \tag{6.66}$$

so that we will always find an rigorous upper bound for the ground state energy.

6.4.1 Ground state energy of the helium atom

We now apply the variational method to improve a perturbative computation of the ground state energy of the helium atom, which is a system consisting of a nucleus with charge $Ze = 2e$ and two electrons. Treating the nucleus as infinitely heavy and neglecting relativistic effects like the spin-orbit interaction we consider the Hamiltonian

$$H = -\frac{\hbar^2}{2m_e}(\Delta_1 + \Delta_2) - \frac{2e^2}{r_1} - \frac{2e^2}{r_2} + \frac{e^2}{r_{12}}, \tag{6.67}$$

which consists of the kinetic energies $T_i = -\frac{\hbar^2}{2m_w}\Delta_i$, the Coulomb energies $V_i = -2e^2/|\vec{x}_i|$ due to the attraction by the nucleus and the mutual repulsion $V_{12} = e^2/|\vec{x}_1 - \vec{x}_2|$ of the electrons. If we omit the repulsive interaction among the electrons in a first step, the Hamiltonian becomes the sum of two commuting operators

$$H_0 \;=\; -\frac{\hbar^2}{2m_e}(\Delta_1 + \Delta_2) - \frac{2e^2}{r_1} - \frac{2e^2}{r_2} = (T_1 + V_1) + (T_2 + V_2) \tag{6.68}$$

for two independent particles and the Schrödinger equation $H_0|u\rangle = E_0|u\rangle$ is solved by product wave functions

$$u(\vec{x}_1, \vec{x}_2) = u_{n_1 l_1 m_1}(\vec{x}_1)\, u_{n_2 l_2 m_2}(\vec{x}_2). \tag{6.69}$$

The energy thus becomes the sum of the two respective energy eigenvalues,

$$E_0 = -Z^2\mathcal{R}\left(\frac{1}{n_1^2} + \frac{1}{n_2^2}\right), \tag{6.70}$$

where $\mathcal{R} = \hbar^2/2m_e a_0^2 = m_e e^4/2\hbar^2 = 13.6\,eV$ is the Rydberg constant. For the approximate ground state wave function $u_0^{pert}(\vec{x}_1, \vec{x}_2) = u_{100}(\vec{x}_1)u_{100}(\vec{x}_2)$ this implies $E_0^{pert} \approx -108.8\,eV$ where the superscript refers to the Rayleigh–Schrödinger perturbation theory with

$$H = H_0 + V, \qquad V = V_{12} = \frac{e^2}{r_{12}}. \tag{6.71}$$

For the wave function we have ignored so far the spin degree of freedom and the Pauli exclusion principle. In chapter 10 (many particle theory) we will learn that wave functions of identical spin 1/2 particles have to be anti-symmetrized under the simultaneous exchange of all of their quantum numbers (position and spin), which is the mathematical implementation of Pauli's exclusion principle. For the ground state of the helium atom the wave function is symmetric under the exchange $\vec{x}_1 \leftrightarrow \vec{x}_2$ and total antisymmetry implies antisymmetrization of the spin degrees of freedom so that

$$u_0(\vec{x}_1, \vec{x}_2, s_{1z}, s_{2z}) = u_{100}(\vec{x}_1)\; u_{100}(\vec{x}_2)\; |0,0\rangle_{12}, \tag{6.72}$$

where $|0,0\rangle_{12}$ is the singlet state in spin space. Since this will not influence any of our results we will, however, ignore the spin degrees of freedom for the rest of the calculation.

Taking into account now the repulsion term V between the electrons in (6.71) as a perturbation, the first order ground state energy correction becomes

$$E_1 = \langle u_0|V_{12}|u_0\rangle = e^2 \int d^3x_1 d^3x_2 \frac{|u_{100}(\vec{x}_1)|^2\ |u_{100}(\vec{x}_2)|^2}{|\vec{x}_1 - \vec{x}_2|} \tag{6.73}$$

where

$$u_{100}(\vec{x}) = \frac{1}{\sqrt{\pi}} \left(\frac{Z}{a_0}\right)^{\frac{3}{2}} e^{-\frac{Z}{a_0} r} \tag{6.74}$$

is the wave function of a single electron in the Coulomb field of a nucleus with atomic number Z and a_0 is the Bohr radius.

The integrals for the energy correction (6.73) are best carried out in spherical coordinates,

$$E_1 = \left(\frac{Z^3}{a_0^3\pi}\right)^2 \int_0^\infty dr_1 r_1^2 e^{-\frac{2Z}{a_0} r_1} \int_0^\infty dr_2 r_2^2 e^{-\frac{2Z}{a_0} r_2} \int d\Omega_1 d\Omega_2 \frac{e^2}{|\vec{x}_1 - \vec{x}_2|}. \tag{6.75}$$

If we first perform the angular integration $d\Omega_1$ it is useful to recall that a spherically symmetric charge distribution[5] at radius r_1 creates a constant (force-free) potential $-q/r_1$ in the interior $r < r_1$ and a Coulomb potential $-q/r$ of a point charge located at the origin for $r > r_1$. Performing the Ω_1–integration we thus obtain

$$\begin{aligned} E_1 &= 4\pi \left(\tfrac{Z^3 e}{a_0^3\pi}\right)^2 \int_0^\infty dr_2\, r_2^2 \int_0^\infty dr_1\, r_1^2 \int d\Omega_2\ e^{-\frac{2Z}{a_0} r_1}\ e^{-\frac{2Z}{a_0} r_2} \cdot \begin{cases} \frac{1}{r_2} & r_2 > r_1 \\ \frac{1}{r_1} & r_2 < r_1 \end{cases} && (6.76) \\ &= 2\,(4\pi)^2 \left(\tfrac{Z^3 e}{a_0^3\pi}\right)^2 \int_0^\infty dr_1\, r_1^2 \int_{r_1}^\infty dr_2\, r_2\ e^{-\frac{2Z}{a_0} r_1}\ e^{-\frac{2Z}{a_0} r_2}, && (6.77) \end{aligned}$$

where the contribution of the domain $r_2 < r_1$ is accounted for by the prefactor 2 in the second line and the trivial Ω_2-integration has also been done. With $\int r e^{-cr} = -\frac{1+cr}{c^2} e^{-cr}$ and $c = \frac{2Z}{a_0}$ we find

$$E_1 = 32 \left(\tfrac{Z}{a_0}\right)^6 e^2 \int_0^\infty dr_1\, r_1^2 \left(\left(\tfrac{a_0}{2Z}\right)^2 + r_1 \tfrac{a_0}{2Z}\right) e^{-\frac{4Z}{a_0} r_1} \tag{6.78}$$

and with $\int_0^\infty dr\, r^{n-1} e^{-cr} = \Gamma(n)/c^n = (n-1)!/c^n$ the energy correction becomes

$$E_1 = 32\, \frac{Ze^2}{a_0} \left(\frac{2!}{2^2 4^3} + \frac{3!}{24^4}\right) = \frac{5}{8} \frac{Ze^2}{a_0} \approx 34.015\ \text{eV}\,. \tag{6.79}$$

With $E_0 = 8\mathcal{R}$ and $\mathcal{R} = 13.606\,\text{eV}$ our perturbative result for the ground state energy of the helium atom becomes

$$E_{He}^{(pert)} = E_0 + E_1 + \ldots \approx -74.83\,\text{eV} \tag{6.80}$$

This is about 5% higher than the experimental value

$$E_{He}^{(exp)} \approx -79.015\,\text{eV}, \tag{6.81}$$

which is not too bad for our simple approach, in particular if we note that the first perturbative correction E_1, with almost 1/3 of E_0, is quite large.

[5] Since all solutions to the homogeneous Laplace equation are superpositions of $r^l Y_{lm}$ and $r^{-l-1} Y_{lm}$ spherical symmetry implies $l = 0$ so that the potential is constant in the interior $r < r_{charge}$, as there is no singularity at the origin, and proportional to $1/r$ for $r > r_{charge}$, as the potential has to vanish for $r \to \infty$.

6.4.2 Applying the variational method and the virial theorem

In order to improve our perturbative result we note that the second electron partially screens the positive charge of the nucleus so that the electrons on average feel the attraction of an effective charge $q_{eff} < Ze$ and are less tightly bound. This suggests to use the ground state wave function of a hydrogen-like atom with the atomic number Z of the nucleus replaced by a continuous parameter b. Our starting point is thus the family $u(b)$ of normalized trial wave functions

$$u(\vec{x}_1, \vec{x}_2; b) = \frac{b^3}{\pi a_0^3} e^{-\frac{b}{a_0}(r_1+r_2)}, \tag{6.82}$$

where we recover the case (6.72) for $b = Z$ and expect to find $b < Z$ at the minimum of $E(b)$. The expectation value

$$\langle u(b)|V_{12}|u(b)\rangle = \frac{5}{8}\frac{be^2}{a_0} \tag{6.83}$$

directly follows from (6.79) by replacing Z by b. But in order to find the expectation value of $H_0 = T_1 + V_1 + T_2 + V_2$ we need to decompose the ground state energy $T + V$ into the kinetic contribution, for which we simply can replace Z by b, and the potential contribution, which is proportional to the charge Ze of the nucleus. The decomposition can be obtained as follows.

The virial theorem: If the potential $V(\vec{x})$ of a Hamiltonian of the form

$$H = T + V, \qquad T = \tfrac{P^2}{2m} \tag{6.84}$$

is homogeneous of degree n, i.e.

$$V(\lambda\vec{x}) = \lambda^n V(\vec{x}) \qquad \forall\ \lambda \in \mathbb{R}, \tag{6.85}$$

then the expectation values of T and V are related by[6]

$$2\,\langle u|T|u\rangle \;=\; n\,\langle u|V|u\rangle \tag{6.89}$$

so that $\langle u|T|u\rangle = \frac{n}{n+2}E$ and $\langle u|V|u\rangle = \frac{2}{n+2}E$ for every bound state $|u\rangle$.

[6] The proof of the quantum mechanical virial theorem is based on the Euler formula

$$\sum_i x^i \partial_i V(\vec{x}) = nV(\vec{x}) \tag{6.86}$$

for a homogeneous potential of degree n and on the fact that the expectation value of a commutator $[H, A]$ vanishes for bound states $|u_i\rangle$,

$$\langle u_i|\,[H, A]\,|u_i\rangle = \langle u_i|HA - AH|u_i\rangle = \langle u_i|E_iA - AE_i|u_i\rangle = 0. \tag{6.87}$$

The theorem then follows for $A = \vec{X}\vec{P}$ because

$$[\vec{X}\vec{P}, \tfrac{P^2}{2m}] = 2i\hbar\tfrac{P^2}{2m}, \qquad [\vec{X}\vec{P}, V] = \tfrac{\hbar}{i}X^i\partial_i V \qquad \Rightarrow \qquad [\vec{X}\vec{P}, H] = i\hbar(2T - nV). \tag{6.88}$$

For further details see, for example, chapter 4 of [Grau].

Since the Coulomb potential is homogeneous of degree $n = -1$ we find

$$T_i = -\tfrac{1}{2}V_i = Z^2\mathcal{R} = \tfrac{Z^2e^2}{2a_0} \quad \mapsto \quad \langle u(b)|T_i|u(b)\rangle = \tfrac{b^2e^2}{2a_0}, \;\; \langle u(b)|V_i|u(b)\rangle = -\tfrac{bZe^2}{a_0}. \tag{6.90}$$

The energy functional $E(b) = \langle u(b)| \left(T_1 + V_1 + T_2 + V_2 + V_{12}\right) |u(b)\rangle$ thus becomes

$$E(b) = \tfrac{e^2}{a_0}\left(b^2 - 2bZ + \tfrac{5}{8}b\right) = \tfrac{e^2}{a_0}\left((b - Z + \tfrac{5}{16})^2 - (Z - \tfrac{5}{16})^2\right). \tag{6.91}$$

The minimal value $E_{min} = \tfrac{e^2}{a_0}(Z - \tfrac{5}{16})^2$ is obtained for $b = Z - \tfrac{5}{16}$. For the helium atom we thus obtain

$$E_{He}^{(var)} = -\tfrac{e^2}{a_0}\left(\tfrac{27}{16}\right)^2 \approx 77.5\,\mathrm{eV}, \tag{6.92}$$

which is only 2% above the experimental value (6.81). The effecitve charge becomes $b \approx \tfrac{27}{16}$.

6.5 Time dependent perturbation theory

We now turn to non-stationary situations. In particular we will be interested in the response of a system to time dependent perturbations

$$H(t) = H_0 + W(t), \tag{6.93}$$

where the unperturbed Hamiltonian H_0 is not explicitly time dependent. For simplicity we assume that the unpertubed system has discrete and non-degenerate eigenstates

$$H_0|\varphi_n\rangle = E_n|\varphi_n\rangle. \tag{6.94}$$

If the perturbation is turned on at an initial time t_0 this implies

$$i\hbar\frac{\partial}{\partial t}|\varphi(t)\rangle = H_0|\varphi(t)\rangle \qquad \text{for} \qquad t < t_0 \tag{6.95}$$

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \left(H_0 + W(t)\right)|\psi(t)\rangle \qquad \text{for} \qquad t > t_0 \tag{6.96}$$

where the state $|\psi_i(t)\rangle$ is defined by the initial condition

$$|\psi_i(t = t_0)\rangle = |\varphi_i(t_0)\rangle \tag{6.97}$$

if the system is originally in the stationary state $|\varphi_i\rangle$. We first consider two limiting situations:

- In the **sudden approximation** we assume that a time independent perturbation is switched on very rapidly,

$$t_{switch} \ll t_{response} \qquad \Rightarrow \qquad W(t) \simeq \theta(t - t_0)W' \tag{6.98}$$

so that we can describe the time dependence by a step function $\theta(t-t_0)$. A physical example would be a radioactive decay, where the reorganization of the electron shell takes much longer than the nuclear reaction. Hence the Hamiltonian suddenly changes to a new time-independent form $H' = H_0 + W'$. For $t > t_0$ the system has a new set of stationary solutions $|\psi_f\rangle$, and since the wave function has no time to evolve under a time-dependent force the transition probability into a final state

$$\mathcal{P}_{i\to f} = |\langle\psi_f|\varphi_i\rangle|^2 \tag{6.99}$$

is determined by the operlap (scalar product) of the wave functions.

- The **adiabatic limit** is the other extremal situation,

$$t_{switch} \gg t_{response}, \tag{6.100}$$

for which the time variation of the external conditions is so slow that it cannot induce a transition and the system evolves by a continuous deformation of the energy eigenstate because we have, at each time, an almost stationary situation. More quantitatively, the transition probability will be negligable if the energy uncertainty that is due to the time variation of H is small in comparison to differences between energy levels.

In the rest of this section we will consider small time-dependent perturbations $W(t) = \lambda V(t)$, where a small parameter λ can be introduced to control the perturbative expansion, but it is equivalent to simply count powers of W. Since the perturbation is small we can, at each instant of time at which we perform a measurement, use eigenstates $|\varphi_f\rangle$ of H_0 to represent the possible outcomes of the reduction of the wave function. Our aim hence is to determine the probability

$$\mathcal{P}_{i\to f} = |\langle\varphi_f|\psi_i(t)\rangle|^2 \tag{6.101}$$

for finding the system in a final eigenstate $|\varphi_f\rangle$ after having evolved from $|\varphi_i\rangle$ under the influence of $H = H_0 + W$ according to (6.96) with boundary condition (6.97). For simplicity we set $t_0 = 0$. It is convenient to perform the perturbative computation of (6.101) in the interaction picture

$$|\psi(t)\rangle_I = e^{\frac{i}{\hbar}H_0 t}\,|\psi(t)\rangle = U_0^\dagger(t)|\psi(t)\rangle, \qquad U_0(t) = e^{-\frac{i}{\hbar}H_0 t}, \tag{6.102}$$

so that

$$i\hbar\,\partial_t|\psi(t)\rangle_I = W_I(t)|\psi(t)\rangle_I \qquad \text{with} \qquad W_I(t) = e^{\frac{i}{\hbar}H_0 t}W(t)e^{-\frac{i}{\hbar}H_0 t} \tag{6.103}$$

as we found in (3.146–3.152).

Our next step is to transform the Schrödinger equation (6.103) of the interaction picture into an integral equation by integrating it over the interval from $t_0 = 0$ to t,

$$|\psi_i(t)\rangle_I = |\varphi_i\rangle + \frac{1}{i\hbar}\int_0^t dt'\; W_I(t')|\psi_i(t')\rangle_I, \tag{6.104}$$

where we used the boundary condition (6.97). For small $W_I(t)$ we can solve this equation by iteration, i.e. we insert $|\psi_i\rangle_I = |\varphi_i\rangle + \mathcal{O}(W)$ on the r.h.s. and continue by inserting the resulting higher order corrections of $|\psi_i\rangle_I$. We thus obtain the **Neumann series**[7]

$$\begin{aligned} |\psi_i(t)\rangle = \quad & \underbrace{|\varphi_i\rangle}_{\text{initial state}} + \underbrace{\frac{1}{i\hbar}\int_0^t dt'\; W_I(t')|\varphi_i\rangle}_{\text{first order correction}} + \\ & + \underbrace{\frac{1}{(i\hbar)^2}\int_0^t dt' \int_0^{t'} dt''\; W_I(t')W_I(t'')|\varphi_i\rangle}_{\text{second order correction}} + \ldots \end{aligned} \tag{6.107}$$

The transition amplitude now becomes

$$\mathcal{A}_{i\to f} = \langle\varphi_f|\psi_i(t)\rangle = \delta_{if} - \frac{i}{\hbar}\int_0^t dt'\; \langle\varphi_f|\; W_I(t')\; |\varphi_i\rangle + \mathcal{O}(W_I^2). \tag{6.108}$$

In the remainder of this section we focus on the leading contribution to the transition from an initial state $|\varphi_i\rangle$ to a final state $|\varphi_f\rangle$ with $f \neq i$,

$$\begin{aligned} \mathcal{A}^{(1)}_{i\to f} &= -\frac{i}{\hbar}\int_0^t dt'\; \langle\varphi_f|\; e^{\frac{i}{\hbar}H_0t'}W(t')e^{-\frac{i}{\hbar}H_0t'}\; |\varphi_i\rangle & (6.109)\\ &= -\frac{i}{\hbar}\int_0^t dt'\; e^{\frac{i}{\hbar}(E_f-E_i)t'}\langle\varphi_f|\, W(t')\,|\varphi_i\rangle & (6.110) \end{aligned}$$

where we used $H_0|\varphi_i\rangle = E_i|\varphi_i\rangle$ and $\langle\varphi_f|H_0 = \langle\varphi_f|E_f$ to evaluate the time evolution operators. For **first order transitions** we thus obtain the probability

$$\mathcal{P}^{(1)}_{i\to f} = \frac{1}{\hbar^2}\left|\int_0^t dt'\; e^{i\omega_{fi}t'}\langle\varphi_f|W(t')|\varphi_i\rangle\right|^2 \tag{6.111}$$

with the **Bohr angular frequency**

$$\omega_{fi} = \frac{E_f - E_i}{\hbar}. \tag{6.112}$$

[7] Introducing the **time ordering** operator T by

$$TA(t_1)B(t_2) = \theta(t_1-t_2)A(t_1)B(t_2) + \theta(t_2-t_1)B(t_2)A(t_1) = \begin{cases} A(t_1)B(t_2) & \text{if } t_1 > t_2 \\ B(t_2)A(t_1) & \text{if } t_1 < t_2 \end{cases} \tag{6.105}$$

the Neumann series can be subsumed in terms of a formal expression for the time evolution operator

$$|\psi_i(t)\rangle_I = U_I(t)|\varphi_i\rangle, \qquad U_I(t) = Te^{-\frac{i}{\hbar}\int_0^t dt' W_I(t')} \tag{6.106}$$

as is easily checked by expansion of the exponential.

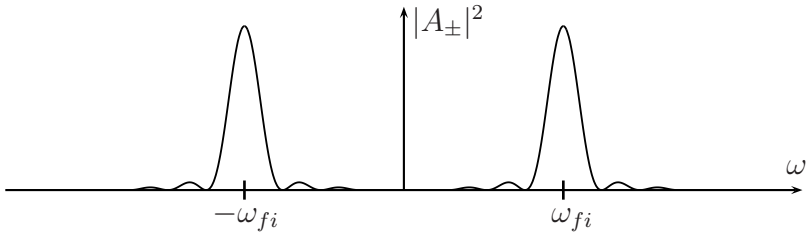


Figure 6.4: The functions $|A_\pm|^2 = \frac{\sin^2(\Delta\omega t/2)}{(\Delta\omega/2)^2} \to 2\pi t\delta(\Delta\omega)$ of height $t^2/2$ and width $4\pi/t$.

Note that the transition probability (6.111) is related to the Fourier transform at ω_{fi} of the matrix element $\langle\varphi_f|W(t')|\varphi_i\rangle$ restricted to $0 < t' < t$.

Periodic perturbations. In practice we will often be interested in the response to periodic external forces of the form

$$W(t) = \theta(t)(W_+ e^{i\omega t} + W_- e^{-i\omega t}) \qquad \text{with} \qquad W_-^\dagger = W_+. \tag{6.113}$$

Then the time integration can be performed with the result

$$\mathcal{P}^{(1)}_{i\to f} = \frac{1}{\hbar^2}\Big| A_+ \langle\varphi_f|W_+|\varphi_i\rangle + A_- \langle\varphi_f|W_+|\varphi_i\rangle \Big|^2 \tag{6.114}$$

in terms of the integrals

$$A_\pm = \int_0^t dt'\ e^{i(\omega_{fi}\pm\omega)t'} = \frac{e^{i(\omega_{fi}\pm\omega)t} - 1}{i(\omega_{fi}\pm\omega)} = e^{\frac{i}{2}(\omega_{fi}\pm\omega)t}\ \frac{\sin\big((\omega_{fi}\pm\omega)t/2\big)}{(\omega_{fi}\pm\omega)/2}. \tag{6.115}$$

Figure 6.4 shows that the functions $A_\pm(\omega)$ are well-localized about $\omega = \mp\omega_{fi}$, respectively, and converge to δ-functions

$$|A_\pm|^2 = \Big(\tfrac{\sin\Delta\omega t/2}{\Delta\omega/2}\Big)^2 \quad\to\quad 2\pi t\delta(\Delta\omega) \qquad \text{with} \qquad \Delta = \omega \pm \omega_{fi} \tag{6.116}$$

for late times $t \gg 1/\omega_{fi}$, where the prefactor follows from the integral

$$\textstyle\int_{-\infty}^{\infty} d\xi \left(\frac{\sin(t\xi)}{\xi}\right)^2 = \pi t \qquad \Rightarrow \qquad \lim_{t\to\infty} \frac{1}{t}\left(\frac{\sin(t\xi)}{\xi}\right)^2 = \pi\delta(\xi). \tag{6.117}$$

For $t \to \infty$ the interference terms between A_+W_+ and A_-W_- in (6.114) can hence be negleted,

$$\mathcal{P}_{i\to f} \ \to\ \frac{2\pi t}{\hbar}\Big(\delta(E_f - E_i - \hbar\omega)\big|\langle f|W_-|i\rangle\big|^2 + \delta(E_f - E_i + \hbar\omega)\big|\langle f|W_+|i\rangle\big|^2\Big). \tag{6.118}$$

For frequencies $\omega \approx \pm\omega_{fi}$ the transition probabilities become very large so that the contribution of A_-W_- is called *resonant term* (absorption of an energy quantum $\hbar\omega_{fi}$) while A_+W_+ is called *anti-resonant* (emission of an energy quantum $\hbar\omega_{fi}$). Since the probability becomes linear in t it is useful to introduce the **transition rate**

$$\Gamma_{i\to f} = \lim_{t\to\infty}(\tfrac{1}{t}\mathcal{P}_{i\to f}). \tag{6.119}$$

For discrete energy levels we thus obtain **Fermi's golden rule**

$$\Gamma_{i\to f} = \frac{2\pi}{\hbar} \left|\langle f|W_\pm|i\rangle\right|^2 \delta(E_f - E_i \pm \hbar\omega) \tag{6.120}$$

which was derived by Pauli in 1928, and called *golden rule* by Fermi in his 1950 book *Nuclear Physics*. The δ-function infinity for discrete energy levels is of course an unphysical artefact of our approximation. In reality spectral lines have finite width. For transitions to a continuum of energy levels we introduce the concept of a level density $\rho(E)$ by summing over transitions to a set F of final states,

$$\Gamma_{i\to F} = \sum_{f\in F} \Gamma_{i\to f} \;\to\; \int_{f\in F} df\; \Gamma_{i\to f} = \int_F dE\; \rho(E)\, \Gamma_{i\to f}. \tag{6.121}$$

Inserting this into the golden rule we can perform the energy integration and obtain the integrated rate

$$\Gamma_i(\omega) = \rho(E_f)\frac{2\pi}{\hbar} \left|\langle f|W_\pm|i\rangle\right|^2, \qquad E_f = E_i \pm \hbar\omega. \tag{6.122}$$

If $f \in F$ is characterized by additional continuous quantum numbers β, like for example the solid angle covered by a detector, the level density can be generalized to $df(E,\beta) = \rho(E,\beta)\, dE\, d\beta$ and the integrated transition rate is obtained by integrating over the relevant range of β's.

6.5.1 Absorption and emission of electromagnetic radiation

We now want to compute the rate for atomic transitions of electrons irradiated by an electromagnetic wave. The relevant Hamiltonian can be written as

$$H = \frac{1}{2m_e}\left(P^2 - \frac{e}{c}\vec{A}\right)^2 + V(r) - \frac{e\hbar}{2m_e c}\vec{\sigma}\vec{B} \tag{6.123}$$

with

$\vec{A}$	...	vector potential of the electromagnetic radiation,
$V(\vec{r})$	...	central potential created by the nucleus,
$\frac{e\hbar}{2m_e c}\vec{\sigma}\vec{B}$	...	magnetic interaction with the radiation field.

We split the Hamiltonian as

$$H = H_0 + W(t) \qquad \text{with} \qquad H_0 = \frac{P^2}{2m_e} + V(r) \tag{6.124}$$

and interaction term

$$W(t) = \underbrace{-\frac{e}{2m_e c}(\vec{P}\vec{A} + \vec{A}\vec{P})}_{W_A} \underbrace{-\frac{e\hbar}{2m_e c}\vec{\sigma}\vec{B}}_{W_B} + \frac{e^2}{2m_e c^2}\vec{A}^2. \tag{6.125}$$

The last term can be ignored because it is quadratic in the perturbation $\vec{A}$. For the vector potential of the electromagnetic field we take a plane wave

$$\vec{A}(\vec{x},t) = \vec{\epsilon}\, A_0 \left(e^{-i(\omega t - \vec{k}\vec{x})} + e^{i(\omega t - \vec{k}\vec{x})} \right) \tag{6.126}$$

with frequency ω and polarization vector $\vec{\varepsilon}$. Since $\vec{B} = \mathrm{curl}\,\vec{A} = i\vec{k} \times \vec{A}$ the magnetic interaction term W_B can be neglected for optical light, for which $\hbar k \ll |\langle f|\vec{P}|i\rangle| \sim \hbar/a_0$. In the radiation gauge

$$\mathrm{div}\vec{A} = \tfrac{i}{\hbar}\vec{P}\vec{A} = 0 \qquad \Leftrightarrow \qquad \vec{\varepsilon}\vec{k} = 0, \tag{6.127}$$

hence $\vec{P}\vec{A} + \vec{A}\vec{P} = 2\vec{A}\vec{P}$ and the relevant matrix element becomes (up to a phase)

$$\langle f|W_A|i\rangle = -A_0 \frac{e}{m_e c}\, \vec{\varepsilon}\langle f|\vec{P}e^{\pm i\vec{k}\vec{x}}|i\rangle. \tag{6.128}$$

For optical transitions expectation values of $\vec{x}$ are of the order of $a_0 \ll 1/k$ so that we can drop contributions from the exponential

$$\langle f|\vec{P}e^{\pm i\vec{k}\vec{x}}|i\rangle \approx \langle f|\vec{P}|i\rangle. \tag{6.129}$$

This is called **electric dipole approximation** because the matrix element of $\vec{P}$ can be related to the matrix element of the dipole

$$\vec{D}_{fi} = \langle f|\vec{X}|i\rangle \tag{6.130}$$

using

$$[H_0, \vec{X}] = \tfrac{\hbar}{im_e}\vec{P} \quad \Rightarrow \quad \langle f|\vec{P}|i\rangle = i\tfrac{m_e}{\hbar}\langle f|\,[H_0,\vec{X}]\,|i\rangle = i\tfrac{m_e}{\hbar}(E_f - E_i)\langle f|\vec{X}|i\rangle. \tag{6.131}$$

Inserting everything into Fermi's golden rule we obtain the transition rate

$$\Gamma_{i\to f} = \frac{2\pi}{\hbar}\left|\frac{A_0 e}{m_e c}\frac{m_e}{\hbar}(E_f - E_i)\,\vec{\varepsilon}\langle f|\vec{X}|i\rangle\right|^2 \delta(E_f - E_i \pm \hbar\omega) \tag{6.132}$$

$$= \frac{2\pi e^2\omega^2}{\hbar\, c^2}\delta(E_f - E_i \pm \hbar\omega)A_0^2\left|\vec{\varepsilon}\langle f|\vec{X}|i\rangle\right|^2. \tag{6.133}$$

The **selection rules** for the dipole approximation are obtained by considering the matrix elements

$$\langle n'l'm'|\vec{X}|nlm\rangle. \tag{6.134}$$

Under a parity transformation $\vec{X} \to -\vec{X}$ the spherical harmonics transform as

$$Y_{lm}(\pi - \theta, \varphi + \pi) = (-1)^l Y_{lm}(\theta,\varphi). \tag{6.135}$$

The matrix element (6.134) hence transforms with a factor $(-1)^{l-l'+1}$ so that spherical symmetry implies that $l - l'$ must be odd. Since

$$[L_z, X_3] = 0, \quad [L_z, X_i \pm iX_2] = \pm\hbar(X_1 \pm iX_2) \tag{6.136}$$

the components X_3 and $X_\pm = X_1 \pm iX_2$ of $\vec{X}$ change the magnetic quantum number by $m' - m \in \{0, \pm 1\}$. More generally, we will learn in the chapter 9 that the vector operator $\vec{X}$ corresponds to addition of angular momentum 1, so that $|l' - l| \leq 1$. Combining all constraints we find

$$l' - l = 1, -1, \qquad m'_l - m_l = 1, 0, -1. \tag{6.137}$$

Moreover, since we neglected magnetic interactions, spin in conserved $m'_s = m_s$. These selection rules translate to

$$j' - j = 1, 0, -1, \qquad j = 0 \Rightarrow j' = 1 \tag{6.138}$$

in the total angular momentum basis $|jlsm_j\rangle$.

In the present chapter we could only compute induced absorption and emission. Spontaneous emission will be discussed in chapter 10.

Chapter 7

Relativistic Quantum Mechanics

In the previous chapters we have investigated the Schrödinger equation, which is based on the non-relativistic energy-momentum relation. We now want to reconcile the principles of quantum mechanics with special relativity. Schrödinger actually first considered a relativistic equation for de Broglie's matter waves, but was deterred by discovering some apparently unphysical properties like the existence of plane waves with unbounded negative energies $E = -\sqrt{p^2c^2 + m^2c^4}$: Classically a particle on the positive branch of the square root will keep $E \geq mc^2$ forever but quantum mechanically interactions can induce a jump to the negative branch releasing $\delta E \geq 2mc^2$. Schrödinger hence arrived at his famous equation in the non-relativistic context.

Two years later Paul A.M. Dirac found a linearization of the relativistic energy–momentum relation, which explained the gyromagnetic ratio $g = 2$ of the electron as well as the fine structure of hydrogen. While his equation missed its original task of eliminating negative energy solutions, it was too successful to be wrong so that Dirac went on, inspired by Pauli's exclusion principle, to solve the problem of unbounded negative energies by inventing the particle-hole duality that is nowadays familiar from semiconductors. In the relativistic context the holes are called anti-particles. While Dirac first tried to identify the anti-particle of the electron with the proton (the only other particle known at that time) this did not work for several reasons and he concluded that there must exist a positively charged particle with the same mass as the electron. The positron was then discovered in cosmic rays in 1932. Dirac thus made the first prediction of a new particle on theoretical grounds and, more generally, showed the existence of antimatter as an implication of the consistency of quantum mechanics with special relativity. This initiated the long development of the modern quantum field theory of elementary particles and interactions.

After a brief discussion of the problem with negative energy solutions we now construct the Dirac equation and analyze its nonrelativistic limit. The issue of Lorentz transformations and

further symmetries will be taken up in the chapter on symmetries and transformation groups.

The Klein-Gordon-equation. If we start with the relativistic energy-momentum relation

$$E^2 = m^2c^4 + c^2\vec{p}^2. \tag{7.1}$$

and use the correspondence rule $E \to i\hbar\partial_t$ and $\vec{p} \to \frac{\hbar}{i}\vec{\nabla}$ we obtain

$$(i\hbar\partial_t)^2\,\psi(\vec{x},t) = (m^2c^4 - c^2\hbar^2\Delta)\psi(\vec{x},t), \tag{7.2}$$

which can be written as

$$\boxed{\left(\Box + \frac{m^2c^2}{\hbar^2}\right)\psi(\vec{x},t) = 0} \tag{7.3}$$

in terms of the d'Alembert operator $\Box := \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta$, briefly called "*box*." While Schrödinger already knew this relativistic wave equation, it was later named after Klein and Gordon who first published it. An immediate problem for the use of (7.3) as an equation for quantum mechanical wave functions is the existence of negative energy solutions

$$\psi_E(\vec{x},t) = e^{-\frac{i}{\hbar}Et + \frac{i}{\hbar}\vec{p}\vec{x}}, \qquad E = -\sqrt{mc^2 + p^2c^4} \;\leq\; -mc^2 \tag{7.4}$$

with an energy that is unbounded from below. Once interactions are turned on electrons could thus emit an infinite amount of energy, which is clearly unphysical.

If we try to avoid the negative energy solutions of the Klein-Gordon equation (7.3) by using an expansion of the positive square root,

$$E = mc^2\sqrt{1 + \frac{p^2}{m^2c^2}} = mc^2 + \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \frac{p^6}{16m^5c^4} - \ldots \;\geq\; mc^2, \tag{7.5}$$

the Hamiltonian becomes an infinite series with derivatives of arbitrary order and we loose locality. More concretely, it can be shown that localized wavepackets cannot be constructed without contributions from plane waves with negative energies [Itzykson,Zuber].

7.1 The Dirac-equation

Dirac tried to avoid the problem with the negative energy solutions by linearization of the equation for the energy. We get an idea for how this could work by recalling the equation $\sigma_i\sigma_j = \delta_{ij}\mathbb{1} + i\varepsilon_{ijk}\sigma_k$ for the Pauli matrices σ_i, which implies $(\vec{\sigma}\vec{v})^2 = \sigma_i v_i \sigma_j v_j = \vec{v}^2\mathbb{1}$. For massless particles we thus obtain the relativistic energy-momentum relation from a linear equation

$$E\psi = \pm c\vec{\sigma}\vec{p}\,\psi \qquad \Rightarrow \qquad E^2\mathbb{1} = c^2(\vec{p}\vec{\sigma})^2 = c^2p^2\mathbb{1}. \tag{7.6}$$

Upon quantization the suggested Hamiltonian $H = \pm c\vec{p}\vec{\sigma}$ becomes a 2×2 matrix of momentum operators because

$$\vec{\sigma} = \left(\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right) \;\Rightarrow\; \vec{p}\,\vec{\sigma} = p_i\sigma_i = \begin{pmatrix}p_3 & p_1 - ip_2\\ p_1 + ip_2 & -p_3\end{pmatrix} = \tfrac{\hbar}{i}\begin{pmatrix}\partial_3 & \partial_1 - i\partial_2\\ \partial_1 + i\partial_2 & -\partial_3\end{pmatrix}. \quad (7.7)$$

Equation (7.6) indeed shows up as the massless case of the Dirac equation. It is called Weyl equation and it describes the massless neutrinos of the standard model of particle physics.

In 1928 Dirac made the following ansatz for a linear relation between energy and momenta

$$E \cdot \mathbb{1} = c\, p_i\, \alpha_i + m\, c^2\, \beta \quad (7.8)$$

with four dimensionless Hermitian matrices α_i and β. Taking squares and demanding the relativistic energy-momentum relation (7.1) we find

$$E^2\mathbb{1} = (m^2c^4 + c^2p_ip_i)\mathbb{1} = c^2\alpha_i\alpha_j p_ip_j + mc^3p_i(\alpha_i\beta + \beta\alpha_i) + \beta^2m^2c^4, \quad (7.9)$$

which is equivalent to the matrix equations

$$\boxed{\beta^2 = \mathbb{1}, \qquad \{\alpha_i, \beta\} = 0, \qquad \{\alpha_i, \alpha_j\} = 2\delta_{ij}\mathbb{1},} \quad (7.10)$$

where $\{A, B\} = AB + BA$ denotes the anticommutator. Assuming the existence of a solution we arrive at the free Dirac equation

$$i\hbar\partial_t\psi = H\psi, \qquad H = \frac{\hbar}{i}c\alpha_i\partial_i\psi + \beta mc^2 \quad (7.11)$$

with $H = H^\dagger$. The coupling to electromagnetic fields is achieved as in the non-relativistic case with the replacement $\vec{p} \to \frac{\hbar}{i}\vec{\nabla} - \frac{e}{c}\vec{A}$ and $E \to i\hbar\partial_t - V$. With $V = e\phi$ this leads to the **interacting Dirac equation**

$$\boxed{\left(i\hbar\frac{\partial}{\partial t} - e\phi\right)\psi = c\alpha_i\left(\frac{\hbar}{i}\partial_i - \frac{e}{c}A_i\right)\psi + \beta mc^2\psi,} \quad (7.12)$$

for charged particles in an electromagnetic field, where we still need to find matrices α_i and β representing the Dirac algebra (7.10).

While $\alpha_i = \sigma_i$ would satisfy $\{\alpha_i, \alpha_j\} = \delta_{ij}\mathbb{1}$ it is impossible to find a further 2×2 matrix β solving (7.10). A simple argument shows that the dimension of the Dirac matrices has to be even: Since $\beta^2 = \mathbb{1}$ and $\alpha_i\beta = -\beta\alpha_i$ cyclicity of the trace $\mathrm{tr}(M\beta) = \mathrm{tr}(\beta M)$ implies

$$\mathrm{tr}\,\alpha_i = \mathrm{tr}\,\alpha_i\beta^2 = \mathrm{tr}(\alpha_i\beta)\beta = \mathrm{tr}\,\beta(\alpha_i\beta) = -\,\mathrm{tr}\,\beta(\beta\alpha_i) = -\,\mathrm{tr}\,\alpha_i \quad (7.13)$$

and hence $\mathrm{tr}\,\alpha_i = 0$ (similarly $\mathrm{tr}\,\beta = \mathrm{tr}(\beta\alpha_1)\alpha_1 = \mathrm{tr}\,\alpha_1(\beta\alpha_1) = -\,\mathrm{tr}\,\alpha_1(\alpha_1\beta) = -\,\mathrm{tr}\,\beta$ shows that the trace of β also has to vanish). On the other hand, $\alpha_i^2 = \beta^2 = 1$ implies that all

eigenvalues of α_i and β must be ± 1 so that $\mathrm{tr}\,\alpha_i = \mathrm{tr}\,\beta = 0$ entails that half of the eigenvalues are $+1$ and the other half are -1. This is only possible for even-dimensional matrices. Dirac hence tried an ansatz with 4×4 matrices and found the solution

$$\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}, \qquad \beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \qquad \begin{matrix} \alpha_i = \alpha_i^\dagger = \alpha_i^{-1} \\ \beta = \beta^\dagger = \beta^{-1} \end{matrix}, \tag{7.14}$$

where we used a block notation with 2×2 matrices $\mathbb{1}, 0$ and σ_i as matrix entries. In full gory detail the Dirac matrices read $\alpha_1 = \left(\begin{smallmatrix} 0&0&0&1 \\ 0&0&1&0 \\ 0&1&0&0 \\ 1&0&0&0 \end{smallmatrix}\right)$, $\alpha_2 = \left(\begin{smallmatrix} 0&0&0&-i \\ 0&0&i&0 \\ 0&-i&0&0 \\ i&0&0&0 \end{smallmatrix}\right)$, $\alpha_3 = \left(\begin{smallmatrix} 0&0&1&0 \\ 0&0&0&-1 \\ 1&0&0&0 \\ 0&-1&0&0 \end{smallmatrix}\right)$, $\beta = \left(\begin{smallmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&-1&0 \\ 0&0&0&-1 \end{smallmatrix}\right)$, but one should *never use these explicit expressions* and rather work with the defining equations (7.10), or possibly with the block notation (7.14) if a splitting of the 4-component spinor wave function $\psi = (\psi_1, \psi_2, \psi_3, \psi_4)^T$ into two 2-component spinors is unavoidable like in the non-relativistic limit (see below). It can be shown that (7.14) is the unique irreducible unitary representation of the Dirac algebra (7.10), up to unitary equivalence $\alpha_i \to U\alpha_i U^{-1}$, $\beta \to U\beta U^{-1}$ with $UU^\dagger = \mathbb{1}$.

Relativistic spin. In order to derive the spin operator for relativistic electrons we observe that the orbital angular momentum $\vec{L} = \vec{X} \times \vec{P}$ is not conserved for the free Dirac Hamiltonian

$$[H, L_i] = [c\vec{\alpha}\vec{P} + \beta mc^2, \varepsilon_{ijk} X_j P_k] = \tfrac{\hbar}{i} c \varepsilon_{ijk} \alpha_j P_k \neq 0 \tag{7.15}$$

because $[P_l, X_j] = \frac{\hbar}{i}\delta_{lj}$. A conserved operator $\vec{J} = \vec{L} + \vec{S}$ can be constructed by observing that

$$\begin{aligned} [H, \varepsilon_{ijk}\alpha_j\alpha_k] = \varepsilon_{ijk}[c\vec{\alpha}\vec{P}, \alpha_j\alpha_k] &= \varepsilon_{ijk} c P_l \left(\{\alpha_l, \alpha_j\}\alpha_k - \alpha_j\{\alpha_l, \alpha_k\}\right) \\ &= \varepsilon_{ijk} c P_l \left(2\delta_{lj}\alpha_k - 2\delta_{lk}\alpha_j\right) = 4c\varepsilon_{ijk} P_j \alpha_k \end{aligned} \tag{7.16}$$

because $[A, BC] = ABC + BAC - BAC - BCA = \{A, B\}C - B\{A, C\}$. The spin operator

$$S_i = \frac{\hbar}{4i}\varepsilon_{ijk}\alpha_j\alpha_k = \frac{\hbar}{2}\begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} \tag{7.17}$$

therefore yields a conserved total angular momentum $\vec{J} = \vec{L} + \vec{S}$.

Lorentz covariant form of the Dirac equation. Multiplication with the matrix $\frac{1}{c}\beta$ from the left recasts the relation $E - e\phi = c\vec{\alpha}(\vec{p} - \frac{e}{c}\vec{A}) + \beta mc^2$ into

$$\left(\beta(\tfrac{1}{c}E - \tfrac{e}{c}\phi) - \beta\vec{\alpha}(\vec{p} - \tfrac{e}{c}\vec{A}) - mc\right)\psi = 0. \tag{7.18}$$

The standard combination of coordinates, energy-momenta and gauge potentials into 4-vectors

$$x^\mu = (ct, \vec{x}), \qquad \partial_\mu = (\tfrac{1}{c}\partial_t, \vec{\nabla}), \qquad p^\mu = (\tfrac{1}{c}E, \vec{p}), \qquad A^\mu = (\phi, \vec{A}), \tag{7.19}$$

which entails the relativistic version $p_\mu \to P_\mu = i\hbar\partial_\mu$ of the quantum mechanical correspondence (2.3), suggests the combination of β and $\beta\alpha_i$ into a 4-vector γ^μ of 4×4 matrices as

$$\gamma^\mu = (\beta, \beta\vec{\alpha}) \qquad \Rightarrow \qquad \boxed{\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\mathbb{1}} \qquad \text{with} \qquad g^{\mu\nu} = \begin{pmatrix} 1&0&0&0 \\ 0&-1&0&0 \\ 0&0&-1&0 \\ 0&0&0&-1 \end{pmatrix}. \tag{7.20}$$

Putting everything together and dividing (7.18) by $\hbar$ we obtain the Lorentz-covariant equation

$$\boxed{\left(\gamma^\mu(i\partial_\mu - \frac{e}{\hbar c}A_\mu) - \frac{c}{\hbar}m\right)\psi = 0.} \tag{7.21}$$

Since $(\beta\alpha_i)^\dagger = \alpha_i\beta = -\beta\alpha_i$ the gamma matrices γ^μ are unitary, but for $\mu \neq 0$ not Hermitian

$$\boxed{(\gamma^\mu)^\dagger = (\gamma^\mu)^{-1} \equiv \gamma_\mu,} \tag{7.22}$$

where the second equality should be interpreted as a numerical coincidence due to (7.20) and our convention $g_{\mu\nu} = \mathrm{diag}(1,-1,-1,-1)$ and not as a covariant equation! When we will come to the discussion of symmetries we will see that the non-Hermiticity of γ^μ with $\mu \neq 0$ is related to the fact that Lorentz-boosts are represented by non-unitary transformations, as one might expect to be required by consistency of probability density interpretations with Lorentz contractions.

The Dirac sea. The Dirac algebra $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\mathbb{1}$ implies $(\gamma^\mu p_\mu)^2 = \gamma^\mu\gamma^\nu p_\mu p_\nu = p^2$ and

$$(\gamma^\mu p_\mu + mc)(\gamma^\nu p_\nu - mc) = (p^2 - m^2c^2)\mathbb{1} = \frac{1}{c^2}\big(E^2 - (p^2c^2 + m^2c^4)\big)\mathbb{1}. \tag{7.23}$$

Every component $\psi_1, \ldots, \psi_4$ of a solution ψ of the free Dirac equation $(p^\nu\gamma_\nu - mc)\psi = 0$ is hence a solution of the Klein Gordon equation. In turn, the operator $p^\mu\gamma_\mu + mc$ can be used to construct plane wave solutions

$$\psi(t,\vec{x}) = e^{\frac{i}{\hbar}(Et - \vec{p}\vec{x})}\,(\gamma^\mu p_\mu + mc)\,\psi_0 \qquad \text{with} \qquad E = \pm\sqrt{p^2c^2 + m^2c^4} \tag{7.24}$$

in terms of constant spinors ψ_0. For each momentum $\vec{p}$ two of the four spinor polarizations have positive and two have negative energy. This is most easily seen in the rest frame $\vec{p} = 0$ where $\gamma^\mu p_\mu + mc = mc(\mathbb{1} \pm \beta)$ for $E = \pm mc$. For positive energies $\frac{1}{2}(\mathbb{1} + \beta)$ projects onto $\psi_{1,2}$ while for negative energies $\frac{1}{2}(\mathbb{1} - \beta)$ projects onto $\psi_{3,4}$.

As negative energy states are thus unavoidable in the realtivistic quantum theory Dirac concluded that what we observe as the vacuum is a state where all (infinitely many!) negative energy eigenstates are filled by electrons. This vacuum is called *Dirac sea*. Pauli's exclusion principle then forbids transitions to negative energy states because they are already occupied. But the existence of this sea implies that, when interactions are turned on, a sufficient amount of energy $E \geq 2mc^2$ can be used to bring an electron into a positive energy state while leaving a hole in the vacuum. A missing negative charge in a uniform background density, however, acts like a positive charge so that the holes are perceived as positively charged particles, called positrons, and a particle-antiparticle pair has been created out of the vacuum.

The same story can now be told for other particles, but for bosons there is no Pauli exclusion principle and the Dirac sea has to be replaced by the more powerful concept of field quantization

where wave functions are replaced by (superpositions of particle creation and annihilation) operators [Itzykson,Zuber]. This is true, in particular, for the quanta of the electromagnetic field, called photons. We will further discuss the concept of particle creation and annihilation in chapter 10 (many partile theory). While the relativistic Dirac sea should hence not be taken too literally it is still a very powerful intuitive concept and later found important applications in solid state physics.

7.2 Nonrelativistic limit and the Pauli-equation

In this section we show that the Pauli-equation, as well as the fine structure of the hydrogen atom, can be obtained from a non-relativistic approximation of the Dirac-equation. Assuming that the potentials $V = e\phi$ and $\vec{A}$ are time-independent we make the stationary ansatz

$$\psi(\vec{x},t) = w(\vec{x})\, e^{-\frac{i}{\hbar}Et} = \begin{pmatrix} u(\vec{x}) \\ v(\vec{x}) \end{pmatrix} e^{-\frac{i}{\hbar}Et} \tag{7.25}$$

for energy eigenfunctions ψ where we decomposed the 4-spinor $w(\vec{x})$ into two 2-component spinors $u = \binom{u_1}{u_2}$ and $v = \binom{v_1}{v_2}$. For conveniece we introduce the notation

$$\vec{\Pi} = \vec{P} - \frac{e}{c}\vec{A} \qquad \text{with} \qquad \vec{P} = \frac{\hbar}{i}\vec{\nabla} \tag{7.26}$$

for the gauge-invariant physical momentum $\vec{\Pi} = \vec{p} - \frac{e}{c}\vec{A}$ and insert $\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}$, $\beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}$ into the stationary Dirac-equation $Ew = Hw$ with $H = c\vec{\alpha}\vec{\Pi} + V + \beta mc^2$. Putting everything together we obtain

$$\begin{pmatrix} E - V - mc^2 & 0 \\ 0 & E - V + mc^2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 & c\vec{\sigma}\vec{\Pi} \\ c\vec{\sigma}\vec{\Pi} & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}. \tag{7.27}$$

For positive energies

$$E' \equiv E - mc^2 > 0 \tag{7.28}$$

equation (7.25) is now written as a coupled system of two differential equations

$$(E' - V) \quad u = c(\vec{\sigma}\vec{\Pi})\, v \tag{7.29}$$

$$(E' + 2m_0c^2 - V)\, v = c(\vec{\sigma}\vec{\Pi})\, u \tag{7.30}$$

for two spinor wave functions u and v.

In the *non-relativistic approximation* we now assume that all energies, momenta and electromagnetic potentials are small

$$V,\ \vec{A},\ \vec{p},\ E' \ll mc^2. \tag{7.31}$$

Then we can solve (7.30) for the *small components* v of the 4-spinor w as

$$v = \frac{f(\vec{x})}{2mc}\,\vec{\sigma}\vec{\Pi}\,u \qquad \text{with} \qquad f(\vec{x}) = \frac{1}{1+\frac{E'-V(\vec{x})}{2mc^2}} \approx 1 - \frac{E'-V(\vec{x})}{2mc^2}. \tag{7.32}$$

Inserting v into equation (7.29) yields

$$(E'-V)\,u = \frac{(\vec{\sigma}\vec{\Pi})f(\vec{x})(\vec{\sigma}\vec{\Pi})}{2m}\,u \tag{7.33}$$

which can be interpreted as a non-relativistic Schrödinger equation with the Hamilton operator

$$H_{non-rel} = \tfrac{1}{2m}(\vec{\sigma}\vec{\Pi}f(\vec{x})\vec{\sigma}\vec{\Pi}) + V \tag{7.34}$$

and non-relativistic energy E'.

In the evaluation of the resulting operator we now assume a *centrally symmetric potential* $V(\vec{x}) = V(r)$. With $\frac{\partial f(r)}{\partial x_i} = f'(r)\frac{x_i}{r}$, $\sigma_i\sigma_j = \delta_{ij} + i\varepsilon_{ijk}\sigma_k$ and $\vec{\Pi} = \frac{\hbar}{i}\vec{\nabla} - \frac{e}{c}\vec{A}$ we find

$$(\vec{\sigma}\vec{\Pi})f(r)(\vec{\sigma}\vec{\Pi}) = \Pi_i f \Pi_j \sigma_i\sigma_j = \left(f\Pi_i\Pi_j + \frac{\hbar}{i}\frac{\partial f}{\partial x_i}\Pi_j\right)(\delta_{ij} + i\varepsilon_{ijk}\sigma_k) \tag{7.35}$$

whose decomposition according to (anti)symmetry in ij yields four terms,

$$\boxed{\Pi_i f \Pi_j \sigma_i\sigma_j = f\Pi_i\Pi_j\delta_{ij} + if\varepsilon_{ijk}\sigma_i\Pi_j\Pi_k - i\hbar f'\frac{x_i}{r}\Pi_j\delta_{ij} + \hbar f'\frac{x_i}{r}\Pi_j\varepsilon_{ijk}\sigma_k} \tag{7.36}$$

which we now discuss in turn. We will neglect higher order corrections that lead to energy corrections of higher order in the fine structure constant $\alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137}$ and that are hence suppressed by additional powers of $\frac{\alpha}{2\pi} \approx 10^{-3}$.

- **Angular momentum term.** We begin with the evaluation of the first term

$$f\Pi^2 = f\Pi_i\Pi_j\delta_{ij} = f\left(\vec{p}^{\,2} - \frac{e}{c}(\vec{p}\vec{A} + \vec{A}\vec{p}) + \frac{e^2}{c^2}\vec{A}^{\,2}\right). \tag{7.37}$$

  For a constant magnetic field $\vec{B} = \nabla\times\vec{A}$ we choose the vector potential $A_i = -\frac{1}{2}\varepsilon_{ijk}x_jB_k$ that satisfies the Coulomb gauge condition $\operatorname{div}\vec{A} = \partial_iA_i = 0$ and neglect the last term, which is quadratic in B. Since $\partial_iA_i\psi = (\partial_iA_i)\psi + A_i\partial_i\psi = A_i\partial_i\psi$ in Coulomb gauge and

$$-2\,\tfrac{e}{c}\vec{A}\vec{p} = -2\,\tfrac{e}{c}(-\tfrac{1}{2}\varepsilon_{ijk}x_jB_k)p_i = -\tfrac{e}{c}\vec{B}\vec{L} \tag{7.38}$$

  the leading contribution $\frac{1}{2m}f \approx \frac{1}{2m}$ yields the leading kinetic term and the angular momentum coupling

$$\frac{1}{2m}\vec{\Pi}^2 = \frac{1}{2m}\left(\vec{p}^{\,2} - \frac{e}{c}(\vec{\mathcal{L}}\vec{B})\right) + \mathcal{O}(B^2). \tag{7.39}$$

  Relativistic kinetic energy. Since $\vec{p}^{\,2}/2m$ is the leading non-relativistic term we should

also keep its combination with the next-to-leading term in the expansion $f = 1 - \frac{E'-V}{2mc^2} + \ldots$. In the resulting correction term we insert $E' = V + p^2/2m + \ldots$ and obtain

$$\boxed{H_{rel} = -\frac{\vec{p}^4}{8m^3c^2}} \tag{7.40}$$

in agreement with (6.20).

- **Gyromagnetic ratio.** The next term that we need to consider is $\frac{1}{2m}$ times

$$i\varepsilon_{ijk}\sigma_i\Pi_j\Pi_k = -\frac{e}{c}i\varepsilon_{ijk}\sigma_k(p_iA_j + A_ip_j) = -\frac{e}{c}i\varepsilon_{ijk}\sigma_k(p_iA_j - A_jp_i). \tag{7.41}$$

Since $(p_iA_j - A_jp_i)\psi = (p_iA_j)\psi$ this expression becomes

$$-\frac{e}{c}\varepsilon_{ijk}\hbar(\partial_iA_j)\sigma_k = -\frac{e\hbar}{c}(\mathrm{rot}\vec{A})_k\sigma_k = -\frac{e\hbar}{c}\vec{B}\vec{\sigma} = -2\frac{e}{c}\vec{B}\vec{\mathcal{S}} \tag{7.42}$$

which amounts to a g-factor $g = 2$ in the magnetic coupling $-\frac{e}{2mc}(\vec{L} + g\vec{S})\vec{B}$.

- **Darwin term.** For the contribution $\frac{1}{2m}\frac{\hbar}{i}f'\frac{x_i}{r}\Pi_i$ of the third term in (7.36) we insert the derivative $f' \approx V'/2mc^2$ of the r.h.s. of (7.32) and find

$$\frac{1}{2m}\frac{\hbar}{i}f'\frac{x_i}{r}\Pi_i \approx \frac{\hbar}{4im^2c^2}\frac{V'}{r}(\vec{x}\vec{p} - \frac{e}{c}\vec{x}\vec{A}) \approx -\frac{\hbar^2}{4m^2c^2}\partial_iV\partial_i \tag{7.43}$$

where we used $\frac{x_i}{r}V' = \partial_iV$ and dropped the vector potential contribution, which is quadratic in the electromagnetic fields. This term is equivalent to the Darwin term

$$\boxed{H_{Darwin} = \frac{\hbar^2}{8m^2c^2}\Delta V} \tag{7.44}$$

because expectation values of the Hamiltonians agree by partial integration of $\int u\partial_iV\partial_iu = \frac{1}{2}\int\partial_iV\partial_i(u^2) = -\frac{1}{2}\int(\Delta V)u^2$ for real eigenfunctions u.

- **Spin-orbit coupling.** Since $\vec{S} = \frac{\hbar}{2}\vec{\sigma}$ the term $\frac{\hbar}{i}\frac{f'}{r}i\varepsilon_{ijk}x_ip_j\sigma_k$ yields

$$\frac{\hbar}{i}\frac{f'}{r}i\varepsilon_{ijk}x_ip_j\sigma_k = \frac{f'}{r}\hbar\mathcal{L}_k\sigma_k \approx \frac{dV}{dr}\frac{1}{mc^2r}\vec{\mathcal{L}}\vec{\mathcal{S}}, \tag{7.45}$$

which completes our evaluation of the terms in eq. (7.36).

Collecting all relevant contributions we thus obtain the *Pauli-equation*

$$\boxed{H_{Pauli}\,u = \left(\frac{\vec{p}^2}{2m} + V - \frac{e}{2mc}\vec{B}(\vec{L} + 2\vec{\mathcal{S}}) + \frac{dV}{dr}\frac{\vec{L}\vec{\mathcal{S}}}{2m^2c^2r}\right)u,} \tag{7.46}$$

which contains the spin-orbit coupling and explains the observed gyromagnetic ratio. In addition we find the *relativistic energy correction* (7.40) and the *Darwin term* (7.44), which together with (7.46) explain the complete fine structure of the energy levels of the hydrogen atom.

Chapter 8

Scattering Theory

I ask you to look both ways. For the road to a knowledge of the stars leads through the atom; and important knowledge of the atom has been reached through the stars.
-Sir Arthur Eddington (1882 - 1944)

Most of our knowledge about microscopic physics originates from scattering experiments. In these experiments the interactions between atomic or sub-atomic particles can be measured. This is done by letting them collide with a fixed target or with each other. In this chapter we present the basic concepts for the analysis of scattering experiments.

We will first analyze the asymptotic behavior of scattering solutions to the Schrödinger equation and define the differential cross section. With the method of partial waves the scattering amplitudes are then obtained from the phase shifts for spherically symmetric potentials. The Lippmann–Schwinger equation and its formal solution, the Born series, provides a perturbative approximation technique which we apply to the Coulomb potential. Eventually we define the scattering matrix and the transition matrix and relate them to the scattering amplitude.

8.1 The central potential

The physical situation that we have in mind is an incident beam of particles that scatters at some localized potential $V(\vec{x})$ which can represent a nucleus in some solid target or a particle in a colliding beam. For fixed targets we can usually focus on the interaction with a single nucleus. In beam-beam collisions it is more difficult to produce sufficient luminosity, but this has to be dealt with in the ultrarelativistic scattering experiments of particle physics for kinematic

reasons.[1] We will mostly confine our interest to elastic scattering where the particles are not excited and there is no particle production. It is easiest to work in the center of mass frame, where a spherically symmetric potential has the form $V(r)$ with $r = |\vec{x}|$. For a fixed target experiment the scattering amplitude can then easily be converted to the laboratory frame for comparison with the experimental data. Because of the quantum mechanical uncertainty we can only predict the *probability* of scattering into a certain direction, in contrast to the deterministic scattering angle in classical mechanics. With particle beams that contain a sufficiently large number of particles we can, however, measure the probability distribution (or differntial cross section) with arbitrary precision.

8.1.1 Differential cross section and frames of reference

Imagine a beam of monoenergetic particles being scattered by a target located at $\vec{x} = 0$. Let the detector cover a solid angle $d\Omega$ in direction (θ, φ) from the scattering center. We choose a coordinate system

$$\vec{x} = (r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta), \qquad \vec{k}_{in} = \frac{\sqrt{2mE}}{\hbar}\,\vec{e}_3 \tag{8.2}$$

so that the incoming beam travels along the z-axis. The number of particles per unit time entering the detector is then given by $N d\Omega$. The flux of particles F in the incident beam is defined as the number of particles per unit time, crossing a unit area placed normal to the direction of incidence. To characterize the collisions we use the *differential scattering cross-section*

$$\frac{d\sigma}{d\Omega} = \frac{N}{F}, \tag{8.3}$$

which is defined as the ratio of the number of particles scattered into the direction (θ, φ) per unit time, per unit solid angle, divided by the incident flux. The *total scattering cross-section*

$$\sigma_{tot} = \int \left(\frac{d\sigma}{d\Omega}\right) d\Omega = \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \sin\theta \frac{d\sigma}{d\Omega} \tag{8.4}$$

is defined as the integral of the differential scattering cross-section over all solid angles. Both the differential and the total scattering cross-sections have the dimension of an area.

Center-of-Mass System. As shown in fig. 8.1 we denote by $\vec{p}_1$ and $\vec{p}_2$ the momenta of the incoming particles and of the target, respectively. The center of mass momentum is $\vec{p}_g = \vec{p}_1 + \vec{p}_2 = \vec{p}_{1L}$ with the target at rest $\vec{p}_{2L} = 0$ in the laboratory frame. As we derived

[1] Using the notation of figure 8.1 below with an incident particle of energy $E_1 = E_{\rm in} = \sqrt{c^2\vec{p}_{1L}^{\,2} + m_1^2c^4}$ hitting a target with mass m_2 at rest in the laboratory system the total energy

$$E^2 = c^2(p_{1L} + p_{2L})^2 = (E_1 + m_2c^2)^2 - c^2\vec{p}_{1L}^{\,2} = m_1^2c^4 + m_1^2c^4 + 2E_{\rm in}m_2c^2 \tag{8.1}$$

available for particle production in the center of mass system is only $E \approx \sqrt{2E_{in}m_2c^2}$ for $E_{\rm in} \gg m_1c^2, m_2c^2$.

Figure 8.1: Scattering angle for fixed target and in the center of mass frame.

in section 4, the kinematics of the reduced 1-body problem is given by the reduced mass $\mu = m_1 m_2/(m_1 + m_2)$ and the momentum

$$\vec{p} = \frac{\vec{p}_1 m_2 - \vec{p}_2 m_1}{m_1 + m_2}. \tag{8.5}$$

Obviously $\varphi = \varphi_L$, while the relation between θ in the center of mass frame and the angle θ_L in the fixed target (laboratory) frame can be obtained by comparing the momenta $\vec{p}'_1$ of the scattered particles. With $p_i = |\vec{p}_i|$ the transversal momentum is

$$p'_{1L} \sin\theta_L = p'_1 \sin\theta. \tag{8.6}$$

The longitudinal momentum is $\vec{p}'_1 \cos\theta$ in the center of mass frame. In the laboratory frame we have to add the momentum due to the center of mass motion with velocity $\vec{v}_g$, where

$$\vec{p}_{1L} = \vec{p}_1 + m_1 \vec{v}_g = \vec{p}_g = (m_1 + m_2)\vec{v}_g \qquad \Rightarrow \qquad m_2 \vec{v}_g = \vec{p}_1. \tag{8.7}$$

Restricting to elastic scattering where $|\vec{p}_1{}'| = |\vec{p}_1|$ we find for the longitudinal motion

$$p'_{1L} \cos\theta_L = p'_1 \cos\theta + m_1 v_g \overset{el.}{=} p'_1(\cos\theta + \frac{m_1}{m_2}) \tag{8.8}$$

We hence find the formula

$$\tan\theta_L^{elastic} = \frac{\sin\theta}{\cos\theta + \tau} \qquad \text{with} \qquad \tau = \frac{m_1}{m_2} = \frac{m_{in}}{m_{target}} \tag{8.9}$$

for the scattering angle in the laboratory frame for elastic scattering. According to the change of the measure of the angular integration the differential cross section also changes by a factor

$$\left(\frac{d\sigma}{d\Omega}\right)_L (\theta_L(\theta)) = \left|\frac{d(\cos\theta)}{d(\cos\theta_L)}\right| \frac{d\sigma}{d\Omega}(\theta) = \frac{(1 + 2\tau\cos\theta + \tau^2)^{3/2}}{|1 + \tau\cos\theta|} \frac{d\sigma}{d\Omega}(\theta) \tag{8.10}$$

where we used $\cos\theta_L = 1/\sqrt{1 + \tan^2\theta} = (\cos\theta + \tau)/\sqrt{1 + 2\tau\cos\theta + \tau^2}$.

8.1.2 Asymptotic expansion and scattering amplitude

We now consider the scattering of a beam of particles by a fixed center of force and let m denote the reduced mass and $\vec{x}$ the relative coordinate. If the beam of particles is switched on for a long time compared to the time one particle needs to cross the interaction area, steady-state

conditions apply and we can focus on stationary solutions of the time-independent Schrödinger equation

$$\left[-\frac{\hbar^2}{2m}\Delta + V(\vec{x})\right] u(\vec{x}) = Eu(\vec{x}), \qquad \psi(\vec{x},t) = e^{-i\omega t}u(\vec{x}). \tag{8.11}$$

The energy eigenvalues E is related by

$$E = \frac{1}{2}m\vec{v}^2 = \frac{\vec{p}^2}{2m} = \frac{\hbar^2\vec{k}^2}{2m} \tag{8.12}$$

to the incident momentum $\vec{p}$, the incident wave vector $\vec{k}$ and the incident velocity $\vec{v}$. For convenience we introduce the reduced potential

$$U(\vec{x}) = 2m/\hbar^2 \cdot V(\vec{x}) \tag{8.13}$$

so that we can write the Schrödinger equation as

$$[\nabla^2 + k^2 - U(\vec{x})]u(\vec{x}) = 0. \tag{8.14}$$

For potentials that asymptotically decrease faster then r^{-1}

$$|V_{as}(r)| \leq c/r^\alpha \qquad \text{for} \quad r \to \infty \qquad \text{with} \quad \alpha > 1, \tag{8.15}$$

we can neglect $U(\vec{x})$ for large r and the Schrödinger equation reduces to the Helmholtz equation of a free particle

$$[\Delta + k^2]u_{as}(\vec{x}) = 0. \tag{8.16}$$

Potentials satisfying (8.15) are called *finite range*. (The important case of the Coulomb potential is, unfortunately, of infinite range, but we will be able to treat it as the limit $\alpha \to 0$ of the finite range Yukawa potential $e^{-\alpha r}/r$.) For large r we can decompose the wave function into a part u_{in} describing the incident beam and a part u_{sc} for the scattered particles

$$u(\vec{x}) \;\to\; u_{in}(\vec{x}) + u_{sc}(\vec{x}) \qquad \text{for} \qquad r \to \infty. \tag{8.17}$$

Since we took the z-axis as the direction of incidence and since the particles have all the same momentum $p = \hbar k$ the incident wave function can be written as

$$u_{in}(\vec{x}) = e^{i\vec{k}\cdot\vec{x}} = e^{ikz}, \tag{8.18}$$

where we were free to normalize the amplitude of u_{in} since all equations are linear.

Far from the scattering center the scattered wave function represents an outward radial flow of particles. We can parametrize it in terms of the *scattering amplitude* $f(k,\theta,\varphi)$ as

$$u_{sc}(\vec{x}) = f(k,\theta,\varphi)\frac{e^{ikr}}{r} + \mathcal{O}(\frac{1}{r^\alpha}), \tag{8.19}$$

where (r, θ, φ) are the polar coordinates of the position vector $\vec{x}$ of the scattered particle. The asymptotic form u_{as} of the scattering solution thus becomes

$$u_{as} = (e^{i\vec{k}\cdot\vec{x}})_{as} + f(k, \theta, \varphi)\frac{e^{ikr}}{r}. \tag{8.20}$$

The scattering amplitude can now be related to the differential cross-section. From chapter 2 we know the probability current density for the stationary state

$$\vec{j}(\vec{x}) = \frac{\hbar}{2im}(\psi^*\vec{\nabla}\psi - \psi\vec{\nabla}\psi^*) = \frac{\hbar}{m}\,\mathrm{Re}\,(u^*\vec{\nabla}u) \tag{8.21}$$

with the gradient operator in spherical polar coordinates (r, θ, φ) reading

$$\vec{\nabla} = \vec{e}_r\,\frac{\partial}{\partial r} + \vec{e}_\theta\,\frac{1}{r}\frac{\partial}{\partial\theta} + \vec{e}_\varphi\,\frac{1}{r\sin\theta}\frac{\partial}{\partial\varphi}. \tag{8.22}$$

For large r the scattered particle current flows in radial direction with

$$j_r = \frac{\hbar k}{mr^2}|f(k, \theta, \varphi)|^2 + \mathcal{O}(\frac{1}{r^3}). \tag{8.23}$$

Since the area of the detector is $r^2 d\Omega$ the number of particles $N d\Omega$ entering the detector per unit time is

$$N d\Omega = \frac{\hbar k}{m}|f(k, \theta, \varphi)|^2 d\Omega. \tag{8.24}$$

For $|\psi_{in}(\vec{x})|^2 = 1$ the incoming flux $F = \hbar k/m = v$ is given by the particle velocity. We thus obtain the differential cross-section

$$\boxed{\frac{d\sigma}{d\Omega} = |f(k, \theta, \varphi)|^2} \tag{8.25}$$

as the modulus squared of the scattering amplitude.

8.2 Partial wave expansion

For a spherically symmetric central potential $V(\vec{x}) = V(r)$ we can use rotation invariance to simplify the computation of the scattering amplitude by an expansion of the angular dependence in spherical harmonics. Since the system is completely symmetric under rotations about the direction of incident beam (chosen along the z-axis), the wave function and the scattering amplitude do not depend on φ. Thus we can expand both $u_{\vec{k}}(r, \theta)$ and $f(k, \theta)$ into a series of Legendre polynomials, which form a complete set of functions for the interval $-1 \leqslant \cos\theta \leqslant +1$,

$$u_{\vec{k}}(r, \theta) = \sum_{l=0}^{\infty} R_l(k, r)P_l(\cos\theta), \tag{8.26}$$

$$f(k, \theta) = \sum_{l=0}^{\infty}(2l + 1)f_l(k)P_l(\cos\theta), \tag{8.27}$$

where the factor $(2l+1)$ in the definition of the *partial wave amplitudes* $f_l(k)$ corresponds to the degeneracy of the magnetic quantum number. (Some authors use different conventions, like either dropping the factor $(2l+1)$ or including an additional factor $1/k$ in the definition of f_l.) The terms in the series (8.26) are known as a *partial waves*, which are simultaneous eigenfunctions of the operators $\mathcal{L}^2$ and $\mathcal{L}_z$ with eigenvalues $l(l+1)\hbar^2$ and 0, respectively. Our aim is now to determine the amplitudes f_l in terms of the radial functions $R_l(k,r)$ for solutions (8.27) to the Schrödinger equation.

The radial equation. We recall the formula for the Laplacian in spherical coordinates

$$\Delta = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{\mathcal{L}^2}{\hbar^2 r^2} \qquad \text{with} \quad -\frac{\mathcal{L}}{\hbar^2} = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2} \tag{8.28}$$

With the separation ansatz

$$u_{Elm}(\vec{x}) = R_{El}(r)Y_{lm}(\theta,\varphi) \tag{8.29}$$

for the time-independent Schrödinger equation with central potential in spherical coordinates

$$\left\{-\frac{\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{\mathcal{L}^2}{\hbar^2 r^2}\right] + V(r)\right\} u(\vec{x}) = Eu(\vec{x}), \tag{8.30}$$

and $\mathcal{L}^2 Y_{lm}(\theta,\varphi) = l(l+1)\hbar^2 Y_{lm}(\theta,\varphi)$ we find the *radial equation*

$$\left(-\frac{\hbar^2}{2m}\left(\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr}\right) + \frac{l(l+1)\hbar^2}{2mr^2} + V(r)\right) R_{El}(r) = ER_{El}(r). \tag{8.31}$$

and its reduced form

$$\boxed{\left(\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2} - U(r) + k^2\right) R_l(k,r) = 0} \tag{8.32}$$

with $k = \sqrt{2mE/\hbar^2}$ and the reduced potential $U(r) = (2m/\hbar^2)V(r)$.

Behavior near the center. For potentials less singular than r^{-2} at the origin the behavior of $R_l(k,r)$ at $r=0$ can be determined by expanding R_l into a power series

$$R_l(k,r) = r^s \sum_{n=0}^{\infty} a_n r^n. \tag{8.33}$$

Substitution into the radial equation (8.32) leads to the quadratic indicial equation with the two solutions $s = l$ and $s = -(l+1)$. Only the first one leads to a non-singular wave function $u(r,\theta)$ at the origin $r=0$.

Introducing a new radial function $\tilde{R}_{El}(r) = rR_{El}(r)$ and substituting into (8.31) leads to the equation

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dr^2} + V_{eff}(r)\right)\tilde{R}_{El}(r) = E\tilde{R}_{El}(r) \tag{8.34}$$

which is similar to the one-dimensional Schrödinger equation but with $r \geq 0$ and an effective potential

$$V_{eff} = V(r) + \frac{l(l+1)\hbar^2}{2mr^2} \tag{8.35}$$

containing the repulsive *centrifugal barrier* term $l(l+1)\hbar^2/2mr^2$ in addition to the interaction potential $V(r)$.

Free particles and asymptotic behavior. We now solve the radial equation for $V(r) = 0$ so that our solutions can later be used either for the representation of the wave function of a free particle at any radius $0 \leq r < \infty$ or for the asymptotic form as $r \to \infty$ of scattering solutions for finite range potentials. Introducing the dimensionless variable $\rho = kr$ with $R_l(\rho) = R_{El}(r)$ for $U(r) = 0$ the radial equation (8.31) turns into the *spherical Bessel differential equation*

$$\left[\frac{d^2}{d\rho^2} + \frac{2}{\rho}\frac{d}{d\rho} + \left(1 - \frac{l(l+1)}{\rho^2}\right)\right] R_l(\rho) = 0, \tag{8.36}$$

whose independent solutions are the *spherical Bessel functions*

$$j_l(\rho) = (-\rho)^l \left(\frac{1}{\rho}\frac{d}{d\rho}\right)^l \frac{\sin\rho}{\rho} \tag{8.37}$$

and the *spherical Neumann functions*

$$n_l(\rho) = -(-\rho)^l \left(\frac{1}{\rho}\frac{d}{d\rho}\right)^l \frac{\cos\rho}{\rho}. \tag{8.38}$$

Their leading behavior at $\rho = 0$,

$$\lim_{\rho\to 0} j_l(\rho) \;\to\; \frac{\rho^l}{1\cdot 3\cdot 5\cdot \ldots \cdot (2l+1)}, \tag{8.39}$$

$$\lim_{\rho\to 0} n_l(\rho) \;\to\; -\frac{1\cdot 3\cdot 5\cdot \ldots \cdot (2l-1)}{\rho^{l+1}} \tag{8.40}$$

can be obtained by expanding $\rho^{-1}\sin\rho$ and $\rho^{-1}\cos\rho$ into a power series in ρ. In accord with our previous result for the ansatz (8.33) the spherical Neumann function $n_l(\rho)$ has a pole of order $l+1$ at the origin and is therefore an *irregular* solution, whereas the spherical Bessel function $j_l(\rho)$ is the *regular* solution with a zero of order l at the origin. The radial part of the wave function of a free particle can hence only contain spherical bessel functions $R^{free}_{El}(r) \propto j_l(kr)$.

8.2.1 Expansion of a plane wave in spherical harmonics

In order to use the spherical symmetry of a potential $V(r)$ we need to expand the plane wave representing the incident particle beam in terms of spherical harmonics. Since $e^{i\vec{k}\cdot\vec{x}}$ is a regular solution to the free Schrödinger equation we can make the ansatz

$$e^{i\vec{k}\cdot\vec{x}} = \sum_{l=0}^{\infty}\sum_{m=-l}^{+l} c_{lm} j_l(kr) Y_{lm}(\theta,\varphi), \tag{8.41}$$

where the radial part is given by the spherical Bessel functions with constants c_{lm} that have to be determined. Choosing $\vec{k}$ in the direction of the z-axis the wave function $\exp(i\vec{k}\cdot\vec{r}) = \exp(ikr\cos\theta)$ is independent of φ so that only the Y_{lm} with $m = 0$, which are proportional to the Legendre polynomials $P_l(\theta)$, can contribute to the expansion

$$e^{ikr\cos\theta} = \sum_{l=0}^{\infty} a_l j_l(kr) P_l(\cos\theta). \tag{8.42}$$

With $\rho = kr$ and $u = \cos\theta$ this becomes

$$e^{i\rho u} = \sum_{l=0}^{\infty} a_l j_l(\rho) P_l(u). \tag{8.43}$$

One way of determining the coefficients a_l is to differentiate this ansatz with respect to ρ,

$$iue^{i\rho u} = \sum_{l} a_l \frac{dj_l}{d\rho} P_l. \tag{8.44}$$

The left hand side of (8.44) can now be evaluated by inserting the series (8.43) and using the recursion relation

$$(2l+1)uP_l^m = (l+1-m)P_{l+1}^m + (l+m)P_{l-1}^m \tag{8.45}$$

of the Legendre polynomials for $m = 0$. This yields

$$i\sum_{l=0}^{\infty} a_l j_l \left(\frac{l+1}{2l+1}P_{l+1} + \frac{l}{2l+1}P_{l-1}\right) = \sum_{l=0}^{\infty} a_l j_l' P_l \tag{8.46}$$

and, since the Legendre polynomials are linearly independent, for the coefficient of P_l

$$a_l j_l' = i\left(\frac{l}{2l-1}a_{l-1}j_{l-1} + \frac{l+1}{2l+3}a_{l+1}j_{l+1}\right). \tag{8.47}$$

The derivative j_l' can now be expressed in terms of $j_{l\pm1}$ by using the recursion relations

$$j_{l-1} = \left(\frac{d}{d\rho} + \frac{l+1}{\rho}\right) j_l = \frac{1}{\rho^{l+1}}\frac{d}{d\rho}(\rho^{l+1} j_l) \tag{8.48}$$

and

$$(2l+1)j_l = \rho[j_{l+1} + j_{l-1}], \tag{8.49}$$

which imply

$$j_l' = j_{l-1} - \frac{l+1}{\rho}j_l = j_{l-1} - \frac{l+1}{2l+1}(j_{l+1} + j_{l-1}) = \frac{l}{2l+1}j_{l-1} - \frac{l+1}{2l+1}j_{l+1} \tag{8.50}$$

[the equations (8.48-8.50) also holds for the spherical Neumann functions n_l]. Substituting this expression for j_l' into eq. (8.47) we obtain the two equivalent recursion relations

$$\frac{a_l}{2l+1} = i\frac{a_{l-1}}{2l-1} \qquad \text{and} \qquad \frac{a_l}{2l+1} = -i\frac{a_{l+1}}{2l+3} \tag{8.51}$$

as coefficients of the independent functions $j_{l-1}(\rho)$ and $j_{l+1}(\rho)$, respectively. These relations have the solution $a_l = (2l+1)i^l a_0$. The coefficient a_0 is obtained by evaluating our ansatz at $\rho = 0$: Since $j_l(0) = \delta_{l0}$ and $P_0(u) = 1$ eq. (8.43) implies $a_0 = 1$, so that the expansion of a plane wave in spherical harmonics becomes

$$\boxed{e^{ikr\cos\theta} = \sum_{l=0}^{\infty}(2l+1)i^l j_l(kr)P_l(\cos\theta).} \tag{8.52}$$

Using the addition theorem of spherical harmonics

$$\frac{2l+1}{4\pi}P_l(\cos\alpha) = \sum_{m=-l}^{+l} Y^*_{lm}(\theta_1,\varphi_1)Y_{lm}(\theta_2,\varphi_2) \tag{8.53}$$

with α being the angle between the directions (θ_1,φ_1) and (θ_2,φ_2) this result can be generalized to the expansion of the plane wave in any polar coordinate system

$$e^{i\vec{k}\cdot\vec{x}} = 4\pi\sum_{l=0}^{\infty}\sum_{m=-l}^{+l} i^l j_l(kr)Y^*_{lm}(\theta_{\vec{k}},\varphi_{\vec{k}})Y_{lm}(\theta_{\vec{x}},\varphi_{\vec{x}}), \tag{8.54}$$

where the arguments of Y^*_{lm} and Y_{lm} are the angular coordinates of $\vec{k}$ and $\vec{x}$, respectively.

8.2.2 Scattering amplitude and phase shift

The computation of the scattering data for a given potential requires the construction of the regular solution of the radial equation. In the next section we will solve this problem for the example of the square well, but first we analyse the asymptotic form of the partial waves in order to find out how to extract and interpret the relevant data.

For large r we can neglect the potential $U(r)$ and it is common to write the asymptotic form of the radial solutions as a linear combination of the spherical Bessel and Neumann functions

$$R_l(k,r) = B_l(k)j_l(kr) + C_l(k)n_l(kr) + \mathcal{O}(r^{-\alpha}) \tag{8.55}$$

with coefficients $B_l(k)$ and $C_l(k)$ that depend on the incident momentum k. Inserting the asymptotic forms

$$j_l(kr) = \frac{1}{kr}\sin\left(kr - \frac{l\pi}{2}\right) + \mathcal{O}(\frac{1}{r^2}), \tag{8.56}$$

$$n_l(kr) = -\frac{1}{kr}\cos\left(kr - \frac{l\pi}{2}\right) + \mathcal{O}(\frac{1}{r^2}), \tag{8.57}$$

we can write

$$R^{as}_l(k,r) = \frac{1}{kr}\left[B_l(k)\sin\left(kr - \frac{l\pi}{2}\right) - C_l(k)\cos\left(kr - \frac{l\pi}{2}\right)\right] \tag{8.58}$$

$$= A_l(k)\frac{1}{kr}\sin\left(kr - \frac{l\pi}{2} + \delta_l(k)\right) \tag{8.59}$$

where

$$A_l(k) = [B_l^2(k) + C_l^2(k)]^{1/2} \tag{8.60}$$

and

$$\delta_l(k) = -\tan^{-1}[C_l(k)/B_l(k)]. \tag{8.61}$$

The $\delta_l(k)$ are called *phase shifts*. We will see that they are real functions of k and completely characterize the strength of the scattering of the lth partial wave by the potential $U(r)$ at the energy $E = \hbar^2 k^2/2m$. In order to relate the phase shifts to the scattering amplitude we now insert the asymptotic form of the expansion (8.52) of the plane wave

$$\begin{aligned} e^{i\vec{k}\vec{x}} &= \sum_{l=0}^{\infty}(2l+1)i^l j_l(kr)P_l(\cos\theta). && (8.62)\\ &\to \sum_{l=0}^{\infty}(2l+1)i^l(kr)^{-1}\sin\left(kr-\frac{l\pi}{2}\right)P_l(\cos\theta). && (8.63)\end{aligned}$$

into the scattering ansatz (8.20)

$$u_{\vec{k}}^{as}(r,\theta) \to e^{i\vec{k}\vec{x}} + f(k,\theta)\frac{e^{ikr}}{r}. \tag{8.64}$$

With $\sin x = (e^{ix} - e^{-ix})/(2i)$ and the partial wave expansions (8.26)–(8.27) of $u(\vec{x})$ and $f(\theta,\varphi)$ we can write the radial function $R_l(k,r)$, i.e. the coefficient of $P_l(\cos\theta)$, asymptotically as

$$\begin{aligned} R_l^{as}(k,r) &= (2l+1)i^l(kr)^{-1}\sin\left(kr-\frac{l\pi}{2}\right) + \frac{2l+1}{r}e^{ikr}f_l(k) && (8.65)\\ &= \frac{2l+1}{2ikr}\left(i^l\left(\frac{e^{ikr}}{i^l} - \frac{e^{-ikr}}{(-i)^l}\right) + 2ike^{ikr}f_l\right) && (8.66)\end{aligned}$$

Rewriting (8.59) in terms of exponentials

$$R_l^{as}(k,r) = \frac{A_l}{2ikr}\left(\frac{e^{i(kr+\delta_l)}}{i^l} - \frac{e^{-i(kr+\delta_l)}}{(-i)^l}\right) \tag{8.67}$$

comparison of the coefficients of e^{-ikr} implies

$$A_l(k) = (2l+1)i^l e^{i\delta_l(k)}. \tag{8.68}$$

The coefficients of $e^{ikr}/(2ikr)$ are $(2l+1)(1+2ikf_l)$ and $A_l e^{i\delta_l}/i^l$, respectively. Hence

$$f_l(k) = \frac{e^{2i\delta_l(k)} - 1}{2ik} = \frac{1}{k}e^{i\delta_l}\sin\delta_l. \tag{8.69}$$

The scattering amplitude

$$f(k,\theta) = \sum_{l=0}^{\infty}(2l+1)f_l(k)P_l(\cos\theta) = \frac{1}{2ik}\sum_{l=0}^{\infty}(2l+1)(e^{2i\delta_l} - 1)P_l(\cos\theta) \tag{8.70}$$

hence *depends only on the phase shifts* $\delta_l(k)$ and the asymptotic form of $R_l(k,r)$ takes the form

$$R_l^{as}(k,r) = -\frac{1}{2ik}A_l(k)e^{-i\delta_l(k)}\left[\frac{e^{-i(kr-l\pi/2)}}{r} - S_l(k)\frac{e^{i(kr-l\pi/2)}}{r}\right] \tag{8.71}$$

where we defined

$$S_l(k) = e^{2i\delta_l(k)}. \tag{8.72}$$

S_l is the partial wave contribution to the *S-matrix*, which we will introduce in the last section of this chapter. Reality of the phase shift $|S_l| = 1$ expresses equality of the incoming and outgoing particle currents, i.e. conservation of particle number or unitarity of the S matrix. For inelastic scattering we could write the radial wave fuction as (8.71) with $S_l = s_l e^{i\delta_l}$ for $s_l \leq 1$ describing the loss of part of the incoming current into inelastic processes like energy transfer of particle production. (The complete scattering matrix, including the contribution of inelastic channels, would however still be unitary as a consequence of the conservation of probability.)

The optical theorem. The total cross section for scattering by a central potential can be written as

$$\sigma_{\rm tot} = \int |f(k,\theta)|^2 d\Omega = 2\pi \int_{-1}^{+1} d(\cos\theta) f^*(k,\theta) f(k,\theta). \tag{8.73}$$

Using (8.70) and the orthogonality property of the Legendre polynomials

$$\int_{-1}^{+1} d(\cos\theta) P_l(\cos\theta) P_{l'}(\cos\theta) = \frac{2}{2l+1}\delta_{ll'} \tag{8.74}$$

we find

$$\sigma_{\rm tot} = \sum_{l=0}^{\infty} 4\pi(2l+1)|f_l(k)|^2 = \sum_{l=0}^{\infty}\sigma_l \qquad \text{with} \qquad \sigma_l = \frac{4\pi}{k^2}(2l+1)\sin^2\delta_l. \tag{8.75}$$

Since (8.69) implies $\mathrm{Im}\, f_l = k|f_l|^2$ we can set $\theta = 0$ in (8.70) and use the fact that $P_l(1) = 1$ to obtain the *optical theorem*

$$\boxed{\sigma_{\rm tot} = \frac{4\pi}{k}\mathrm{Im} f(k,\theta=0),} \tag{8.76}$$

The optical theorem can be shown to hold also for inelastic scattering with $\sigma_{tot} = \sigma_{el} + \sigma_{inel}$. The proof relates the total cross section to the interference of the incoming with the forward-scattered amplitude so that (8.76) is a consequence of the unitarity of the S-matrix [Hittmair].

8.2.3 Example: Scattering by a square well

The centrally symmetric square well is a potential for which the phase shifts can be calculated by analytical methods. Starting with the radial equation (8.32) and the reduced potential

$$U(r) = \begin{cases} -U_0, & r < a \quad (U_0 > 0) \\ 0, & r > a, \end{cases} \tag{8.77}$$

we can write the radial equation inside the well as

$$\left[\frac{d^2}{dr^2}+\frac{2}{r}\frac{d}{dr}-\frac{l(l+1)}{r^2}+K^2\right]R_l(k,r)=0 \qquad \text{for} \quad r<a \tag{8.78}$$

with $K^2=k^2+U_0$. Inside the well the regular solution is thus

$$R_l^{in}(K,r)=N_l\, j_l(Kr), \qquad r<a \tag{8.79}$$

where N_l is related to the exact solution in the exterior region

$$R_l^{ext}(k,r)=B_l(k)[j_l(kr)-\tan\delta_l(k)n_l(kr)], \qquad r>a \tag{8.80}$$

by the matching condition at $r=a$. Continuity of R and R' at $r=a$ hence implies

$$N_l\, j_l(Ka) = B_l(j_l(ka)-\tan\delta_l\, n_l(ka)), \tag{8.81}$$

$$KN_l\, j_l'(Ka) = kB_l\left(j_l'(ka)-\tan\delta_l\, n_l'(ka)\right). \tag{8.82}$$

The ratio of these two equations yields an equation for $\tan\delta_l(k)$ whose solution is

$$\tan\delta_l(k)=\frac{kj_l'(ka)j_l(Ka)-Kj_l(ka)j_l'(Ka)}{kn_l'(ka)j_l(Ka)-Kn_l(ka)j_l'(Ka)} \tag{8.83}$$

with $K=\sqrt{k^2+U_0}$.

In the low energy limit $k=\sqrt{2mE}/\hbar\to 0$ we can insert the leading behavior $j_l(\rho)\propto\rho^l$ and $n_l(\rho)\propto\rho^{-l-1}$, and thus for the derivatives $j_l'(\rho)\propto\rho^{l-1}$ and $n_l'(\rho)\propto\rho^{-l-2}$, to conclude that $\tan\delta_l(k)$ goes to zero like a constant times $k^l/k^{-l-1}=k^{2l+1}$. In this limit the cross section,

$$\boxed{\sigma_l\propto k^{4l},} \tag{8.84}$$

is dominated by $l=0$ so that the scattering probability approximately goes to a θ-independent constant. With

$$j_0(\rho)=\tfrac{\sin\rho}{\rho}, \quad n_0(\rho)=-\tfrac{\cos\rho}{\rho}, \qquad j_0'(\rho)=\tfrac{\rho\cos\rho-\sin\rho}{\rho^2}, \quad n_0'(\rho)=\tfrac{\rho\sin\rho+\cos\rho}{\rho^2} \tag{8.85}$$

and the abbreviations $x=ka$, $X=Ka$ we find

$$\tan\delta_0=\tfrac{((x\cos x-\sin x)\sin X-\sin x(X\cos X-\sin X))/(xX)}{(x\sin x+\cos x)\sin X+\cos x(X\cos X-\sin X))/(xX)}=\tfrac{x\cos x\sin X-X\sin x\cos X}{x\sin x\sin X+X\cos x\cos X}. \tag{8.86}$$

Dividing numerator and denominator by $\cos x\cos X$ we obtain the result

$$\tan\delta_0(k)=\frac{k\tan(Ka)-K\tan(ka)}{K+k\tan(ka)\tan(Ka)}. \tag{8.87}$$

For $k\to 0$ we observe that $\tan\delta_0$ becomes proportial to k. The limit

$$\boxed{a_s=-\lim_{k\to 0}\frac{\tan\delta_0(k)}{k}} \tag{8.88}$$

is called *scattering length* and it determines the limit of the partial cross section

$$\sigma_0 = \frac{4\pi}{k^2}\sin^2\delta_0 = \frac{4\pi}{k^2}\frac{1}{1+\cot^2\delta_0(k)} \quad \stackrel{k\to 0}{\longrightarrow} \quad 4\pi a_s^2. \tag{8.89}$$

For the square well we find

$$a_s = \left(1 - \frac{\tan(a\sqrt{U_0})}{a\sqrt{U_0}}\right) a, \tag{8.90}$$

The coefficient of the next term of the expansion

$$k\cot\delta_0(k) = -\frac{1}{a_s} + \frac{1}{2}r_0k^2 + \cdots \tag{8.91}$$

defines the *effective range* r_0. This definition of the scattering length a_s and the effective range r_0 can be used for all short-range potentials.

Another exactly solvable potential is the hard-sphere potential

$$U(r) = \begin{cases} +\infty, & r < a, \\ 0, & r > a, \end{cases} \tag{8.92}$$

for which the total cross section can be shown to obey

$$\sigma(k) \;\to\; \begin{cases} 4\pi a^2, & k \to 0, \\ 2\pi a^2, & k \gg 1/a. \end{cases} \tag{8.93}$$

For $k \to 0$ the scattering length a_s hence coincides with a and the cross section is 4 times the classical value. For $ka \gg 1$ the wave lengths of the scattered particles goes to 0 and one might naively expect to observe the classical area $a^2\pi$. The fact that quantum mechanics yield twice that value is in accord with refraction phenomena in optics and can be attributed to interference between the incoming and the scattered beam close to the forward direction. This effect is hence called refraction scattering, or shadow scattering.

8.2.4 Interpretation of the phase shift

For a weak and slowly varying potential we may think of the phase shift as arising from the change in the effective wavelength $k \sim \sqrt{2m(E-V(x))}/\hbar$ due to the presence of the potential. For an attractive potential we hence expect an advanced oscillation and a positive phase shift $\delta_l > 0$, while a repulsive potential should lead to retarded oscillation and a negative phase shift $\delta_l < 0$. Comparing this expectation with the result (8.90) for the square well and using $\tan x \approx x + \frac{1}{3}x^3$ for small U_0 we find $a_s \approx -\frac{1}{3}a^3U_0$ so that indeed the scattering length (8.88) becomes negative and the phase shift δ_0 positive for an attractive potential $U_0 > 0$. It can also be shown quite generally that small angular momenta dominate the scattering at low energies and that the partial cross sections σ_l are negligible for $l > ka$ where a is the range of the potential.

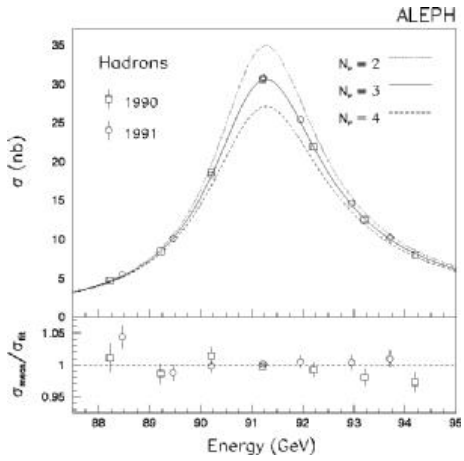


Figure 8.2: Z boson resonance in e^+e^- scattering at LEP and light neutrin number.

As we increase the energy the phase shift varies and the partial cross sections

$$\sigma_l(E) = \frac{4\pi}{k^2}(2l+1)\sin^2\delta_l = \frac{4\pi}{k^2}(2l+1)\frac{1}{1+\cot^2\delta_l} \tag{8.94}$$

go through maxima and zeros as the phase shift δ_l goes through odd and even multiples of π, respectively. For small energies the single cross section σ_0 dominates so that we can get minima where the target becomes almost transparent. This is called *Ramsauer-Townsend effect*.

A rapid move of the phase shift through an odd multiple of π, i.e.

$$\cot\delta_l \approx \big((n+\tfrac{1}{2})\pi - \delta_l\big) \approx \frac{E_R - E}{\Gamma(E)/2} + \mathcal{O}(E_R - E)^2 \qquad \text{for} \qquad \delta_l \approx (n+\tfrac{1}{2})\pi \tag{8.95}$$

with $\Gamma(E_R)$ small at the resonance energy E_R, leads to a sharp peak in the cross section with an angular distibution characteristic for the angular momentum chanel l. This is called *resonance scattering* and described by the *Breit-Wigner resonance formula*

$$\sigma_l(E) = \frac{4\pi}{k^2}(2l+1)\frac{\Gamma^2/4}{(E-E_R)^2+\Gamma^2/4}. \tag{8.96}$$

A resonance can be thought of as a metastable bound state with positive energy whose lifetime is $\hbar/\Gamma$. For a sharp resonance the inverse width Γ^{-1} is indeed related to the dwelling time of the scattered particles in the interaction region. Note that σ_{max} at a resonance is determined by the momentum k of the scattered particles and not by properties of the target. A striking example of a resonance in particle physics is the peak in electron-positron scattering at the Z-boson mass which was analyzed by the LEP-experiment ALEPH as shown in fig. 8.2. Since the Z boson has no electric charge but couples to the weakly interacting particles its lifetime is very sensitive to the number of light neutrinos, which are otherwise extremely hard to observe. This experiment confirmed with great precision the number $N_\nu = 3$ of such species, which is also required for nucleosynthesis, about one second after the big bang, to produce the right amount of helium and other light elements as observed in the interstellar gas clouds.

Resonances can be interpreted as poles in the scattering amplitudes that are close to the real axis (with the imaginary part related to the lifetime). Poles on the positive imaginary axis, on the other hand, correspond to bound states for the potential $V(x)$. The information of the number of such bound states is also contained in the phase shift. For the precise statement we fix the ambiguity modulo 2π in the definition (8.61) of δ_l by requiring continuity. The *Levinson theorem* then states that

$$\delta_l(0) - \delta_l(\infty) = n_l\,\pi \qquad \text{for} \qquad l > 0, \tag{8.97}$$

where n_l denotes the number of bound states with angular momentum l [Chadan-Sabatier]. The theorem also holds for $l = 0$ except for a shift $n_l \to n_l + \frac{1}{2}$ in the formula (8.97) if there is a so-called bound state a zero energy with $l = 0$. While we consider in this chapter the problem of determining the scattering data from the potential, in inverse problem of obtaining information on the potential from the scattering data is physically equally important, but mathematically quite a bit more complicated. *Inverse scattering theory* has been a very active field of research in the last decades with a number of interesting interrelations to other fields like integrable systems [Chadan-Sabatier].

8.3 The Lippmann-Schwinger equation

We can use the method of *Green's functions* to solve the stationary Schrödinger equation (8.14)

$$(\nabla^2 + k^2)u(\vec{x}) = U(\vec{x})u(\vec{x}). \tag{8.98}$$

Using the defining equation of the Green's function for the Helmholtz equation

$$(\nabla^2 + k^2)G_0(k, \vec{x}, \vec{x}') = \delta(\vec{x} - \vec{x}') \tag{8.99}$$

we can write down the general solution of equation (8.98) as a convolution integral

$$u(\vec{x}) = u_{hom}(\vec{x}) + \int G_0(k, \vec{x}, \vec{x}')U(\vec{x}')u(\vec{x}')\; d^3x' \tag{8.100}$$

where u_{hom} is a solution of the homogenous Schrödinger equation

$$(\nabla^2 + k^2)u_{hom}(\vec{x}) = 0. \tag{8.101}$$

We will see that the scattering boundary condition (8.20) is equivalent to taking $u_{hom}(\vec{x})$ to be an incident plane wave

$$u_{hom}(\vec{x}) = \phi_{\vec{k}}(\vec{x}) \equiv e^{i\vec{k}\vec{x}} \tag{8.102}$$

if $G_0 = G_0^{ret}$ is the retarded Green's function. The existence of solutions to the homogeneous equation is of course related to the ambiguity of G_0, as we will see explicitly in the following computation.

Since (8.99) is a linear differential equation with constant coefficients we can determine the Green's function by Fourier transformation. Because of translation invariance

$$G_0(k,\vec{x},\vec{x}') = G_0(k,\vec{R}) \qquad \text{with} \qquad \vec{R} = \vec{x} - \vec{x}', \tag{8.103}$$

hence

$$G_0(k,\vec{x}-\vec{x}') = \frac{1}{(2\pi)^3}\int e^{i\vec{K}\cdot\vec{R}}\,\tilde{g}_0(k,\vec{K})d\vec{K} \tag{8.104}$$

$$\delta(\vec{x}-\vec{x}') = \frac{1}{(2\pi)^3}\int e^{i\vec{K}\cdot\vec{R}}d\vec{K}. \tag{8.105}$$

Substituting the Fourier representations into the defining equation of the Green's function (8.99) we find that

$$\tilde{g}_0(k,\vec{K}) = \frac{1}{k^2-K^2}. \tag{8.106}$$

Since $\tilde{g}_0$ has a pole on the real axis we give a small imaginary part to k and define

$$G_0^{\pm}(k,\vec{x},\vec{x}') = \frac{1}{(2\pi)^3}\int \frac{e^{i\vec{K}\cdot(\vec{x}-\vec{x}')}}{k^2-K^2\pm i\varepsilon}d\vec{K}. \tag{8.107}$$

Let (K,Θ,Φ) be the spherical coordinates of $\vec{K}$ and let the z-axis be along $\vec{R} = \vec{x} - \vec{x}'$. Then

$$G_0^{\pm}(k,\vec{R}) = \frac{1}{(2\pi)^3}\int_0^{\infty} dK K^2 \int_0^{\pi} \mathrm{d}\Theta \sin\Theta \int_0^{2\pi} \mathrm{d}\Phi \frac{e^{iKR\cos\Theta}}{k^2-K^2\pm i\varepsilon}. \tag{8.108}$$

Performing the angular integrations and observing that the integrand is an even function of K we can extend the integral from $-\infty$ to $+\infty$ and obtain

$$G_0^{\pm}(k,\vec{R}) = \frac{1}{8\pi^2 iR}\int_{-\infty}^{+\infty} \frac{K(e^{iKR}-e^{-iKR}))}{k^2-K^2\pm i\varepsilon}dK. \tag{8.109}$$

With the partial fraction decomposition $\frac{1}{k^2-K^2} = -\frac{1}{2K}\left(\frac{1}{K-k}+\frac{1}{K+k}\right)$ we can split the integral into two parts

$$G_0(k,R) = \frac{i}{16\pi^2 R}(I_1 - I_2), \tag{8.110}$$

with

$$I_1 = \int_{-\infty}^{+\infty} e^{iKR}\left(\frac{1}{K-k}+\frac{1}{K+k}\right)\mathrm{d}K \tag{8.111}$$

$$I_2 = \int_{-\infty}^{+\infty} e^{-iKR}\left(\frac{1}{K-k}+\frac{1}{K+k}\right)\mathrm{d}K \tag{8.112}$$

The integrals can now be evaluated using the Cauchy integral formula if we close the integration path with a half-circle in the upper or lower complex half-plane, respectively, so that the contribution from the arcs at infinity vanish. The ambiguity of the Green's function arises from different choices of the integration about the poles of the integrand on the real axis, and different pole prescriptions obviously differ by terms localized at $K^2 = k^2$ and hence by a

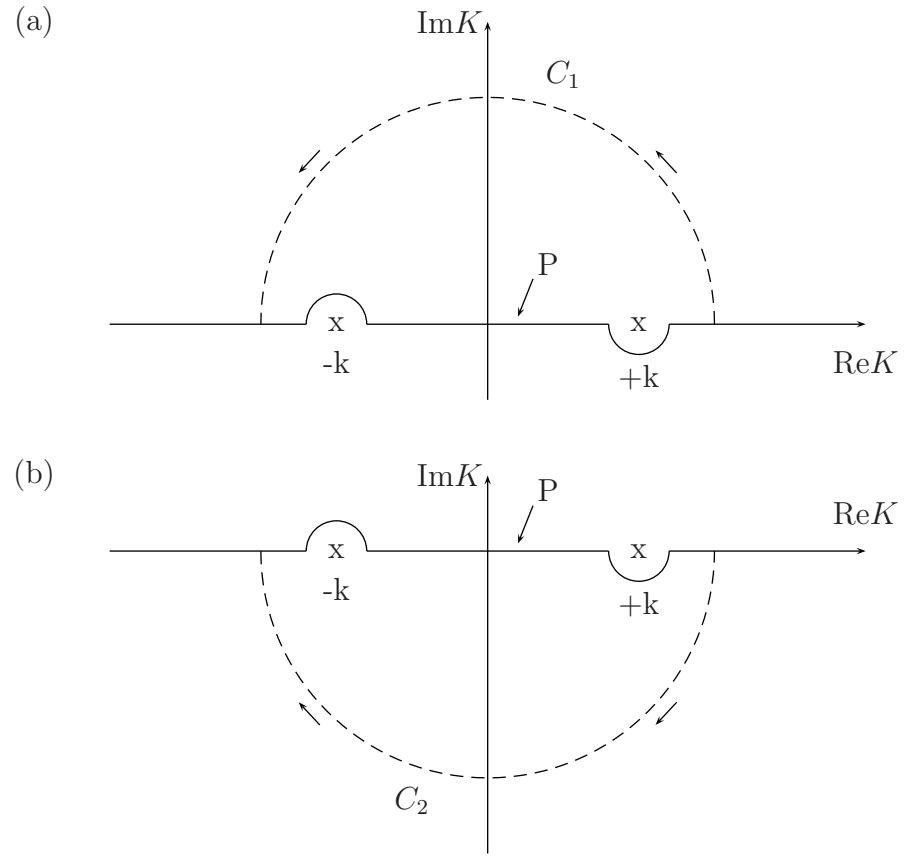


Figure 8.3: (a) The contour (P+C_1) for calculating the integral I_1 by avoiding the poles $K = \pm k$ and closing via a semi-circle in infinity. (b) the contour for calculating the integral I_2.

superposition of plane wave solutions to the homogeneous equation. The integration contour in the complex K-plane shown in fig. 8.3 corresponds to a small positive imaginary part of k and hence to G_0^+. Since e^{iKR} vanishes on C_1 and e^{-iKR} vanishes on C_2 we find

$$I_1 = 2\pi i e^{ikR} \tag{8.113}$$

$$I_2 = -2\pi i e^{ikR} \tag{8.114}$$

With a similar calculation for $k \to k - i\varepsilon$ the Green's function in the original variables $\vec{x}$ and $\vec{x}'$ becomes

$$G_0^{\pm}(k, \vec{x}, \vec{x}') = -\frac{1}{4\pi} \frac{\pm e^{ik|\vec{x}-\vec{x}'|}}{|\vec{x} - \vec{x}'|}. \tag{8.115}$$

so that $G_0^+ = G_0^{ret}$ corresponds to retarded boundary conditions. With $U = \frac{2m}{\hbar^2} V$ we can now write the integral equation for the wave function as

$$\boxed{u_{\vec{k}}(\vec{x}) = e^{i\vec{k}\cdot\vec{x}} - \frac{m}{2\pi\hbar^2} \int \frac{e^{ik|\vec{x}-\vec{x}'|}}{|\vec{x} - \vec{x}'|} V(\vec{x}') u_{\vec{k}}(\vec{x}') \mathrm{d}\vec{x}'.} \tag{8.116}$$

This integral equation is known as the *Lippmann-Schwinger equation* for potential scattering. It is equivalent to the Schrödinger equation *plus* the scattering boundary condition (8.20).

We can now relate this integral representation to the scattering amplitude by considering the situation where the distance of the detector $r \to \infty$ is much larger than the range of the potential to which the integration variable $\vec{x}'$ is essentially confined so that $r' \ll r$. Hence

$$|\vec{x} - \vec{x}'| = \sqrt{r^2 - 2\vec{x}\vec{x}' + r'^2} = r - \frac{\vec{x}\vec{x}'}{r} + \mathcal{O}(\frac{1}{r}). \tag{8.117}$$

Since $\vec{x}$ points in the same direction (θ, φ) as the wave vector $\vec{k}'$ of the scattered particles we have $\vec{k}' = k\vec{x}/r$ for elastic scattering and hence

$$\frac{e^{ik|\vec{x}-\vec{x}'|}}{|\vec{x} - \vec{x}'|} \xrightarrow[r\to\infty]{} \frac{e^{ikr}}{r} e^{-i\vec{k}'\cdot\vec{x}'} + \cdots, \tag{8.118}$$

where terms of order in $1/r^2$ have been neglected. Substituting this expansion into the Lippmann-Schwinger equation we find

$$u_{\vec{k}}(\vec{x}) \xrightarrow[r\to\infty]{} e^{i\vec{k}\cdot\vec{x}} - \frac{1}{4\pi}\frac{e^{ikr}}{r}\int e^{-i\vec{k}'\cdot\vec{x}'} U(\vec{x}') u_{\vec{k}}(\vec{x}') \mathrm{d}\vec{x}'. \tag{8.119}$$

Comparing with the ansatz (8.20) we thus obtain the integral representation

$$\begin{aligned} f(k,\theta,\phi) &= -\frac{1}{4\pi}\int e^{-i\vec{k}'\cdot\vec{x}} U(\vec{x}) u_{\vec{k}}(\vec{x}) \mathrm{d}\vec{x} \\ &= -\frac{1}{4\pi}\langle \phi_{\vec{k}'} | U | u_{\vec{k}} \rangle = -\frac{m}{2\pi\hbar^2}\langle \phi_{\vec{k}'} | V | u_{\vec{k}} \rangle \end{aligned} \tag{8.120}$$

for the scattering amplitude, where $\langle \phi_{\vec{k}'}| = e^{-i\vec{k}'\cdot\vec{x}}$ and $|\phi_{\vec{k}'}\rangle = e^{i\vec{k}'\cdot\vec{x}} = (2\pi)^{3/2}|k'\rangle$.

8.4 The Born series

The *Born series* is the iterative solution of the Lippmann-Schwinger equation by the ansatz

$$u(\vec{x}) = \sum_{n=0}^{\infty} u_n(\vec{x}) \qquad \text{for} \qquad u_0(\vec{x}) = \phi_{\vec{k}}(\vec{x}) = e^{i\vec{k}\cdot\vec{x}}, \tag{8.121}$$

which yields

$$u_1(\vec{x}) = \int G_0^+(k, \vec{x}, \vec{x}') U(\vec{x}') u_0(\vec{x}')\, d\vec{x}', \tag{8.122}$$

$$\vdots$$

$$u_n(\vec{x}) = \int G_0^+(k, \vec{x}, \vec{x}') U(\vec{x}') u_{n-1}(\vec{x}') \mathrm{d}\vec{x}', \tag{8.123}$$

so that the n^{th} term u_n is formally of order $\mathcal{O}(V^n)$. It usually converges well for weak potentials or at high energies. Insertion of the Born series into of the formula (8.120) yields

$$f = -\frac{1}{4\pi}\langle \phi_{\vec{k}'} | U + UG_0^+U + UG_0^+UG_0^+U + \cdots | \phi_{\vec{k}} \rangle \tag{8.124}$$

and keeping only the first term we obtain the (first) *Born approximation*

$$f^B = -\frac{1}{4\pi}\langle\phi_{\vec{k}'}|U|\phi_{\vec{k}}\rangle. \tag{8.125}$$

to the scattering amplitude.

Phase shift in Born approximation. The Lippmann Schwinger equation (8.116) can also be analysed using partial waves. We assume that our potential is centrally symmetric and expand the scattering wave function $u_{\vec{k}}$ in Legendre polynomials (see equation (8.26)). With the normalisation

$$R_l(k,r) \xrightarrow[r\to\infty]{} j_l(kr) - \tan\delta_l(k) n_l(kr) \tag{8.126}$$

$$\xrightarrow[r\to\infty]{} \frac{1}{kr}\left[\sin\left(kr - \frac{l\pi}{2}\right) + \tan\delta_l(k)\cos\left(kr - \frac{l\pi}{2}\right)\right], \tag{8.127}$$

we find that each radial function satisfies the radial integral equation

$$R_l(k,r) = j_l(kr) + \int_0^\infty G_l(k,r,r')U(r')R_l(k,r')r'^2\mathrm{d}r', \tag{8.128}$$

where

$$G_l = k j_l(kr_<) n_l(kr_>) \qquad \text{with} \qquad r_< \equiv \min(r,r') \quad \text{and} \quad r_> \equiv \max(r,r') \tag{8.129}$$

is the partial wave contribution to the Green's function

$$\frac{e^{ik|\vec{x}-\vec{x}'|}}{|\vec{x}-\vec{x}'|} = ik\sum_{l=0}^{\infty}(2l+1)\, j_l(kr_<)\Big(j_l(kr_>) + i n_l(kr_>)\Big)P_l(\cos\theta). \tag{8.130}$$

We solve this equation by iteration, starting with $R_l^{(0)}(k,r) = j_l(kr)$. When we analyse equation (8.128) for $r \to \infty$ we obtain the integral representation

$$\tan\delta_l(k) = -k\int_0^\infty j_l(kr)U(r)R_l(k,r)r^2\mathrm{d}r. \tag{8.131}$$

Substituting the iteration for R_l into the integral equation yields a Born series whose first term

$$\boxed{\tan\delta_l^B(k) = -k\int_0^\infty [j_l(kr)]^2 U(r) r^2\mathrm{d}r.} \tag{8.132}$$

is the *first Born approximation to* $\tan\delta_l$.

Total scattering cross section in first Born approximation. With the *wave vector transfer*

$$\vec{q} = \vec{k} - \vec{k}' \tag{8.133}$$

the first Born approximation of the scattering amplitude can be written as the Fourier transform

$$f^B = -\frac{1}{4\pi}\int e^{i\vec{q}\cdot\vec{x}}U(\vec{x})\mathrm{d}\vec{x} \tag{8.134}$$

of the potential. For elastic scattering with $k = k'$ and $\vec{k}\cdot\vec{k}' = k^2\cos\theta$ we find

$$q = 2k\sin\frac{\theta}{2}, \tag{8.135}$$

with θ being the scattering angle. For a central potential it is now useful to introduce polar coordinates with angles (α, β) such that $\vec{q}$ is the polar axis. We thus find that

$$\begin{aligned} f^B(q) &= -\frac{1}{4\pi}\int_0^\infty \mathrm{d}r r^2 U(r)\int_0^\pi \mathrm{d}\alpha\sin\alpha\int_0^{2\pi}\mathrm{d}\beta e^{iqr\cos\alpha} \\ &= -\frac{1}{2}\int_0^\infty \mathrm{d}r r^2 U(r)\int_{-1}^{+1}\mathrm{d}(\cos\alpha)e^{iqr\cos\alpha} \\ &= -\frac{1}{q}\int_0^\infty r\sin(qr)U(r)\mathrm{d}r \end{aligned} \tag{8.136}$$

only depends on $q(k,\theta)$. The total cross-section in the first Born approximation hence becomes

$$\boxed{\sigma^B_{\mathrm{tot}}(k) = 2\pi\int_0^\pi |f^B(q)|^2\sin\theta\mathrm{d}\theta = \frac{2\pi}{k^2}\int_0^{2k}|f^B(q)|^2 q\mathrm{d}q} \tag{8.137}$$

where we used the differential $dq = k\cos\frac{\theta}{2}d\theta$ of (8.135) and $\sin\theta\, d\theta = 2\sin\frac{\theta}{2}\,\cos\frac{\theta}{2}\,d\theta = \frac{q}{k}\frac{dq}{k}$.

8.4.1 Application: Coulomb scattering and the Yukawa potential

Since the Coulomb potential has infinite range we apply the Born approximation to the *Yukawa potential*

$$U(r) = C\frac{e^{-\alpha r}}{r} = C\frac{e^{-r/a}}{r} \qquad \text{with} \qquad a = \alpha^{-1}, \tag{8.138}$$

which can be regarded as a screened Colomb potential. At the end of the calculation we can then try to send the screening length $a \to \infty$. For the Born approximation (8.136) we obtain

$$f^B = -\frac{1}{q}\int_0^\infty r\sin(qr)\frac{C}{r}e^{-\alpha r}dr = -\frac{C}{q}\,\mathrm{Im}\int_0^\infty e^{iqr-\alpha r}dr = -\frac{C}{q}\,\mathrm{Im}\,\frac{1}{\alpha - iq} = -\frac{C}{\alpha^2+q^2} \tag{8.139}$$

and the corresponding differential cross section

$$\frac{\mathrm{d}\sigma^B}{\mathrm{d}\Omega} = \frac{C^2}{(\alpha^2+q^2)^2} \tag{8.140}$$

of the Yukawa potential.

The Coulomb potential. The electrostatic force between charges Q_A and Q_B corresponds to the potential

$$V_{Coulomb}(r) = \frac{Q_A Q_B}{4\pi\varepsilon_0}\frac{1}{r} \tag{8.141}$$

which corresponds to

$$C = \frac{2m}{\hbar^2}\frac{Q_A Q_B}{4\pi\varepsilon_0} \tag{8.142}$$

but obviously violates the finite range condition. Nevertheless, there is a finite limit $\alpha \to 0$ for which we obtain the scattering amplitude $f^B = -C/q^2$ and the differential cross-section in first Born approximation as

$$\frac{\mathrm{d}\sigma_c^B}{\mathrm{d}\Omega} = \frac{C^2}{q^4} = \left(\frac{\gamma}{2k}\right)^2 \frac{1}{\sin^4(\theta/2)} = \left(\frac{Q_A Q_B}{4\pi\varepsilon_0}\right)^2 \frac{1}{16E^2\sin^4(\theta/2)} \tag{8.143}$$

where

$$\gamma = \frac{Q_A Q_B}{(4\pi\varepsilon_0)\hbar v} = \frac{C}{2k} \tag{8.144}$$

is a dimensionless quantity.

- This result for the differential cross-section for scattering by a Coulomb potential is identical with the formula that *Rutherford* obtained 1911 by using classical mechanics.
- The *exact* quantum mechanical treatment of the Coulomb potential yields the same result for the differential cross-section. The scattering amplitude f_c however differs by a phase factor. It can be shown that

  $$f_c = -\frac{\gamma}{2k\sin^2(\theta/2)}\frac{\Gamma(1+i\gamma)}{\Gamma(1-i\gamma)}e^{-i\gamma\log[\sin^2(\theta/2)]} \tag{8.145}$$

  where Γ denotes the Gamma-function [Hittmair].
- The Rutherford differential cross-section scales with the energy E at all angles by the factor $(Q_A Q_B/16\pi\varepsilon_0 E)^2$ so that the angular distribution is independent of the energy.
- The phase correction in (8.145) becomes observable in the scattering of identical particles due to interference terms. This will be discussed in chapter 10.

8.5 Wave operator, transition operator and S-matrix

In this section we introduce the scattering matrix S and relate it to the scattering amplitude via the transition matrix T. We start with the observation that the Greens function $G_0^{\pm}$ can be interpreted as the inverse of $E - H_0 \pm i\varepsilon$ up to a factor $\frac{2m}{\hbar^2}$. Indeed, with $k = \sqrt{\frac{2mE}{\hbar^2}}$ and the free Hamiltonian $H_0 = -\frac{\hbar^2}{2m}\Delta$ we find for a momentum eigenstate with $\vec{p}|K\rangle = \hbar\vec{K}|K\rangle$ that

$$(E - H_0)|K\rangle = \frac{\hbar^2}{2m}(k^2 - K^2)|K\rangle \tag{8.146}$$

so that

$$\lim_{\varepsilon\to 0}\ \frac{1}{E - H_0 \pm i\varepsilon} = \frac{2m}{\hbar^2}\, G_0^{\pm} \tag{8.147}$$

follows by Fourier transformation and regularization with a small imaginary part of the energy. More explicitly, the matrix elements of the operator $(E - H_0 \pm i\varepsilon)^{-1}$ in position space are

$$\langle x|\frac{1}{E - H_0 \pm i\varepsilon}|x'\rangle = \langle x|\frac{1}{E - H_0 \pm i\varepsilon}\int d^3K|K\rangle\langle K|x'\rangle \tag{8.148}$$

$$= \int d^3K\ \langle x|\frac{1}{\frac{\hbar^2}{2m}(k^2 - K^2) \pm i\varepsilon}|K\rangle\ \frac{e^{-i\vec{K}\vec{x}'}}{(2\pi)^{3/2}} \tag{8.149}$$

$$= \frac{2m}{\hbar^2}\int \frac{d^3K}{(2\pi)^3}\frac{e^{i\vec{K}(\vec{x}-\vec{x}')}}{k^2 - K^2 \pm i\varepsilon} = \frac{2m}{\hbar^2}G_0^{\pm}(\vec{x} - \vec{x}'). \tag{8.150}$$

With $z = E \pm i\varepsilon$ this is proportional to the resolvent $R_z(H_0) = (H_0 - z)^{-1}$ of H_0, which is a bounded operator for $\varepsilon > 0$. The Lippmann–Schwinger equation can now be written as

$$|u_\pm\rangle = |u_0\rangle + \frac{1}{E - H_0 \pm i\varepsilon}V|u_\pm\rangle \tag{8.151}$$

where $|u_+\rangle$ corresponds to the scattering solution with retarded boundary conditions.

Wave operator and transition matrix. The Born series for the solutions of (8.151) is

$$|u_\pm\rangle = |u_0\rangle + \sum_{n=1}^{\infty}(\frac{1}{E - H_0 \pm i\varepsilon}V)^n|u_0\rangle \tag{8.152}$$

In order to sum up the geometric operator series we use the matrix formula

$$1 + \tfrac{1}{A}V + (\tfrac{1}{A}V)^2 + \ldots = (1 - \tfrac{1}{A}V)^{-1} = (\tfrac{1}{A}(A - V))^{-1} = (A - V)^{-1}A$$
$$= \tfrac{1}{A-V}(A - V + V) = 1 + \tfrac{1}{A-V}V \tag{8.153}$$

for $A = E - H_0 \pm i\varepsilon$. Since $A - V = E - H \pm i\varepsilon$ with $H = H_0 + V$ the Born series can thus be summed up in terms of the resolvent of the full Hamiltonian

$$|u_\pm\rangle = |u_0\rangle + \frac{1}{E - H \pm i\varepsilon}V|u_0\rangle = \Omega_\pm|u_0\rangle, \tag{8.154}$$

where we introduced the *wave operator* or *Møller operator*

$$\Omega_\pm = 1 + \frac{1}{E - H \pm i\varepsilon}V \tag{8.155}$$

which maps plane waves $|u_0\rangle$ to exact stationary scattering solution $|u_0\rangle \to |u_\pm\rangle = \Omega_\pm|u_0\rangle$. If we insert this representation for the scattering solution into the formula (8.120) we need to be careful about the normalization of the wave function. In the present section we prefer to work with momentum eigenstates normalized as $\langle\vec{k}'|\vec{k}\rangle = \delta^3(\vec{k}' - \vec{k})$ which yield a factor $(2\pi)^3$ as compared to the plane waves $e^{\pm ikx}$ with normalized amplitude $|\phi_k(x)| = 1$. We hence obtain

$$f = -2\pi^2\langle\vec{k}'|U|u_+\rangle = -4\pi^2\frac{m}{\hbar^2}\ \langle\vec{k}'|V\Omega_+|k\rangle \tag{8.156}$$

for the scattering amplitude, which suggests to define the transition operator T as

$$T := V\Omega_+ = V(1 + \tfrac{1}{E-H+i\varepsilon}V) = (1 + V\tfrac{1}{E-H+i\varepsilon})V = \Omega_-^\dagger V \tag{8.157}$$

so that

$$f(k,\theta,\varphi) = -4\pi^2 \frac{m}{\hbar^2} \langle k'|T|k\rangle \tag{8.158}$$

with (θ,φ) corresponding to the direction of $\vec{k}'$.

The S–matrix. The idea behind the definition of the scattering matrix in terms of the wave operator is that the incoming scattering state $\Omega_+|k\rangle$ is reduced by the measurement in the detector to a state that is a plane wave in the asymptotic future and hence, as an exact solution to the Schrödinger equation, corresponds to advanced boundary conditions $\Omega_-|k'\rangle$. Since $(\Omega_-|k'\rangle)^\dagger = \langle k'|\Omega_-^\dagger$ the scattering amplitude should correspond to the matrix element $\langle k'|S|k\rangle$ in the momentum eigenstate basis. Hence we define

$$\langle k'|S|k\rangle = \langle k'|\Omega_-^\dagger\Omega_+|k\rangle \qquad \Rightarrow \qquad S = \Omega_-^\dagger\Omega_+. \tag{8.159}$$

It can be shown that this definition of the S-matrix agrees with the limit

$$S = \lim_{\substack{t_1\to+\infty \\ t_0\to-\infty}} U_I(t_1,t_0) \tag{8.160}$$

of the time evolution operator in the interaction picture [Hittmair], which implies unitarity.

Unitarity of the S-matrix. We first prove that $\Omega_\pm$ are isometries, i.e. that the wave operators preserve scalar products. It will then be easy to directly show that $SS^\dagger = S^\dagger S = \mathbb{1}$. For this we introduce a more abstract notation with a complete orthonormal basis $\langle u_i^0|u_j^0\rangle = \delta_{ij}$ of free energy eigenstates $|u_i^0\rangle$ with $H_0|u_i^0\rangle = E_i|u_i^0\rangle$ and the corresponding exact solutions

$$|u_i^\pm\rangle = \Omega_\pm|u_i^0\rangle, \qquad H|u_i^\pm\rangle = E_i|u_i^\pm\rangle, \tag{8.161}$$

for which we compute

$$\langle u_i^+|u_j^+\rangle = \langle u_i^+|u_j^0\rangle + \langle u_i^+|\frac{1}{E_j - H + i\varepsilon}V|u_j^0\rangle \tag{8.162}$$

$$= \langle u_i^+|u_j^0\rangle + \frac{1}{E_j - E_i + i\varepsilon}\langle u_i^+|V|u_j^0\rangle. \tag{8.163}$$

Hermitian conjugation of the Lippmann–Schwinger equation implies, on the other hand,

$$\langle u_i^+|u_j^0\rangle = \langle u_i^0|u_j^0\rangle + \langle u_i^+|V\frac{1}{E_i - H_0 - i\varepsilon}|u_j^0\rangle \tag{8.164}$$

$$= \langle u_i^0|u_j^0\rangle + \frac{1}{E_i - E_j - i\varepsilon}\langle u_i^+|V|u_j^0\rangle \tag{8.165}$$

Hence $\langle u_i^+|u_j^+\rangle = \langle u_i^0|u_j^0\rangle = \delta_{ij}$ and by complex conjugation $\langle u_i^-|u_j^-\rangle = \delta_{ij}$.

In contrast to the finite-dimensional situation the isometry property of $\Omega_\pm$ does not imply unitarity because an isometry in an infinite-dimensional Hilbert space does not need to be surjective. Indeed, the maps $\Omega_\pm$ send plane waves, which form a complete system, to scattering

states, which are not complete if the potential V supports bound states. More explicitly, we can write

$$\Omega_\pm = \Omega_\pm \sum_i |u_i^0\rangle\langle u_i^0| = \sum_i |u_i^\pm\rangle\langle u_i^0|. \tag{8.166}$$

Hence

$$\Omega_\pm^\dagger \Omega_\pm = \sum_{ij} |u_i^0\rangle\langle u_i^\pm|u_j^\pm\rangle\langle u_j^0| = \sum_{ij} |u_i^0\rangle\delta_{ij}\langle u_j^0| = \mathbb{1} \tag{8.167}$$

$$\Omega_\pm \Omega_\pm^\dagger = \sum_{ij} |u_i^\pm\rangle\langle u_i^0|u_j^0\rangle\langle u_j^\pm| = \sum_i |u_i^\pm\rangle\langle u_i^\pm| = \mathbb{1} - P_{b.s.} \tag{8.168}$$

where $P_{b.s.}$ is the projector to the bound states. If the potential V has negative energy solutions these states cannot be produced in a scattering process and are hence missing from the completeness relation in the last sum. Combining these results unitarity of the S matrix

$$S^\dagger S = \Omega_+^\dagger \Omega_- \Omega_-^\dagger \Omega_+ = \Omega_+^\dagger (\mathbb{1} - P_{b.s.})\Omega_+ = \Omega_+^\dagger \Omega_+ = \mathbb{1} \tag{8.169}$$

$$S S^\dagger = \Omega_-^\dagger \Omega_+ \Omega_+^\dagger \Omega_- = \Omega_-^\dagger (\mathbb{1} - P_{b.s.})\Omega_- = \Omega_-^\dagger \Omega_- = \mathbb{1} \tag{8.170}$$

is established.

Relating the S-matrix to the transition matrix. In order to derive the relation between S and T we and write the S-matrix elements S_{ij} as

$$S_{ij} = \langle u_i^-|u_j^+\rangle = \langle u_i^+|u_j^+\rangle + (\langle u_i^-| - \langle u_i^+|)\,|u_j^+\rangle. \tag{8.171}$$

With $\langle u_i^+|u_j^+\rangle = \delta_{ij}$ and

$$\langle u_i^-| = \langle u_i^0|\Omega_-^\dagger = \langle u_i^0|(1 + V\frac{1}{E_i - H + i\varepsilon}), \tag{8.172}$$

$$\langle u_i^+| = \langle u_i^0|\Omega_+^\dagger = \langle u_i^0|(1 + V\frac{1}{E_i - H - i\varepsilon}). \tag{8.173}$$

we obtain

$$S_{ij} = \delta_{ij} + \langle u_i^0|V(\frac{1}{E_i - H + i\varepsilon} - \frac{1}{E_i - H - i\varepsilon})|u_j^+\rangle. \tag{8.174}$$

Since $H|u_j^+\rangle = E_j|u_j^+\rangle$ and

$$\lim_{\varepsilon\to 0}(\frac{1}{z - i\varepsilon} - \frac{1}{z + i\varepsilon}) = 2\pi i\,\delta(z) \tag{8.175}$$

we conclude $S_{ij} = \delta_{ij} - 2\pi i\delta(E_i - E_j)\langle u_i^0|V|u_j^+\rangle$ and hence

$$\boxed{S_{ij} = \delta_{ij} - 2\pi i\delta(E_i - E_j)T_{ij}.} \tag{8.176}$$

The non-relativistic dispersion $E = (\hbar k)^2/2m$ implies $\delta(E_i - E_j) = \frac{m}{\hbar^2 k}\delta(k_i - k_j)$ so that

$$\langle k'|S|k\rangle = \delta^3(\vec{k} - \vec{k}') + \frac{i}{2\pi k}\delta(k - k')f(\vec{k}', \vec{k}). \tag{8.177}$$

Partial wave decomposition on the energy shell [Hittmair] then leads to $S_l = e^{2i\delta_l} = 1 - 2\pi i T_l$.

Chapter 9

Symmetries and transformation groups

When contemporaries of Galilei argued against the heliocentric world view by pointing out that we do not feel like rotating with high velocity around the sun he argued that a uniform motion cannot be recognized because the laws of nature that govern our environment are invariant under what we now call a Galilei transformation between inertial systems. Invariance arguments have since played an increasing role in physics both for conceptual and practical reasons.

In the early 20th century the mathematician Emmy Noether discovered that energy conservation, which played a central role in 19th century physics, is just a special case of a more general relation between symmetries and conservation laws. In particular, energy and momentum conservation are equivalent to invariance under translations of time and space, respectively. At about the same time Einstein discovered that gravity curves space-time so that space-time is in general not translation invariant. As a consequence, energy is not conserved in cosmology.

In the present chapter we discuss the symmetries of non-relativistic and relativistic kinematics, which derive from the geometrical symmetries of Euclidean space and Minkowski space, respectively. We decompose transformation groups into discrete and continuous parts and study the infinitesimal form of the latter. We then discuss the transition from classical to quantum mechanics and use rotations to prove the Wigner-Eckhart theorem for matrix elements of tensor operators. After discussing the discrete symmetries parity, time reversal and charge conjugation we conclude with the implications of gauge invariance in the Aharonov–Bohm effect.

9.1 Transformation groups

Newtonian mechanics in Euclidean space is invariant under the transformations

$$g_v(t,\vec{x}) = (t, \vec{x} + \vec{v}t) \qquad \text{Galilei transformation,} \tag{9.1}$$

$$g_{\tau,\vec{\xi}}(t,\vec{x}) = (t+\tau, \vec{x} + \vec{\xi}) \qquad \text{time and space translation,} \tag{9.2}$$

$$g_{\mathcal{O}}(t,\vec{x}) = (t, \mathcal{O}\vec{x}) \qquad \text{rotation or orthogonal transformation,} \tag{9.3}$$

where $\mathcal{O}$ is an orthogonal matrix $\mathcal{O}\cdot\mathcal{O}^T = \mathbb{1}$. In special relativity the structure of the invariance group unifies to translations $x^\mu \to x^\mu + \xi^\mu$ and Lorentz transformations

$$x^\mu \to L^\mu{}_\nu x^\nu, \qquad L^T g L = g \quad \text{with} \quad g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{9.4}$$

which leave $x^2 = x^\mu x_\mu$ invariant. A transformation under which the equations of motion of a classical system are invariant is called a symmetry. Transformations, and in particular symmetry transformations, are often invertible and hence form a group under composition.[1]

Infinitesimal transformations. For continuous groups, whose elements depend continuously on one or more parameters, it is useful to consider infinitesimal transformations. Invariance under *infinitesimal* translations, for example, implies invariance under *all translations*. For the group $O(3) \equiv O(3,\mathbb{R})$ of real orthogonal transformations in 3 dimensions this is, however, not true because $\mathcal{O}\cdot\mathcal{O}^T = \mathbb{1}$ only implies $\det\mathcal{O} = \pm 1$ and transformations with a negative determinant $\det\mathcal{O} = -1$, which change of orientation, can never be reached in a continuous process by composing small transformations with determinant $+1$. The orthogonal group $\mathcal{O}(3)$ hence decomposes into two connected parts and its subgroup $SO(3,\mathbb{R})$ of *special* orthogonal matrices R (*special* means the restriction to $\det R = 1$) is the component that contains the identity. In three dimensions every special orthogonal matrix corresponds to a rotation $R_{\vec{\alpha}}$ about a fixed axis with direction $\vec{\alpha}$ by an angle $|\alpha|$ for some vector in $\vec{\alpha} \in \mathbb{R}^3$. Obviously any such rotation can be obtained by a large number of small rotations so that

$$R_{\vec{\alpha}} = (R_{\frac{1}{n}\vec{\alpha}})^n = \lim_{n\to\infty}(\mathbb{1} + \tfrac{1}{n}\delta R_{\vec{\alpha}})^n = \exp(\delta R_{\vec{\alpha}}) \tag{9.5}$$

where we introduced the infinitesimal rotations

$$\delta R(\vec{\alpha})x_i = \varepsilon_{ijk}\alpha_j x_k = \delta R(\vec{\alpha})_{ik}x_k, \qquad \delta R(\vec{\alpha})_{ik} = \varepsilon_{ijk}\alpha_j. \tag{9.6}$$

Like for the derivative f' of a function f in differential calculus, an infinitesimal transformation δT is the linear term in the expansion $T(\varepsilon\alpha) = \mathbb{1} + \varepsilon\delta T(\alpha) + \mathcal{O}(\varepsilon^2)$ and hence is linear and

[1] Since translations $g_{\vec{\xi}}$ and orthogonal transformations $g_{\mathcal{O}}$ do not commute they generate the Euclidean group $E(3)$ as a *semidirect product*, each of whose elements can be written uniquely as a composition $g_{\mathcal{O}} \circ g_{\vec{\xi}}$. Lorentz transformations and translations in Minkowski space similarly generate the *Poincaré group*.

obeys the Leibniz rule for products and the chain rule for functions,[2]

$$\delta T\,(f\cdot g)=\delta T(f)\cdot g+f\cdot\delta T(g),\qquad \delta T(f(x))=\frac{df(x)}{dx^i}\,\delta T(x^i) \tag{9.7}$$

In accord with (9.6) the infinitesimal form δR of an orthogonal transformation $RR^T=\mathbb{1}$ is given by an antisymmetric matrix since $(\mathbb{1}+\varepsilon\delta R)(\mathbb{1}+\varepsilon\delta R^T)=\mathbb{1}+\varepsilon(\delta R+\delta R^T)+\mathcal{O}(\varepsilon^2)$. Similarly, the infinitesimal form $\delta U=iH$ of a unitary transformation $UU^\dagger=\mathbb{1}$ is antihermitian

$$U=\mathbb{1}+\varepsilon\,iH+\mathcal{O}(\varepsilon^2),\qquad UU^\dagger=\mathbb{1}\quad\Rightarrow H=H^\dagger. \tag{9.8}$$

In turn, $\exp(iH)$ is unitary if H is Hermitian. The advantage of infitesimal transformations is that they just add up for combined transformations,

$$T_1T_2=\mathbb{1}+\varepsilon(\delta T_1+\delta T_2)+\mathcal{O}(\varepsilon^2). \tag{9.9}$$

In particular, an infinitesimal rotation about an arbitrary axis $\vec\alpha$ can be written as a linear combination of infinitesimal rotations about the coordinate axes

$$\delta R(\vec\alpha)=\alpha_j\delta R_j,\qquad (\delta R_j)_{ik}=\delta R(\vec e_j)_{ik}=\varepsilon_{ijk}. \tag{9.10}$$

Since the finite transformations are recovered by exponentiation the Baker–Campbell–Hausdorff formula

$$e^A\,e^B=e^{A+B+\frac12[A,B]+\frac{1}{12}([A,[A,B]]-[B,[A,B]])+\text{ multiple commutators}} \tag{9.11}$$

shows that a nonabelian group structure of finite transformations corresponds to nonvanishing commutators of the infinitesimal transformations. In the following an infintesimal transformation will not always be indicated by a variation symbol, but it should be clear from the context which transformations are finite and which are infinitesimal.

Discrete transformations As we observed for the orthogonal group, invariance under infinitesimal transformations only implies invariance for the connected part of a transformation group and a number of discrete "large" transformations, which cannot be obtained by combining many small transformations, may have to be investigated seperately. In nonrelativistic mechanics the relavant transformations are time reversal $T: t\to -t$ and parity $P:\vec x\to-\vec x$, which is equivalent to a reflecion $\vec x\to\vec x-2\vec n(\vec x\cdot\vec n)$ at a mirror with normalized orthogonal vector $\vec n$ combined with a rotation $R(\pi\vec n)$ about $\vec n$ by the angle π. In 1956 T.D. Lee and C.N. Yang came up with the idea that an apparent problem with parity selection rules in neutral kaon decay might be due to violation of parity in weak interactions and they suggested a number of experiments for testing the conservation of parity in weak processes. By the end of that year

[2] Linear transformations obeying the Leibniz rule on some associative algebra, like the commutative algebra of functions in classical mechanics or the noncommutative algebras of matrices or operators in quantum mechanics are called *derivations*.

Madame C.S. Wu and collaborators observed the first experimental signs of parity violation in β-decay of polarized ^{60}Co. This experimental result came as a great surprise because parity selection rules had become a standard tool in atomic physics and parity conservation was also well established for strong interactions.

In the relativistic theory there is another discrete transformation, called charge conjugation, which amounts to the exchange of particles and anti-particles. The combination CP of parity and charge conjugation turns out to be even more natural than parity alone, and CP is indeed conserved in many weak processes. But in 1964 it was discovered that CP is also violated in the neutral kaon system.[3] In 1967 Sakharov showed that CP-violation, in addition to thermal non-equilibrium and the existence of baryon number violating processes, is one of the three conditions for the possibility of creating matter in the universe. Invariance under the combination CPT of all three discrete transformations of relativistic kinematics, can be shown to follow from basic axioms of quantum field theory, and indeed no CPT violation has ever been observed.

Active and passive transformations. A transformation like, for example, a translation $\vec{x} \to \vec{x}' = \vec{x} + \vec{\xi}$ can be interpreted in two different ways. On the one hand, we can think of it as a motion where a particle located at the position $\vec{x}$ is moved to the position with coordinates $\vec{x}'$ with respect to some fixed frame of reference. Such a motion is often called an *active transformation*. On the other hand we can leave everything in place and describe the same physical process in terms of new coordinates x'. The resulting coordinate transformation is often called a *passive transformation*. Active and passive transformation are mathematically equivalent in the sense that the formulas look identical. If we physically move our experiment to a new lab, however, our instruments may be sufficiently sensitive to detect the change of the magnetic field of the earth or of other environmental parameters that are not moved in an active transformation of the experiment. If we also move the earth and its magnetic field, then it is most likely that we are still in our old lab and that all that happened was a change of coordinates.

If we simultaneously perform an active and a passive transformation then a scalar quantity like a wave function $\psi(x)$ does not change its form so that

$$\psi(x) = \psi'(x'), \qquad x' = Rx \quad \Rightarrow \quad \psi'(x) = \psi(R^{-1}x). \tag{9.12}$$

Quantum mechanics has its own way of incorporating this relation into its formalism. Since a symmetry transformation has to preserve scalar products we consider unitary transformations

[3] Since quarks have both weak and strong interactions, it is still mysterious why CP violation does not also affect the strong interactions. This is known as the *strong CP problem*. Its only proposed explanation so far has been the Peccei-Quinn symmetry, which postulates a new particle called axion. If they exist, axions might contribute to the observed dark matter in the universe.

$R^\dagger = R^{-1}$ in Hilbert space. Hence

$$(R\psi)(x) = \langle x|R|\psi\rangle = \langle R^\dagger x|\psi\rangle = \langle R^{-1}x|\psi\rangle = \psi(R^{-1}x), \tag{9.13}$$

in accord with (9.12). For discrete symmetries also anti-unitary maps are possible, as will be the case for time reversal and charge conjugation.

9.2 Noether theorem and quantization

Canonical mechanics. For a dynamical system with *Lagrange function* $L = L(q^i, \dot{q}^i, t)$ *Hamilton's principle of least action* states that the functional

$$\phi(\gamma) = \int_{t_0}^{t_1} \mathrm{d}t L(q^i, \dot{q}^i, t) \tag{9.14}$$

has to be extremal among all paths $\gamma = \{q(t)\}$ with fixed initial point $q^i(t_0)$ and fixed final point $q^i(t_1)$. Since $\delta\dot{q}^i = \frac{d}{dt}\delta q^i$ the variation can be written as

$$\delta\phi = \int_{t_0}^{t_1} \mathrm{dt}\left(\frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial \dot{q}^i}\delta\dot{q}^i\right) = \int_{t_0}^{t_1} \mathrm{dt}\left((\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i})\delta q^i + \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}^i}\delta q^i)\right). \tag{9.15}$$

Due to the boundary conditions the variation $\delta q^i(t)$ is zero at the initial and at the final time so that the surface term $\int dt \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}^i}\delta q^i) = (\frac{\partial L}{\partial \dot{q}^i}\delta q^i)|_{t_0}^{t_1}$ vanishes. Extremality of the action $\delta\phi = 0$ for all variations is hence equivalent to the *Euler-Lagrange equations of motion*

$$\frac{\delta L}{\delta q^i} = 0 \qquad \text{with} \qquad \frac{\delta L}{\delta q^i} \equiv \frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = \frac{\partial L}{\partial q^i} - \dot{p}_i \tag{9.16}$$

where we introduced the *variational derivative* $\frac{\delta L}{\delta q^i}$ of L and the *canonical momentum* $p_i \equiv \frac{\partial L}{\partial \dot{q}^i}$. The space parametrized by the *canonical coordinates* q^i is called *configuration space*.

By *Legendre transformation* with respect to $\dot{q}^i$ we obtain the *Hamilton function*

$$H(p_i, q^i, t) = \sum p_i\dot{q}^i - L(q^i, \dot{q}^i, t) \qquad \text{with} \qquad p_i = \frac{\partial L}{\partial \dot{q}^i}, \tag{9.17}$$

as a function of the momenta p_i and the coordinates q^i, which together parametrize the *phase space*. Since the inverse Legendre transformation is given by eliminating the momenta p_i from the equation $\dot{q}^i = \partial H/\partial p_i(p, q)$ the *Hamiltonian equations of motion*

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q^i} \tag{9.18}$$

are equivalent to the Euler-Lagrange equations. The equations (9.18) can also be obtained directly as variational equations $\delta\tilde{L}/\delta p_i = 0$ and $\delta\tilde{L}/\delta q^i = 0$ of the *first order Lagrangian* $\tilde{L}(q, \dot{q}, p) = \dot{q}^i p_i - H(p, q)$. Infinitesimal time evolution, as any infinitesimal transformation,

obeys the Leibniz rule for products and the chain rule for phase space functions $f(q,p,t)$, for which we admit an explicit time dependence. Regarding f as a function of time on a classical trajectory we hence obtain

$$\dot{f} = \frac{\partial f}{\partial q^i}\dot{q}^i + \frac{\partial f}{\partial p_i}\dot{p}_i + \partial_t f = \{H, f\}_{PB} + \partial_t f \tag{9.19}$$

where we defined the *Poisson brackets*

$$\{f, g\}_{PB} \equiv \sum_i \left(\frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} - \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} \right) \tag{9.20}$$

for arbitrary phase space functions $f(p,q)$ and $g(p,q)$.

The Noether theorem. An infinitesimal transformation $q^i \to q^i + \hat{\delta}q^i$ with $\hat{\delta}q^j = f^j(q^i, \dot{q}^i)$ is a symmetry of a dynamical system with Lagrange function $L(q^i, \dot{q}^i)$ if $\hat{\delta}L = \dot{K}$ is a total time derivative because a total derivative does not contribute to the variation (9.15) and hence leaves the equations of motion invariant. The Noether theorem states that these infinitesimal symmetries are in one-to-one correspondence with *constants of motion* Q, which are also called *conserved charges* or *first integrals*. More explicitly, a symmetry $\hat{\delta}q$ with $\hat{\delta}L = \dot{K}$ implies that

$$Q = \hat{\delta}q^i p_i - K \tag{9.21}$$

is a constant of motion. In turn, if some phase space function $Q(q^i, \dot{q}^i)$ is a constant of motion for all classical trajectories then its time derivative is a linear combination of the equations of motion $\dot{Q} = \sum_i \rho^i \frac{\delta L}{\delta q^i}$. The transformation $\hat{\delta}q^i = -\rho^i(q^j, \dot{q}^j)$ is then a symmetry of the Lagrange function, i.e. $\hat{\delta}L = \dot{K}$ with $K = \hat{\delta}q^i p_i - Q$.

Remarks: A constant of motion is only constant for motions that obey the equations of motion! It is important to discern *identities* and *dynamical equations*. For a constant of motion $\dot{Q} = 0$ is a consequence of the equations of motion. This implies that there is an *identity* $\dot{Q} = \sum_i \rho^i \frac{\delta L}{\delta q^i}$ that holds for *arbitrary* functions $q^i(t)$ and not only for solutions to $\frac{\delta L}{\delta q^i} = 0$. A symmetry transformation $\hat{\delta}q^i$ (like e.g. a translation) does not have to vanish at an initial or final time!

Proof: According to (9.15) the equation

$$\delta L \;=\; \frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial \dot{q}^i}\delta \dot{q}^i \;\equiv\; \frac{\delta L}{\delta q^i}\delta q^i + \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\delta q^i\right) \tag{9.22}$$

is an *identity* for an arbitrary variation. $\hat{\delta}L = \dot{K}$ hence implies

$$\frac{\delta L}{\delta q^i}\hat{\delta}q^i + \frac{d}{dt}(p_i \hat{\delta}q^i) = \dot{K}. \tag{9.23}$$

The theorem follows by subtracting the time derivative $\frac{d}{dt}(p_i\hat{\delta}q^i)$ from this equation. □

Hamiltonian version. Coordinate transformations in phase space $(p,q) \to (p',q')$ with functions $p'(p,q)$ and $q'(p,q)$ that leave the form of the Poisson brackets invariant are called

canonical transformations. It can be shown that infinitesimal canonical transformations $\hat{\delta}$ can be written in terms of a generating function $Q(p,q)$ as

$$\hat{\delta}q^i = \{Q, q^i\}_{PB}, \qquad \hat{\delta}p_i = \{Q, p_i\}_{PB}. \tag{9.24}$$

For a fixed phase space function Q the map $g \to \{Q,g\}_{PB}$ is linear and obeys the Leibniz rule, as required for an infinitesimal transformation. In the canonical formalism a symmetry transformation is, by definition, a canonical transformation $\{Q,.\}_{PB}$ with some generating function $Q(p,q)$ that leaves the Hamilton function invariant $\{Q,H\}_{PB} = 0$. This makes the Noether theorem quite trivial, because $\{Q,H\}_{PB} = -\{H,Q\}_{PB} = -\dot{Q} = 0$ is at the same time the condition for Q to be a constant of motion.

The equivalence of the Hamiltonian and the Lagrangian definition of a symmetry, as well as the equality of the Noether charges Q, can be seen by computing a variation $\hat{\delta}$ of the first order Lagrangian $\tilde{L} = \dot{q}^i p_i - H$,

$$\begin{aligned} \hat{\delta}(\dot{q}^i p_i - H) &= \hat{\delta}\dot{q}^i\, p_i + \dot{q}^i \hat{\delta} p_i - \hat{\delta}H & (9.25)\\ &= \frac{d}{dt}(\hat{\delta}q^i\, p_i) - \hat{\delta}q^i \dot{p}_i + \dot{q}^i \hat{\delta} p_i - \partial_{q^i} H \hat{\delta} q^i - \partial_{p_j} H \hat{\delta} p_j, & (9.26) \end{aligned}$$

which can only be equal to $\dot{K}(p,q)$ if $\frac{d}{dt}(K - p_i\hat{\delta}q^i) = \dot{q}^i\partial_{q^i}(K - p_i\hat{\delta}q^i) + \dot{p}_i\partial_{p_i}(K - p_i\hat{\delta}q^i)$ is equal to $-\hat{\delta}q^i\dot{p}_i + \dot{q}^i\hat{\delta}p_i$ so that $\hat{\delta}$ is a canonical transformation

$$\hat{\delta}q^i = \frac{\partial Q}{\partial p_i} \quad \text{and} \quad \hat{\delta}p_i = -\frac{\partial Q}{\partial q^i} \qquad \Rightarrow \qquad \hat{\delta}f = \{Q,f\}_{PB} \tag{9.27}$$

with generating function $Q = p_i\hat{\delta}q^i - K$. The last two terms in (9.26), which do not contain time-derivatives and hence have to cancel each other, now combine to $\hat{\delta}H = \{Q,H\} = 0$. We thus have shown that the Noether charge Q, when expressed as a function of the canonical coordinates q^i and p_i, is the generating function for the transformation $\hat{\delta}$.

Quantization. Since $\{p_i, x_j\}_{PB} = \delta_{ij}$ in classical mechanics and $[P_i, X_j] = \frac{\hbar}{i}\delta_{ij}$ in quantum mechanics canonical quantization replaces Poisson brackets of conjugate phase space variables by $\frac{i}{\hbar}$ times the commutator of the corresponding operators,

$$\{p_i, x_j\}_{PB} = \delta_{ij} = \frac{i}{\hbar}[P_i, X_j]. \tag{9.28}$$

For the generating functions of infinitesimal transformations this amounts to

$$\hat{\delta}q^i = \{Q, q^i\}_{PB} \qquad \to \qquad \hat{\delta}\vec{X} = \frac{i}{\hbar}[Q, \vec{X}]. \tag{9.29}$$

Note that Poisson brackets and commutators both are antisymmetric, satisfy the Jacobi-Identity, and obey the Leibniz rule for each of its two arguments, so that $\{Q,.\}_{PB}$ and $[Q,.]$

both qualify as inifinitesimal transformations. Moreover, the *real* variation $\hat{\delta}q^i$ of the real coordinate q^i naturally leads to the anti-Hermitian operator $\frac{i}{\hbar}Q$ so that the finite transformation $\exp(\frac{i}{\hbar}Q)$ becomes a unitary operator. More precisely, one has to be careful about possible ordering ambiguities if $Q(q^i, p_i)$ is a composite operator. It is always possible to choose Q Hermitian (for example by $Q \to \frac{1}{2}(Q + Q^\dagger)$, or by Weyl ordering). In quantum mechanics it is usually also possible of find a proper quantum version of the symmetry generators, but for an infinite numbers of degrees of freedom quantum violations of classical symmetries, which are called anomalies, can be unavoidable and may lead to important restrictions for the structure of consistent theories.[4]

Energy, momentum and angular momentum. As an example we now compute the generators of translations and rotations. Under a time translation $\hat{\delta}q^i = \dot{q}^i$ of an autonomous system the Lagrange function transforms into its time derivative $\hat{\delta}L = \dot{K} = \dot{L}$, so that the corresponding Noether charge $Q = \hat{\delta}q^i p_i - K = \dot{q}^i p_i - L$ agrees with the Hamilton function. This proves the equivalence of time independence and energy conservation. Uppon quantization (9.29) we find

$$\frac{d}{dt}\vec{X} = \frac{i}{\hbar}[H, \vec{X}], \tag{9.30}$$

which is Heisenberg's equation of motion for the position operator $\vec{X}$. Canonical quantization hence naturally leads to the Heisenberg picture. The corresponding time evolution of the wave function in the Schrödinger picture is given by the Schrödinger equation $\frac{d}{dt}|\psi\rangle = -\frac{\hbar}{i}H|\psi\rangle$.

The generator of a translation $\hat{\delta}_i\vec{x} = \vec{e}_i$ into the coordinate direction $\vec{e}_i$ is the momentum p_i because $\hat{\delta}_i L = 0$, hence $K = 0$, and $\hat{\delta}_i x_j p_j = \vec{e}_i\vec{p} = p_i$ for a translation invariant Lagrange function. Under rotations a centrally symmetric action is also strictly invariant $\hat{\delta}_{\vec{\alpha}}L = \dot{K} = 0$. A rotation about the x_j-axis is given by $\hat{\delta}\vec{x} = \delta R_{e_j}\vec{x}$. With (9.10) we have $\hat{\delta}x_i = \varepsilon_{ijk}x_k$ and thus obtain the corresponding Noether charge

$$L_j = \hat{\delta}x_i p_i - 0 = \varepsilon_{ijk}x_k p_i = \varepsilon_{jki}x_k p_i \tag{9.31}$$

or $\vec{L} = \vec{x}\times\vec{p}$ in accord with the usual definition of angular momentum. The results are collected in the following table.

Symmetry	Noether charge	infinitesimal transformation
time evolution	Hamiltonian H	$\frac{d}{dt}\|\psi\rangle = -\frac{i}{\hbar}H\|\psi\rangle$
translation	momentum P_i	$-\vec{\nabla}\|\psi\rangle = -\frac{i}{\hbar}\vec{P}\|\psi\rangle$
rotation	orbital angular momentum L_i	$\delta R_\alpha\|\psi\rangle = -\frac{i}{\hbar}\vec{\alpha}\vec{L}\|\psi\rangle$

[4] In the standard model of particle interaction, for example, cancellation of certain anomalies between quarks and leptons is indespensible for the consisteny of the theory, while the anomalay in baryon number conservation is in principle observable and enables proton decay, one of Sakharov's conditions for the creation of matter. Anomalies are also the origin of the space-time dimension 10 in superstring theory.

For finite transformations $U = \exp(-\frac{i}{\hbar}H)$ is the time evolution operator and $\exp(-\frac{i}{\hbar}\vec{a}\vec{P})\psi(x)$ yields the Taylor series expansion of $\psi(\vec{x}-\vec{a})$ in the translation vector $\vec{a}$. For infinitesimal rotations

$$\delta R_{\vec{\alpha}}\psi(x) \equiv -\frac{i}{\hbar}\vec{\alpha}\vec{L}\psi(x) = -\frac{i}{\hbar}\alpha^i\varepsilon_{ijk}x_j\frac{\hbar}{i}\nabla_k\psi(x) = -\varepsilon_{ijk}\alpha_j x^k\nabla_i\psi(x), \tag{9.32}$$

in accord with (9.13).

Transformation of operators. A unitary transformation $|\psi\rangle \to T|\psi\rangle$ of states implies that the respective transformation of operators $\mathcal{O}$ is

$$|\psi\rangle \to T|\psi\rangle \qquad \Rightarrow \qquad \mathcal{O} \to T\mathcal{O}T^\dagger \tag{9.33}$$

because matrix elements should not change if we apply both transformations simultaneously. Since $T\mathcal{O}T^{-1} = (\mathbb{1}+\varepsilon\delta T)\mathcal{O}(\mathbb{1}-\varepsilon\delta T)+\mathcal{O}(\varepsilon^2) = \mathbb{1}+\varepsilon[\delta T,\mathcal{O}]+\mathcal{O}(\varepsilon^2)$ the infinitesimal version of this correspondence is

$$|\psi\rangle \to \delta T|\psi\rangle \qquad \Rightarrow \qquad \delta\mathcal{O} = [\delta T,\mathcal{O}]. \tag{9.34}$$

By the active–passive equivalence, an operator transformation $\mathcal{O} \to T\mathcal{O}T^\dagger$ can hence be replaced by the inverse transformation of states, projectors and density matrices

$$|\psi\rangle \to T^\dagger|\psi\rangle \qquad \Rightarrow \qquad P_\psi \to T^\dagger P_\psi T \quad \text{and} \quad \rho \to T^\dagger\rho T \qquad \text{with} \quad P_\psi = |\psi\rangle\langle\psi|, \tag{9.35}$$

which transforms expectation values $\operatorname{tr} P_\psi\mathcal{O} \to \operatorname{tr}(T^\dagger P_\psi T)\mathcal{O} = \operatorname{tr} P_\psi(T\mathcal{O}T^\dagger)$ in the same way. These rules hold for all unitary transformations and not only for symmetry transformations!

9.3 Rotation of spins

If we consider the total angular momentum operator $\vec{J} = \vec{L}+\vec{S}$ for a particle with spin $\vec{S}$ we obtain its finite rotations by

$$e^{-\frac{i}{\hbar}\vec{\alpha}\vec{J}} = e^{-\frac{i}{\hbar}\vec{\alpha}\vec{L}}e^{-\frac{i}{\hbar}\vec{\alpha}\vec{S}} \tag{9.36}$$

because the orbital angular momentum and the spin operator commute. The former operator is responsible for the shift in the position resulting from the rotation while the latter rotates the orientation of the spin. The operator $\exp(-\frac{i}{\hbar}\vec{\alpha}\vec{S})$ is hence called the rotation operator in *spin space*, and it is often sufficient to study its action if the spin orientation rather than the precise position of the particle is relevant for a computation. In the basis $|s,\mu\rangle$ where S^2 and S_z are diagonal we have

$$S_\pm|s,\mu\rangle = \hbar\sqrt{(s\mp\mu)(s\pm\mu+1)}\;|s,\mu\pm1\rangle \qquad \text{and} \qquad S_z|s,\mu\rangle = \hbar\mu\;|s,\mu\rangle \tag{9.37}$$

with $S_\pm = S_x \pm iS_y$ or $S_x = \frac{1}{2}(S_+ + S_-)$ and $S_y = \frac{1}{2i}(S_+ - S_-)$.

Spinors. For spin $s = \frac{1}{2}$ the wave function in the S_z-basis (9.37) can be written as

$$\psi(x) \equiv \psi_+(x)\,|\uparrow\rangle + \psi_-(x)\,|\downarrow\rangle \equiv \begin{pmatrix}\psi_+(x)\\ \psi_-(x)\end{pmatrix} \tag{9.38}$$

with $|\uparrow\rangle = |\frac{1}{2},\frac{1}{2}\rangle = \begin{pmatrix}1\\0\end{pmatrix}$ and $|\downarrow\rangle = |\frac{1}{2},-\frac{1}{2}\rangle = \begin{pmatrix}0\\1\end{pmatrix}$ and the spin operator becomes $\vec{S} = \frac{\hbar}{2}\vec{\sigma}$ in terms of the Pauli matrices

$$\sigma_x = \begin{pmatrix}0&1\\1&0\end{pmatrix},\quad \sigma_y = \begin{pmatrix}0&-i\\i&0\end{pmatrix},\quad \sigma_z = \begin{pmatrix}1&0\\0&-1\end{pmatrix} \qquad \text{with} \qquad \begin{array}{l}\sigma_i\sigma_j = \delta_{ij} + i\varepsilon_{ijk}\sigma_k\\ \operatorname{tr}\sigma_i = 0\end{array}. \tag{9.39}$$

Since $(\vec{\sigma}\vec{\alpha})^2 = \alpha_i\alpha_j\sigma_i\sigma_j = \alpha_i\alpha_j\delta_{ij}\mathbb{1} = \alpha^2\mathbb{1}$ exponentiation of $\delta R_{\vec{\alpha}}|\psi\rangle = -\frac{i}{\hbar}\vec{\alpha}\vec{S}|\psi\rangle = -\frac{i}{2}\vec{\alpha}\vec{\sigma}|\psi\rangle$ yields the finite spin rotations

$$R_\alpha = e^{-\frac{i}{2}\vec{\alpha}\vec{\sigma}} = \mathbb{1}\cos\frac{\alpha}{2} - i\,\vec{e}_\alpha\vec{\sigma}\,\sin\frac{\alpha}{2} \qquad \text{with} \qquad \alpha = |\vec{\alpha}| \quad \text{and} \quad \vec{e}_\alpha = \vec{\alpha}/\alpha, \tag{9.40}$$

which leave the position invariant but mix the spin-up and the spin-down components of the wave function. We observe that a rotation by an angle $\alpha = 2\pi$ transforms $|\psi\rangle \to -|\psi\rangle$. This strange behavior of spinors is not inconsistent because $\pm|\psi\rangle$ only differ by a phase and hence represent the same physical state of the system and cannot be distinguished by any observable. Phases do become observable, however, in interference pattens. The change of sign for a rotation by 2π has indeed been verified experimentally with neutrons interferometry by H. Rauch et al. in 1975, who achieved destructive interference between coherent neutron rays whose spins were rotated by a relative angle 2π.

Remark: The projector $\Pi_{|\uparrow,\vec{n}\rangle}$ onto a state with spin up in the direction of a unit vector $\vec{n}$ can be obtained by an active rotation $R_{\vec{\alpha}}$ of $|\uparrow\rangle\langle\uparrow| = \frac{1}{2}(\mathbb{1} + \sigma_z)$ with $\vec{\alpha} = \frac{\alpha}{\sin\alpha}\vec{e}_z \times \vec{n}$ for $\cos\alpha = n_z$,

$$R_{\vec{\alpha}} = \tfrac{(1+n_z)\mathbb{1} + in_y\sigma_x - in_x\sigma_y}{\sqrt{2(1+n_z)}}, \qquad \Pi_{|\uparrow,\vec{n}\rangle} \equiv |\uparrow,\vec{n}\rangle\langle\uparrow,\vec{n}| = R_{\vec{\alpha}}\tfrac{\mathbb{1}+\sigma_z}{2}R_{\vec{\alpha}}^\dagger = \tfrac{1}{2}(\mathbb{1} + \vec{n}\vec{\sigma}). \tag{9.41}$$

Without lengthy calculation the result directly follows from the fact that $\vec{n}\vec{\sigma} = \frac{2}{\hbar}\vec{n}\vec{S}$ has eigenvalues ±1 on states with spin component $\pm\frac{\hbar}{2}$ in the direction $\vec{n}$.

Vectors. For spin $s = 1$ the analog of the Pauli matrices can again be obtained from (9.37) with $S_x = \frac{1}{2}(S_+ + S_-)$ and $S_y = \frac{1}{2i}(S_+ - S_-)$,

$$S_x^{(1)} = \frac{\hbar}{\sqrt{2}}\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix} \qquad S_y^{(1)} = \frac{\hbar}{i\sqrt{2}}\begin{pmatrix}0&1&0\\-1&0&1\\0&-1&0\end{pmatrix} \qquad S_z^{(1)} = \hbar\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}. \tag{9.42}$$

In order to relate the spherical basis $|1,m\rangle$ to the standard vector basis $\vec{e}_i$ we start with

$$\vec{e}_z = |1,0\rangle \tag{9.43}$$

because $\vec{e}_z$ is an eigenvector of the infinitesimal rotation $\delta R_{\vec{e}_z}$ about the z-axis with eigenvalue 0. Evaluating $S_\pm = S_x \pm iS_y$ on $|1,0\rangle = \vec{e}_z$ we find

$$S_\pm|1,0\rangle = \hbar\sqrt{(1\mp 0)(1\pm 0+1)}|1,\pm 1\rangle = \sqrt{2}\hbar|1,\pm 1\rangle \stackrel{!}{=} (S_x \pm iS_y)\vec{e}_z = -\frac{\hbar}{i}(-\vec{e}_y \pm i\vec{e}_x) \quad (9.44)$$

because $-\frac{i}{\hbar}S_i$ generates an infinitesimal rotation about $\vec{e}_i$. Equality of these expressions implies

$$|1,\pm 1\rangle = \mp\frac{1}{\sqrt{2}}(\vec{e}_x \pm i\vec{e}_y). \quad (9.45)$$

The *spherical components* $V_q^{(1)}$ of a vector $\vec{V}$ in the S_z-basis $|1,m\rangle$ are hence defined by

$$V_0^{(1)} = V_z, \qquad V_{\pm 1}^{(1)} = \mp\frac{1}{\sqrt{2}}(V_x \pm iV_y). \quad (9.46)$$

Since $V_1^{(1)}W_{-1}^{(1)} + V_{-1}^{(1)}W_1^{(1)} = -V_xW_x - V_yW_y$ the scalar product of two vectors becomes

$$\vec{V}\cdot\vec{W} = \sum_{q=-1}^{1}(-1)^q\, V_q^{(1)}W_{-q}^{(1)} \quad (9.47)$$

in the spherical basis. As the matrices $S_i^{(1)}$ in (9.42) neither anticommute nor square to $\mathbb{1}$ there is no simple analog of the formula (9.40) for finite rotations $\exp(-\frac{i}{\hbar}\vec{\alpha}\vec{S}^{(1)})$. A general formula for arbitrary spin is known, however, if we represent the rotation in terms of the Euler angles.

General representation of the rotation group. Every rotation in $\mathbb{R}^3$ can be written as a combination of a rotation by an angle γ about the z-axis followed by a rotation by β about the y-axis and a rotation by α about the z-axis. The angles (α,β,γ) are called *Euler* angles (see appendix A.10 in [Grau]) and the corresponding rotation operator is

$$R = e^{-\frac{i}{\hbar}J_z\alpha}\cdot e^{-\frac{i}{\hbar}J_y\beta}\cdot e^{-\frac{i}{\hbar}J_z\gamma}. \quad (9.48)$$

Since $J_z|j,m\rangle = \hbar m|j,m\rangle$ the matrix elements of this operator in an eigenbasis of J^2 and J_z can be written as

$$\langle j,m'|R|j,m\rangle = e^{-im'\alpha}\langle j,m'|e^{-\frac{i}{\hbar}J_y\beta}|j,m\rangle\, e^{-im\gamma} \quad (9.49)$$

For the non-diagonal rotation operator about the y axis we define

$$d^{(j)}_{m'm}(\beta) = \langle j,m'|e^{-\frac{i}{\hbar}J_y\beta}|j,m\rangle \quad (9.50)$$

Without proof we state the formula

$$d^{(j)}_{m'm}(\beta) = (-1)^{m'-m}\sqrt{\frac{(j-m')!(j+m')!}{(j-m)!(j+m)!}}\sin^{m'-m}\left(\frac{\beta}{2}\right)\cos^{m'+m}\left(\frac{\beta}{2}\right)P_{j-m}^{m'-m,m'+m}(\cos\beta) \quad (9.51)$$

where $P_n^{r,s}(\xi)$ are the *Jacobi* polynomials, which can be defined by

$$P_n^{r,s}(\xi) = \frac{(n+r)!}{(n-r)!}\left(\frac{1+\xi}{2}\right)^2 F\left(-n,-n-s,r+1;\frac{\xi+1}{\xi-1}\right) \quad (9.52)$$

in terms of the hypergeometric function F.

SO(3) and SU(2). The tensor product $\langle\varphi| \otimes |\psi\rangle$ of two spinors corresponds to a 2×2 matrix with 4 degrees of freedom, which we expect to contain a scalar and vector. Since the three traceless Pauli matrices and the unit matrix together form a basis for all 2×2 matrices $\langle\varphi| \otimes |\psi\rangle$ can be written as linear combinations of

$$\langle\varphi|\mathbb{1}|\psi\rangle \qquad \text{and} \qquad \langle\varphi|\vec{\sigma}|\psi\rangle. \tag{9.53}$$

These matrix elements indeed transform as a skalar and a vector, respectively, since

$$\delta R_\alpha(\langle\varphi|\sigma_i|\psi\rangle) = -\frac{i}{2}\alpha_j(\langle\varphi|\sigma_i\sigma_j - \sigma_j\sigma_i|\psi\rangle) = -\frac{i}{2}\alpha_j\langle\varphi|2i\varepsilon_{ijk}\sigma_k|\psi\rangle = \alpha_j\varepsilon_{ijk}\langle\varphi|\sigma_k|\psi\rangle \tag{9.54}$$

and $\delta R_\alpha(\langle\varphi|\psi\rangle) = \frac{i}{2}(\langle\varphi|\alpha\sigma)|\psi\rangle - \frac{i}{2}\langle\varphi|(\alpha\sigma|\psi\rangle) = 0$. Since $\vec{\alpha}\vec{\sigma}$ is an arbitrary traceless Hermitian matrix the exponential $A = \exp(-\frac{i}{2}\vec{\alpha}\vec{\sigma})$ is a arbitrary special unitary matrix $A \in SU(2)$ which can be written as

$$SU(2) \ni A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad AA^\dagger = A\begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix} = \mathbb{1} \quad \Rightarrow \quad \begin{matrix} |a|^2 + |b|^2 = 1 \\ ca^* + db^* = 0 \end{matrix}. \tag{9.55}$$

The last equation implies $c = -\frac{b^*}{a^*}d$ so that $\det A = 1 = ad - bc = \frac{d}{a^*}(aa^* + bb^*) = \frac{d}{a^*}$. We thus obtain $d = a^*$, $c = -b^*$ and

$$A = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \qquad \text{with} \qquad |a|^2 + |b|^2 = 1. \tag{9.56}$$

As a manifold $SU(2)$ is therefore a 3-sphere with radius 1 in $\mathbb{R}^4$ whose coordinates are the real and imaginary parts of a and b. For finite transformations the spinor rotation $|\psi\rangle \to A|\psi\rangle$ leads to the vector rotation

$$\langle\varphi|\sigma_i|\psi\rangle \quad \to \quad \langle\varphi|A^\dagger\sigma_i A|\psi\rangle = \sum_k R_{ik}(A)\langle\varphi|\sigma_k|\psi\rangle. \tag{9.57}$$

The equation

$$A^\dagger\sigma_i A = R_{ik}(A)\sigma_k \tag{9.58}$$

hence defines a map from $A \in SU(2)$ to a rotation $R(A) \in SO(3)$. This map is two-to-one because A and $-A$ lead to the same rotation. This should not come as a surprise because we already know that a rotation by 2π, which is the identity in $SO(3)$, reverses the sign of a spinor. A and $-A$ are antipodal points of the S^3 that represents $SU(2)$. We can therefore think of $SO(3)$ as the 3-sphere with antipodal points identified and $SU(2)$ is a smooth double cover of the rotation group. The mathematical reason for the existence of spinor representations of the rotation group is the fact that $SO(3)$ admits an unbranched double-cover and hence admits so called *projective* or *ray* representations where a full rotation gives back the original state only up to a phase factor. Such objects would be forbidden in classical mechanics, but in quantum mechanics a physical state is not represented by a unique vector $|\psi\rangle \in \mathcal{H}$ but rather by a "ray" of vectors $\lambda|\psi\rangle$ with $\lambda \neq 0$.

9.3.1 Tensor operators and the Wigner Eckhart theorem

Vector and tensor operators are collections of operators labelled by vector or tensor indices that transform accordingly under rotations,

$$[J_i, V_j] = i\hbar\varepsilon_{ijl}V_l, \qquad [J_i, T_{jl}] = i\hbar\varepsilon_{ijm}T_{ml} + i\hbar\varepsilon_{iln}T_{jn}, \qquad \ldots \tag{9.59}$$

The number of indices is the *order* k of the tensor. Since

$$\vec{J}\cdot(T|\psi\rangle) = [\vec{J}, T]\cdot|\psi\rangle + T\cdot\vec{J}\,|\psi\rangle \tag{9.60}$$

the action of a general vector operator is like an addition of an angular momentum $j = 1$ as far as the properties under rotations are concerned. For a vector operator (9.46) we already know that

$$\left[J_z, V_q^{(1)}\right] = q\hbar V_q^{(1)} \tag{9.61}$$

$$\left[J_\pm, V_q^{(1)}\right] = \sqrt{2 - q(q\pm 1)}\hbar V_{q\pm1}^{(1)} \tag{9.62}$$

For higher order $k > 1$ a general tensor decomposes into irreducible parts that are connected by the ladder operators. An irreducible tensor operator T_q^k of the order k is defined by the following commutation relations,

$$\left[J_z, T_q^{(k)}\right] = \hbar q T_q^{(k)}, \tag{9.63}$$

$$\left[J_\pm, T_q^{(k)}\right] = \sqrt{k(k+1) - q(q\pm1)}\hbar T_{q\pm1}^{(k)}, \tag{9.64}$$

where k and q are the analogues of the eigenvalues l and m of the spherical harmonics and $q = -k, \ldots, k$ labels the $2k+1$ spherical components of the irreducible tensor operator $T^{(k)}$.

A tensor operator T_{il} of order 2, for example, has nine elements. Since two spins $j_1 = j_2 = 1$ add up to spin $j \leq 2$ we expect T_{il} to contain irreducible tensors of order $0, 1$ and 2. In this example it is easy to guess that the scalar is the trace $T^{(0)} = \delta^{il}T_{il}$, while the 3 vector degrees of freedom are found by anti-symmetrization $T_i^{(1)} = \frac{1}{2}\varepsilon_{ijl}T_{jl}$. This leaves the traceless symmetric part $T_{il}^{(2)} = \frac{1}{2}(T_{il} + T_{li}) - \frac{1}{3}\delta_{il}T^{(0)}$ to represent the 5 components of the irreducible operator of order 2. The proof of this is analogous to the addition of angular momenta. One first writes the tensor T_{il} in spherical coordinates with $i \to q_1$ and $l \to q_2$. Then the highest component must be $T_2^{(2)} = T_{+1,+1}$ (think of T_{il} as V_iW_l, then $T_2^{(2)} = V_{+1}^{(1)}W_{+1}^{(1)}$). The remaining components $T_q^{(2)}$ are then obtained by acting with J_-.

The Wigner-Eckhart theorem: In a basis $|\alpha, j, m\rangle$ of eigenvectors of J^2 and J_z the matrix elements of an irreducible tensor operator of order k are of the form

$$\boxed{\langle\alpha, j, m|T_q^{(k)}|\alpha', j', m'\rangle = \langle j', m', k, q|j, m, \rangle\langle\alpha, j||T^{(k)}||\alpha' j'\rangle} \tag{9.65}$$

where

- $\langle j', m', k, q | j, m, \rangle$ ... Clebsch-Gordon coefficients (independent of $T_q^{(k)}$),
- $\langle \alpha, j || T^{(k)} || \alpha' j' \rangle$ ... reduced matrix element(independent of $m,\ m',\ q$),
- α ... represents all other quantum numbers.

The Wigner-Eckhart theorem thus factorizes the matrix representation of $T_q^{(k)}$ into a geometric part, which is given by the Clebsch-Gordon coefficients, and a constant, called reduced matrix element, which does not depend on the magnetic quantum numbers.

Proof: For the proof we consider the $(2k+1)(2j'+1)$ vectors

$$T_q^{(k)} |\alpha', j', m'\rangle \qquad (q = -k, ..., k; \quad m' = -j', ..., j') \tag{9.66}$$

and their linear combinations

$$|\sigma, j'', m''\rangle = \sum_{q,m'} T_q^{(k)} |\alpha', j', m'\rangle \langle j', m', k, q | j'', m''\rangle, \tag{9.67}$$

where σ contains j' and α' as well as further quantum numbers that characterize $T^{(k)}$. The crucial point is that the collections $|\sigma, j'', m''\rangle$ of $(2k+1)(2j'+1)$ states for fixed σ and for $|m''| \leq j'' \leq j' + k$ indeed transform according to the irreducible spin-j'' representations, as is suggested by the notation. This follows from (9.60) and the definitions of tensor operators and Clebsch-Gordon coefficients. Since the latter form a unitary matrix we can invert this transformation and obtain

$$T_q^{(k)} |\alpha', j', m'\rangle = \sum_{j'', m''} |\sigma, j'', m''\rangle \langle j'', m'' | j', m', k, q\rangle. \tag{9.68}$$

If we now multiply this equation with $\langle \alpha, j, m|$ we get

$$\begin{aligned} \langle \alpha, j, m | T_q^{(k)} | \alpha', j', m'\rangle &= \sum_{j'', m''} \langle \alpha, j, m | \sigma, j'', m''\rangle \langle j'', m'' | j', m', k, q\rangle & (9.69) \\ &= \langle \alpha, j, m | \sigma, j, m\rangle \langle j, m | j', m', k, q\rangle & (9.70) \end{aligned}$$

with $\langle \alpha, j, m | \sigma, j, m\rangle \equiv \langle \alpha, j || T^{(k)} || \alpha', j'\rangle$ because $\langle \alpha, j, m | \sigma, j'', m''\rangle$ is 0 except for $j = j''$ and $m = m''$. The scalar product $\langle \alpha, j, m | \sigma, j, m\rangle$ does not depend on m, as can be shown by insertion of $[J_+, J_-] = 2J_3$, and its dependence on j' and $T^{(k)}$ is implicitly contained in σ. $\square$

As an application we consider the spherical harmonics $Y_q^{(k)} \equiv Y_{kq}$ with angular momentum k, operating on wave functions by multiplication,

$$\langle \alpha, j, m | Y_q^{(k)} | \alpha', j', m'\rangle = \delta_{\alpha\alpha'} \langle j', m', k, q | j, m\rangle \langle j || Y_k || j'\rangle. \tag{9.71}$$

The reduced matrix element is

$$\langle j || Y_k || j'\rangle = \langle j, 0, k, 0 | j', 0\rangle \sqrt{\frac{(2j+1)(2k+1)}{(2j'+1)}} \tag{9.72}$$

(see [Messiah] II, appendix C).

9.4 Symmetries of relativistic quantum mechanics

In chapter 7 we introduced the Dirac equation

$$i\hbar\dot{\psi} = H\psi, \qquad H = c\alpha_i(P_i - \frac{e}{c}A_i) + \beta mc^2 + e\phi \tag{9.73}$$

with 4-component spinors ψ and Hermitian 4×4 matrices β and α_i with 2×2 block entries

$$\beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \qquad \alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} \tag{9.74}$$

which satisfy the anticommutation relations $\{\alpha_i, \alpha_j\} = \delta_{ij}\mathbb{1}_{4\times 4}$, $\{\alpha_i, \beta\} = 0$, and $\beta^2 = \mathbb{1}_{4\times 4}$. In the relativistic notation we destinguish upper and lower indices $\mu = 0, \ldots, 3$ and combine space-time coordinates, scalar and vector potential, and energy-momentum to 4-vectors according to

$$x^\mu = (ct, \vec{x}), \qquad \partial_\mu = (\frac{1}{c}\partial_t, \vec{\nabla}), \qquad A^\mu = (\phi, \vec{A}), \qquad p^\mu = (\frac{1}{c}E, \vec{p}). \tag{9.75}$$

After multiplication with $\beta/\hbar c$ from the left and with the correspondence rule $p_\mu \to i\hbar\partial_\mu$ the Dirac equation (9.73) becomes [5]

$$\left(i\gamma^\mu\partial_\mu - \frac{e}{\hbar c}\gamma^\mu A_\mu - \frac{c}{\hbar}m\right)\psi(t, \vec{x}) = 0, \tag{9.76}$$

where we introduce the four matrices $\gamma^\mu = (\beta, \beta\vec{\alpha})$, which are unitary $(\gamma^\mu)^\dagger = (\gamma^\mu)^{-1}$ and satisfy the Clifford algebra

$$\{\gamma_\mu, \gamma_\nu\} = 2g_{\mu\nu}, \qquad \text{with} \qquad \gamma_\mu = g_{\mu\nu}\gamma^\nu \tag{9.77}$$

so that $(\gamma^0)^2 = \mathbb{1} = -(\gamma^i)^2$ and different γ's anticommute $\gamma^\mu\gamma^\nu = -\gamma^\nu\gamma^\mu$ if $\mu \neq \nu$. The four matrices γ^μ are the relativistic analog of the three Pauli matrices σ_i. Since relativistic spinors have 4 components (describing the spin-up and the spin-down degrees of freedom of particles *and* antiparticles) the γ^μ are 4×4 matrices. Their unitarity implies that only γ^0 is Hermitian while the three matrices γ^i are anti-Hermitian, which can be expressed in the formula

$$(\gamma^\mu)^\dagger = (\gamma^\mu)^{-1} = \gamma^0\gamma^\mu\gamma^0. \tag{9.78}$$

Explicitly the Dirac matrices read

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \qquad \text{[Pauli representation]}. \tag{9.79}$$

Matrix representations γ^μ of the Clifford algebra (9.77) are far from unique, but it can be shown that all unitary representations are related by unitary similarity transformations $\gamma^\mu \to U\gamma^\mu U^{-1}$. For concrete calculations it is usually much better to use the algebraic relations (9.77–9.78) than to use an explicit form of the γ-matrices.[6]

[5] In quantum electrodynamics it is common to introduce Feynman's *slash* notation $\not{a} \equiv \gamma^\mu a_\mu$ [read: a-slash] for contractions of vectors with γ matrices and to set $\hbar = c = 1$ so that the Dirac operator reads $(i\not{\partial} - e\not{A} - m)$ and $\not{a}^2 = a^2\mathbb{1}_{4\times 4}$, which generalizes the nonrelativistic formula $(\vec{v}\vec{\sigma})^2 = v^2\mathbb{1}_{2\times 2}$.

[6] For certain applications particular representations may, however, be useful: The Pauli representation is convenient for taking the non-relativistic limit (see chapter 7). In a Majorana representation all γ^μ are imaginary

$$\gamma^0 = \begin{pmatrix} \sigma_2 & 0 \\ 0 & \sigma_2 \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} i\sigma_1 & 0 \\ 0 & i\sigma_1 \end{pmatrix}, \quad \gamma^2 = \begin{pmatrix} 0 & i\sigma_3 \\ i\sigma_3 & 0 \end{pmatrix}, \quad \gamma^3 = \begin{pmatrix} i\sigma_3 & 0 \\ 0 & -i\sigma_3 \end{pmatrix} \qquad \text{[Majorana]} \tag{9.80}$$

9.4.1 Lorentz covariance of the Dirac-equation

We want to show now that the Dirac equation (9.76) retains its form under a Lorentz transformation $x'^\mu = L^\mu{}_\nu x^\nu$. For given $L^\mu{}_\nu$ we expect the Dirac spinor ψ to transform linearly $\psi'(x') = \Lambda(L)\psi(x)$ with some 4×4 matrix Λ depending on L. Note that we always use a matrix notation and never write explicit indices for spinors ψ and their linear transformations by multiplications with γ-matrices, which are four 4×4 matrices acting by matrix multiplication on 4-component *spinors* but labeled by a Lorentz vector index μ. One should not be confused by the coincidence that spinors and vectors have the same *number* of components in 4 dimensions. They nevertheless transform differently under Lorentz transformations![7]

For simplicity we consider the free Dirac equation with $A_\mu = 0$. Inserting

$$x'^\nu = L^\nu{}_\mu x^\mu \quad \Rightarrow \quad \partial_\mu = \frac{\partial}{\partial x^\mu} = \frac{\partial x'^\nu}{\partial x^\mu}\frac{\partial}{\partial x'^\nu} = L^\nu{}_\mu \frac{\partial}{\partial x'^\nu} = L^\nu{}_\mu \partial'_\nu \qquad \text{and} \quad \psi = \Lambda^{-1}\psi' \tag{9.82}$$

into the equation $(i\gamma^\mu\partial_\mu - \frac{c}{\hbar}m)\psi = 0$ we obtain

$$(i\gamma^\mu L^\nu{}_\mu \partial'_\nu - \tfrac{c}{\hbar}m)\Lambda^{-1}\psi' = 0. \tag{9.83}$$

This transforms into $(i\gamma^\nu\partial'_\nu - \frac{c}{\hbar}m)\psi' = 0$ by multiplication with Λ from the left provided that $\Lambda\gamma^\mu L^\nu{}_\mu\Lambda^{-1} = \gamma^\nu$ or

$$\boxed{\Lambda^{-1}\gamma^\mu\Lambda = L^\mu{}_\nu\gamma^\nu.} \tag{9.84}$$

This is the relativistic version of the equation (9.58) and defines the spinor transformation $\Lambda(L)$ in terms of a Lorentz transformation L. Like in the non-relativistic case $\pm\Lambda$ correspond to the same $L^\mu{}_\nu$ so that the spin group is a double cover of the Lorentz group. The condition (9.84) also guarantees covariance of the interacting Dirac equation (9.76) because the gauge potential A_μ transforms like the gradient ∂_μ.

Finite transformations $\Lambda(L)$ can be obtained by exponentiation of infinitesimal ones,

$$L^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu + \mathcal{O}(\omega^2) \qquad \Rightarrow \qquad \omega_{\mu\nu} = g_{\mu\rho}\omega^\rho{}_\nu = -\omega_{\nu\mu}. \tag{9.85}$$

Similarly to the electromagnetic field strength $F_{\mu\nu} = -F_{\nu\mu}$, which contains the electric and the magnetic fields as 3-vectors, the antisymmetric tensor $\omega_{\mu\nu}$ contains infinitesimal spacial

---

so that the free Dirac equation becomes real. Weyl representations are block-offdiagonal

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \qquad \text{[Weyl representation]} \tag{9.81}$$

decomposing the massless Dirac equations into two-component equations for left- and right-handed particles.

[7] We already know that spinors in 3 dimensions have 2 compoments, while vectors have 3 components; in higher dimensions $d > 4$, on the other hand, it can be shown that the number of components of spinors grows like $2^{d/2}$, which is much larger than the number d of vector components.

rotations $\delta R_{\vec\rho}$ about an axis $\vec\rho$

$$\delta R_{\vec\rho}\, x^i = \varepsilon_{ijk}\rho^j x^k = \omega^i{}_k x^k \qquad \Rightarrow \qquad \omega_{jk} = \rho^i \varepsilon_{ijk} \tag{9.86}$$

and infinitesimal boosts $\omega_{i0} = v_i/c$ in the direction of a velocity vector $\vec v$, whose finite form for $\vec v = v\vec e_x$ is

$$L^\mu{}_\nu = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \text{with} \quad \gamma = \frac{1}{\sqrt{1-\beta^2}} \ \text{ and } \ \beta = \frac{v}{c}. \tag{9.87}$$

With the ansatz $\Lambda(L) = \mathbb{1} + \frac{1}{2}\omega_{\mu\nu}\Sigma^{\mu\nu} + \mathcal{O}(\omega^2)$ the defining equation $\Lambda^{-1}\gamma^\rho\Lambda = L^\rho{}_\nu\gamma^\nu$ becomes

$$\tfrac{1}{2}\,\omega_{\mu\nu}[\gamma^\rho, \Sigma^{\mu\nu}] = \omega^\rho{}_\nu\gamma^\nu = g^{\rho\mu}\omega_{\mu\nu}\gamma^\nu \qquad \Rightarrow \qquad \tfrac{1}{2}[\gamma_\rho, \Sigma_{\mu\nu}] = \tfrac{1}{2}(g_{\rho\mu}\gamma_\nu - g_{\rho\nu}\gamma_\mu) \tag{9.88}$$

whose solution can be guessed to be of the form $\Sigma_{\mu\nu} = a[\gamma_\mu, \gamma_\nu]$. Since $[\gamma_\mu, \gamma_\nu] = 2\gamma_\mu\gamma_\nu - 2g_{\mu\nu}\mathbb{1}$ and

$$[\gamma_\rho, \Sigma_{\mu\nu}] = 2a[\gamma_\rho, \gamma_\mu\gamma_\nu] = 2a(\{\gamma_\rho, \gamma_\mu\}\gamma_\nu - \gamma_\mu\{\gamma_\rho, \gamma_\nu\}) = 4a(g_{\rho\mu}\gamma_\nu - \gamma_\mu g_{\rho\nu}) \tag{9.89}$$

equation (9.88) is indeed solved for $a = \frac{1}{4}$ and the solution can be shown to be unique. Hence

$$\Sigma_{\mu\nu} = \tfrac{1}{4}[\gamma_\mu, \gamma_\nu]. \tag{9.90}$$

For an alternative derivation of the relativistic spin operator $S_{\mu\nu} = -\frac{\hbar}{i}\Sigma_{\mu\nu}$ we write the spin operator S_i for spacial rotations, which has already been determined in chapter 7, in a Lorentz covariant form. We recall the formula

$$[H, L_i] = -i\hbar c\varepsilon_{ijk}\alpha_j P_k = -[H, S_i], \qquad \text{for} \qquad S_i = \frac{\hbar}{4i}\varepsilon_{ijk}\alpha_j\alpha_k = \frac{\hbar}{2}\begin{pmatrix}\sigma_i & 0 \\ 0 & \sigma_i\end{pmatrix} \tag{9.91}$$

from which we concluded that $\vec J = \vec L + \vec S$ is the conserved angular momentum and $\vec S$ is the spin. With $\vec\rho\vec S = \frac{\hbar}{4i}\rho^i\varepsilon_{ijk}\alpha_j\alpha_k = -\frac{\hbar}{4i}\rho^i\varepsilon_{ijk}\gamma^j\gamma^k$ we conclude that an infinitesimal rotation $\omega_{jk} = \rho^i\varepsilon_{ijk}$ should be given by

$$\delta\psi = -\tfrac{i}{\hbar}\vec\rho\vec S\psi = \tfrac{1}{4}\omega_{jk}\gamma^j\gamma^k\psi = \tfrac{1}{2}\omega_{ij}\Sigma^{ij}\psi, \tag{9.92}$$

which indeed is the specialization of our previous result to spacial rotations $\omega_{i0} = 0$, $\omega_{jk} = \rho^i\varepsilon_{ijk}$.

9.4.2 Spin and helicity

A covariant description of the spin of a relativistic particle can be given in terms of the Pauli-Lubanski vector

$$W_\alpha = -\frac{1}{2}\varepsilon_{\alpha\beta\gamma\delta}J^{\beta\gamma}P^\delta \quad \text{with} \quad \varepsilon^{0123} = 1 = -\varepsilon_{0123} \tag{9.93}$$

where $J^{\mu\nu} = x^\mu P^\nu - x^\nu P^\mu + S^{\mu\nu}$ is the total angular momentum. By evaluation in the center of mass frame it can be shown that the eigenvalues of $W^2 = W^\mu W_\mu$ are

$$W^2 = -m^2c^2\hbar^2\, s(s+1) \tag{9.94}$$

(without proof). For massless particles, however, $W^2 = 0$ so that s cannot be determined by W! The physical reason for this problem is that massless particles can never be in their center of mass frame. In fact, the spin quantum number j refers to the rotation group $SO(3)$ which cannot be used to classify massless particles exactly because of the non-existence of a center of mass frame (indeed, if photons could be described by spin $j = 1$, then the magnetic quantum number would have three allowed values; but we know that photons only have the two tranversal polarizations).

The intrinsic angular momentum of massless particles therefore has to be described by a different conserved quantity. Equation (9.91) implies that $\vec{p} \cdot \vec{\mathcal{S}}$ is a constant of motion

$$[\vec{p} \cdot \vec{\mathcal{S}}, H] = 0. \tag{9.95}$$

If $p = |\vec{p}| \neq 0$, which is always the case for massless particles, we can define the *helicity*

$$s_p = \frac{\vec{p} \cdot \vec{\mathcal{S}}}{p}, \tag{9.96}$$

which is the spin component in the direction of the velocity of the particle. For solutions of the Dirac equation its eigenvalues can be shown to be $s_p = \pm\hbar/2$. For a given momentum a Dirac particle can have two different helicities for positive-energy and two different helicities for negative energy solutions, so that the four degrees of freedom describe particles and antipartiles of both helicities. For the massless neutrinos, however, only positive helicity (left-handed) particles and negative helicity (right-handed) anti-particles exist in the standard model of particle interactions. The massless photons with $s_p = \pm\hbar$ and the gravitons with $s_p = \pm 2\hbar$ are their own antiparticles and they exist with two rather than $2s_p + 1$ polarizations.

9.4.3 Dirac conjugation and Lorentz tensors

If we try to construct a conserved current $j^\mu = (c\rho, \vec{j})$, which satisfies the continuity equation $\dot{\rho} + \mathrm{div}\vec{j} = 0$ and thus generalizes the probability density current of chapter 2, it is natural to consider the quantity $\psi^\dagger \gamma^\mu \psi$, which however is not real

$$(\psi^\dagger \gamma^\mu \psi)^* = (\psi^\dagger \gamma^\mu \psi)^\dagger = \psi^\dagger (\gamma^\mu)^\dagger \psi \neq \psi^\dagger \gamma^\mu \psi \tag{9.97}$$

because the γ^μ is anti-Hermitian for $\mu \neq 0$. It can, hence, also not transform as a 4-vector because Lorentz transformations would mix the real 0-component with the imaginary spacial components.

An appropriate real current can be constructed by replacing the Hermitian conjugate $\psi^\dagger$ by the *Dirac conjugate* spinor

$$\boxed{\overline{\psi} = \psi^\dagger \gamma^0,} \qquad \boxed{j^\mu = \overline{\psi} \gamma^\mu \psi.} \tag{9.98}$$

Now we can use eq. (9.78) and find

$$(j^\mu)^* = (\psi^\dagger\gamma^0\gamma^\mu\psi)^\dagger = \psi^\dagger(\gamma^\mu)^\dagger\gamma^0\psi = \psi^\dagger\gamma^0\gamma^\mu\psi = j^\mu \tag{9.99}$$

so that j^μ is indeed real.

The reason for introducing the Dirac conjugation can also be seen by computing the Lorentz transform of $\psi^\dagger$,

$$\langle\psi'| = \langle\psi|\Lambda^\dagger, \qquad \Lambda^\dagger \neq \Lambda^{-1}. \tag{9.100}$$

Considering infinitesimal transformations we observe that the non-unitarity of Λ again has its origin in the non-Hermiticity γ^μ and thus again can be compensated by conjugation with γ^0.

$$\begin{aligned}\left(\mathbb{1} + \frac{1}{8}\omega_{\mu\nu}[\gamma^\mu,\gamma^\nu]\right)^\dagger &= \mathbb{1} + \frac{1}{8}\omega_{\mu\nu}([\gamma^\mu,\gamma^\nu])^\dagger = \mathbb{1} + \frac{1}{8}\omega_{\mu\nu}[(\gamma^\nu)^\dagger,(\gamma^\mu)^\dagger] = \mathbb{1} - \frac{1}{8}\omega_{\mu\nu}[(\gamma^\mu)^\dagger,(\gamma^\nu)^\dagger] \\ &= \gamma^0(\mathbb{1} - \frac{1}{8}\omega_{\mu\nu}[\gamma^\mu,\gamma^\nu])\gamma^0 = \gamma^0\left(\mathbb{1} + \frac{1}{8}\omega_{\mu\nu}[\gamma^\mu,\gamma^\nu]\right)^{-1}\gamma^0 + \mathcal{O}(\omega^2).\end{aligned} \tag{9.101}$$

and hence for finite transformations

$$\boxed{\Lambda^\dagger = \gamma^0\Lambda^{-1}\gamma^0} \tag{9.102}$$

Λ is unitary for purely spacial rotations, but for Lorentz boosts it is not. For the Lorentz transformation of the Dirac adjoint spinor (9.102) implies

$$\psi' = \Lambda\psi \qquad \Rightarrow \qquad \overline{\psi}' = \psi^\dagger\Lambda^\dagger\gamma^0 = \psi^\dagger\gamma^0\Lambda^{-1} = \overline{\psi}\Lambda^{-1} \tag{9.103}$$

so that $(\overline{\psi}\psi)' = \overline{\psi}\psi$ is a scalar and the current j^μ transforms as a Lorentz vector,

$$(j^\mu)' = \overline{\psi}\Lambda^{-1}\gamma^\mu\Lambda\psi = L^\mu{}_\nu\overline{\psi}\gamma^\nu\psi = L^\mu{}_\nu j^\nu. \tag{9.104}$$

The divergence of the current $j^\mu = \overline{\psi}\gamma^\mu\psi$ can now be computed using the Dirac equation (9.76) and its conjugate

$$0 = \psi^\dagger\left((-i\overleftarrow{\partial}_\mu - \tfrac{e}{\hbar c}A_\mu)(\gamma^\mu)^\dagger - \tfrac{c}{\hbar}m\right)\gamma^0 = \overline{\psi}\left((-i\overleftarrow{\partial}_\mu - \tfrac{e}{\hbar c}A_\mu)\gamma^\mu) - \tfrac{c}{\hbar}m\right), \tag{9.105}$$

where $\psi^\dagger\overleftarrow{\partial}_\mu \equiv \partial_\mu\psi^\dagger$. For the divergence of the current j^μ we thus obtain

$$\begin{aligned}\partial_\mu(\overline{\psi}\gamma^\mu\psi) &= (\partial_\mu\overline{\psi}\gamma^\mu)\psi + \overline{\psi}(\gamma^\mu\partial_\mu\psi) \\ &= \left(\overline{\psi}(i\tfrac{e}{\hbar c}A_\mu\gamma^\mu + i\tfrac{c}{\hbar}m)\right)\psi + \overline{\psi}\left((-i\tfrac{e}{\hbar c}A_\mu\gamma^\mu - i\tfrac{c}{\hbar}m)\psi\right) = 0,\end{aligned} \tag{9.106}$$

which establishes the continuity equation $\partial_\mu j^\mu = 0$.

Lorentz tensors. Similarly to eq. (9.104) we can compute the Lorentz transformation for the insertion of a product of γ-matrices,

$$(\overline{\psi}\gamma^{\mu_1}\dots\gamma^{\mu_k}\psi)' = \overline{\psi}\Lambda^{-1}\gamma^{\mu_1}\Lambda\Lambda^{-1}\gamma^{\mu_2}\Lambda\dots\Lambda^{-1}\gamma^{\mu_k}\Lambda\psi = L^{\mu_1}{}_{\nu_1}\dots L^{\mu_k}{}_{\nu_k}\overline{\psi}\gamma^{\nu_1}\dots\gamma^{\nu_k}\psi \tag{9.107}$$

The expectation values $\overline{\psi}\gamma^{\mu_1}\dots\gamma^{\mu_k}\psi$ hence transform as Lorentz tensors of order k. Since

$$\gamma^\mu\gamma^\nu = \frac{1}{2}[\gamma^\mu,\gamma^\nu] + \frac{1}{2}\{\gamma^\mu,\gamma^\nu\} = \frac{1}{2}[\gamma^\mu,\gamma^\nu] + g^{\mu\nu}\mathbb{1} \tag{9.108}$$

every index symmetrization reduces the number of γ-matrix factors by two so that irreducible tensors are completely antisymmetric in their indices μ_i. In 4 dimensions we can antisymmetrize in at most 4 indices. The complete set of irreducible Lorentz tensors is listed in table 9.1 (we avoid the customary sympols $A^\mu \equiv \widetilde{V}^\mu$ for the axial vector and $P \equiv \widetilde{S}$ for the pseudoscalar to avoid confusion with with gauge potentials and parity).

	Lorentz tensor	$\binom{4}{k}$	C	P	T	CPT
Scalar	$S = \overline{\psi}\psi$	1	S	S	S	S
Vector	$V^\mu = \overline{\psi}\gamma^\mu\psi$	4	$-V^\mu$	V_μ	V_μ	$-V^\mu$
Antisym. tensor	$T^{\mu\nu} = \frac{i}{2}\overline{\psi}[\gamma^\mu,\gamma^\nu]\psi$	6	$-T^{\mu\nu}$	$T_{\mu\nu}$	$-T_{\mu\nu}$	$T^{\mu\nu}$
Axial vector	$\widetilde{V}^\mu = \frac{i}{3!}\varepsilon^{\mu\nu\rho\sigma}\overline{\psi}\gamma_\nu\gamma_\rho\gamma_\sigma\psi$	4	$\widetilde{V}^\mu$	$-\widetilde{V}_\mu$	$\widetilde{V}_\mu$	$-\widetilde{V}^\mu$
Pseudo scalar	$\widetilde{S} = \overline{\psi}\gamma^0\gamma^1\gamma^2\gamma^3\psi$	1	$\widetilde{S}$	$-\widetilde{S}$	$-\widetilde{S}$	$\widetilde{S}$

Table 9.1: Lorentz tensors of order k with $\binom{4}{k}$ components and their CPT transformation.

The total number of components is $\sum_{k=0}^{d}\binom{d}{k} = 2^d = 16$ and it can be shown that the antisymmetrized products of γ matrices are linearly independent.[8] They hence form a basis of all 16 linear operators in the 4-dimensional spinor space. This is the relativistic analog of the fact that the Pauli-matrices together with the unit matrix form a basis for the operators in the 2-dimensional spinor space of nonrelativistic quantum mechanics. The transformation properties under the discrete symmetries C, P and T are indicated in the last columns of table 9.1 and will be discussed in the next section.

9.5 Parity, time reversal and charge-conjugation

The nonrelativistic kinematics is invariant under the discrete symmetries parity $P\vec{x} = -\vec{x}$ and time reversal $Tt = -t$.

Parity. In quantum mechanics the classical action $P\vec{x} = -\vec{x}$ of partity is implemented by a unitary operator $\mathcal{P}$ in Hilbert space transforming $|\psi\rangle \to \mathcal{P}|\psi\rangle$ and hence

$$P\vec{x} = -\vec{x} \qquad \Rightarrow \qquad \mathcal{P}\,\vec{X}\,\mathcal{P}^\dagger = -\vec{X}, \qquad \mathcal{P}\vec{P}\,\mathcal{P}^\dagger = -\vec{P}. \tag{9.109}$$

[8] The proof of linear independence uses the lemma that all products of γ matrices that differ from $\pm\mathbb{1}$ are traceless. A spinor in d dimensions has $2^{[d/2]}$ components ($[d/2]$ is the greatest integer smaller or equal to $d/2$).

While coordinates and momenta are (polar) vectors, i.e. odd under parity, the angular momentum $\vec{x}\times\vec{p}$ is a *pseudo vector*, or *axial vector*), i.e. even under parity

$$L=\vec{x}\times\vec{p}\qquad\Rightarrow\qquad \mathcal{P}\vec{L}\mathcal{P}^\dagger=\vec{L},\qquad \mathcal{P}\,\vec{S}\,\mathcal{P}^\dagger=\vec{S} \tag{9.110}$$

Electromagnetic and strong interactions, as well as gravity, preserve parity. The form of the Maxwell equations implies that the electric field is a vector while the magnetic field transforms as an axial vector

$$\mathcal{P}\,\vec{E}\,\mathcal{P}^\dagger=-\vec{E},\qquad \mathcal{P}\,\vec{B}\,\mathcal{P}^\dagger=\vec{B}. \tag{9.111}$$

In the relativistic notation $A^\mu\to A_\mu$, i.e. A^0 is parity even and the vector potential $\vec{A}$ is parity odd. The spin-orbit coupling $\vec{L}\vec{S}$ and the magnetic coupling $\vec{B}(\vec{L}+2\vec{S})$ are axial-axial couplings and hence allowed by parity, while $\vec{E}\vec{S}$ would be a vector-axial coupling and is hence forbidden by parity. The parity of the spherical harmonics is

$$\mathcal{P}|Y_{lm}\rangle=(-1)^l\,|Y_{lm}\rangle \tag{9.112}$$

which is the basis for parity selections rules in atomic physics.

Parity is violated in weak interactions, as was first observed in the radioactive β-decay of polarized ^{60}Co. Since spin is parity-even the emission probability has to be the same for the angles θ and $\pi-\theta$ if parity is conserved. But experiments show that most electrons are emitted opposite to the spin direction $\theta>\pi/2$.

Time reversal. For a real Hamiltonian the effect of an inversion of the time direction can be compensated in the Schrödinger equation by complex conjugation of the wave fuction

$$t\to t'=-t\qquad\Rightarrow\qquad i\hbar\frac{\partial}{\partial t'}\psi^*=i\hbar\frac{\partial}{-\partial t}\psi^*=H\psi^*. \tag{9.113}$$

Time reversal therefore is implemented in quantum mechanics by an *anti*-unitary operator

$$\mathcal{T}\psi(t,\vec{x})=\psi^*(-t,\vec{x})\qquad\Rightarrow\qquad \langle\mathcal{T}\varphi|\mathcal{T}\psi\rangle=\langle\varphi|\psi\rangle^*,\qquad \mathcal{T}\alpha|\psi\rangle=\alpha^*\mathcal{T}|\psi\rangle \tag{9.114}$$

which implies complex conjugation of scalar products but leaves the norms $\sqrt{\langle\psi|\psi\rangle}$ invariant. For antilinear operators Hermitian conjugation can be defined by $\langle\varphi|\mathcal{T}^\dagger|\psi\rangle=\langle\psi|\mathcal{T}|\varphi\rangle$. Anti-unitary is then equivalent to antilinearity and $\mathcal{T}\mathcal{T}^\dagger=1$. Since velocities and momenta change their signs under time inversion we have

$$\mathcal{T}\,X_i\,\mathcal{T}^{-1}=X_i,\qquad \mathcal{T}\,P_i\,\mathcal{T}^{-1}=-P_i \tag{9.115}$$

The above formulas are compatible with the canonical commutation relations

$$\mathcal{T}[P_i,X_j]\mathcal{T}^{-1}=[-P_i,X_j]=-\frac{\hbar}{i}\delta_{ij}=\mathcal{T}\frac{\hbar}{i}\delta_{ij}\mathcal{T}^{-1}. \tag{9.116}$$

Invariance of the Maxwell equations under time reversal implies

$$\mathcal{T}\,\vec{E}\,\mathcal{T}^{-1}=\vec{E},\qquad \mathcal{T}\,\vec{B}\,\mathcal{T}^{-1}=-\vec{B},\qquad \mathcal{T}\,A^0\,\mathcal{T}^{-1}=A^0,\qquad \mathcal{T}\,\vec{A}\,\mathcal{T}^{-1}=-\vec{A} \tag{9.117}$$

so that gauge potentials transforms in the same way $A^\mu \to A_\mu$ as under parity. In fundamental interactions violation of time reversal invariance has only been observed for weak interactions.

9.5.1 Discrete symmetries of the Dirac equation

In addition to parity and time reversal the relativistic theory has another discrete symmetry, called charge conjugation, which is the exchange of particles and anti-particles.

Parity. Invariance of the Dirac equation under the parity transformation $\vec{x}\to-\vec{x}$ implies that $(i\gamma^\mu\partial'_\mu-m)\psi'(x')=0$ should be equivalent to $(i\gamma^\mu\partial_\mu-m)\psi(x)=0$ for $\psi'=\mathcal{P}\psi$, hence

$$\mathcal{P}^{-1}\left(i\left(\gamma^0\frac{\partial}{\partial x^0}+\gamma^i\frac{\partial}{\partial(-x^i)}\right)-m\right)\mathcal{P}\psi=(i\gamma^\mu\frac{\partial}{\partial x^\mu}-m)\psi \tag{9.118}$$

which implies

$$\mathcal{P}^{-1}\gamma^0\mathcal{P}=\gamma^0,\quad \mathcal{P}^{-1}\gamma^i\mathcal{P}=-\gamma^i \qquad\Rightarrow\qquad \mathcal{P}\,|\psi\rangle=\gamma^0\,|\psi\rangle. \tag{9.119}$$

Equation (9.118) actually fixes the action of the unitary parity operator $\mathcal{P}$ on spinors only up to an irrelevant phase factor and we followed the usual choice.

Charge conjugation. According to Dirac's hole theory exchange of particles and anti-particles should reverse the signs of electric charges and hence the sign of the gauge potential

$$\mathcal{C}\,\vec{E}\,\mathcal{C}^{-1}=-\vec{E},\qquad \mathcal{C}\,\vec{B}\,\mathcal{C}^{-1}=-\vec{B},\qquad \mathcal{C}\,A^\mu\,\mathcal{C}^{-1}=-A^\mu. \tag{9.120}$$

The derivation of the action of $\mathcal{C}$ on spinor starts with the observation that the relative sign between the kinetic and the gauge term is reversed in the conjugated Dirac equation (9.105). Transposition of that equation yields

$$\left((-\gamma^\mu)^T(i\partial_\mu+\tfrac{e}{\hbar c}A_\mu)-\tfrac{c}{\hbar}m\right)\overline{\psi}^T=0. \tag{9.121}$$

This lead to the condition

$$\mathcal{C}^{-1}\,(-\gamma^\mu)^T\,\mathcal{C}=\gamma^\mu \qquad\Rightarrow\qquad \mathcal{C}\,\psi=i\gamma^2\gamma^0\,\overline{\psi}^T=i\gamma^2\psi^* \tag{9.122}$$

in the standard representation[9] (9.79). Charge conjugation is hence an anti-unitary operation. Since $i\gamma^2=\begin{pmatrix}0 & i\sigma_2\\ -i\sigma_2 & 0\end{pmatrix}=\begin{pmatrix}0 & -\varepsilon\\ \varepsilon & 0\end{pmatrix}$ with $\varepsilon=\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$ charge conjugation exchanges positive and negative energy solutions and their chiralities.

[9] While $\mathcal{P}=\gamma^0$ is representation independent the explicit form of the antilinear operators $\mathcal{C}$ and $\mathcal{T}$ in terms of γ-matrices depends on the representation used for

Time reversal. The complex conjugate Dirac equation for $t' = -t$ is $(-i\gamma^{\mu *}\partial'_\mu - m)\psi^* = 0$. With the ansatz $\mathcal{T}|\psi\rangle = B|\psi^*\rangle$ this implies

$$B^{-1}(-i\gamma^{\mu *})B = i\gamma^0\gamma^\mu\gamma^0 \qquad \Rightarrow \qquad B = i\gamma^1\gamma^3, \qquad \mathcal{T}|\psi\rangle = i\gamma^1\gamma^3|\psi^*\rangle \tag{9.123}$$

in the standard representations (with the customary choice of the phase).

The transformation properties for the Lorentz tensors that follow from the formulas (9.119, 9.122, 9.123) for the $\mathcal{P}$, $\mathcal{C}$ and $\mathcal{T}$ are listed in table 9.1. The CPT theorem states that the combination of these three discrete transformations is a symmetry in every local Lorentz-invariant quantum field theory. The proof is based in the fact that all Lorentz tensors of order k (the complete set of fermion bilinears in table 9.1 as well as scalar fields and gauge fields A^μ) transform with a factor $(-1)^k$ under CPT and that Lorentz invariant interaction terms have no free Lorentz indices. Violation of time reversal invariance thus becomes equivalent to CP violation, which was first observed in 1964.[10]

9.6 Gauge invariance and the Aharonov–Bohm effect

An important aspect of the electromagnetic interaction with quantum particles is the fact that the interaction term in the Dirac equation

$$(\gamma^\mu(i\partial_\mu - \tfrac{e}{\hbar c}A_\mu) - \frac{c}{\hbar}m)\psi = 0, \tag{9.126}$$

[10] CP-violation in kaon decay is observed as follows [Nachtmann]. The theory of strong interactions implies that nucleons like the proton $|p\rangle = |uud\rangle$ with $m_p = 938MeV$ and the neutron $|n\rangle = |udd\rangle$ with $m_n = 940MeV$ consist of three quarks, while mesons like the pions $|\pi^+\rangle = |u\overline{d}\rangle$, $|\pi^-\rangle = |d\overline{u}\rangle$ with $m_{\pi^\pm} = 140MeV$ and $|\pi^0\rangle = (|u\overline{u}\rangle - |d\overline{d}\rangle)/\sqrt{2}$ with $m_{\pi^0} = 135MeV$ consist of a quark and an anti-quark. The K mesons $|K^+\rangle = |u\overline{s}\rangle$, $|K^-\rangle = |s\overline{u}\rangle =$ with $m_{K^\pm} = 494MeV$ and $|K^0\rangle = |d\overline{s}\rangle$, $|\overline{K}^0\rangle = |s\overline{d}\rangle$ with $m_{K^0} = 498MeV$ contain the somewhat heavier strange quark s and hence can decay by weak interactions.

Now the states $|K^0\rangle = \mathcal{C}|\overline{K}^0\rangle$ and $|\overline{K}^0\rangle = \mathcal{C}|K^0\rangle$ are each others antiparticles and both are observed to be parity odd $\mathcal{P}|K^0\rangle = -|K^0\rangle$ (according to the parity conserving strong processes in which they are created). The neutral K-mesons can only decay by weak interactions,

$$|K^0\rangle,\ |\overline{K}^0\rangle \quad \to \quad \pi^+\pi^-,\ \ \pi^0\pi^0, \quad \pi^+\pi^-\pi^0,\ \ \pi^0\pi^0\pi^0,\ \ \pi^\pm e^\mp\nu,\ \ \pi^\pm\mu^\mp\nu. \tag{9.124}$$

which break parity as well as charge conjugation. If we assume that weak interactions preserve the combination CP then the CP eigenstates

$$|K^0_{(\pm)}\rangle = \tfrac{1}{\sqrt{2}}(|K^0\rangle \mp |\overline{K}^0\rangle, \quad \mathcal{CP}|K^0_{(\pm)}\rangle = \pm|K^0_{(\pm)}\rangle \qquad \to \qquad |K_S\rangle \equiv |K^0_{(+)}\rangle, \quad |K_L\rangle \equiv |K^0_{(-)}\rangle \tag{9.125}$$

can only decay into CP eigenstates. In particular, $|K_S\rangle$ (S=short lived, with $\tau_s = 0.89 \cdot 10^{-10}s$) can decay into two pions (a CP-even state because $l = 0$ by angular momentum conservation), while $|K_L\rangle$ (L=long lived, with $\tau_L = 5.18 \cdot 10^{-8}s$) can only have the less likely 3-particle decays. But $|K_L\rangle$ is observed to decay into two pions with a probability of about 0.3%. Moreover, the 3-particle decay of $|K_L\rangle$ producing a positively charged lepton is observed to be 0.66% more likely than its decay into the CP conjugate states containing an electron e^- or a muon μ^-.

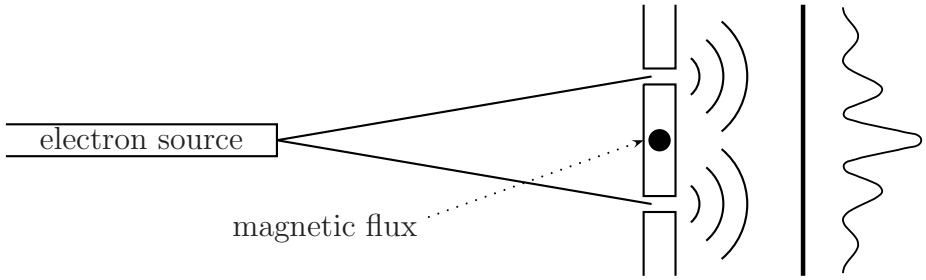


Figure 9.1: In the modified double slit experiment proposed by Aharonov and Bohm one observes a shift of the interference pattern proportional to the magnetic flux although the electrons only move in field-free regions.

like in the Schrödinger equation, explicitly depends on the gauge potential A_μ, which is not observable because the electromagnetic fields $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ are invariant under $A_\mu \to A'_\mu = A_\mu - \partial_\mu \Lambda$. We already noticed in chapter 2 that the complete wave equation is invariant under this gauge transformation if we simultaneoly change the equally unobservable phase of the wave function

$$A'_\mu = A_\mu - \partial_\mu \Lambda(t, \vec{x}) \qquad \Rightarrow \qquad \psi' = e^{\frac{ie}{\hbar c}\Lambda}\psi \tag{9.127}$$

because

$$(i\partial_\mu - \tfrac{e}{\hbar c}A'_\mu)e^{\frac{ie}{\hbar c}\Lambda} = e^{\frac{ie}{\hbar c}\Lambda}(i\partial_\mu - \tfrac{e}{\hbar c}A_\mu). \tag{9.128}$$

In the nonrelativistic limit this gauge invariance, split into space and time components, is inherited to the Schrödinger equation.

In 1959 Y. Aharonov and D. Bohm made the amazing prediction of an apparent action at a distance due to the form of the gauge interaction, which was experimentally verified by R.C. Chambers in 1960. This phenomenon is also used for practical applications like SQIDS (superconducting quantum interference devices) and it implies flux quantization in superconductors [Schwabl].

The experimental setup is a modification of the double slit experiment as shown in figure 9.1 where a magnetic flux is put between the two electron beams. For an infinitely long coil the B-field is confined inside the coil so that the flux lines cannot reach the domain where the electrons move. Nevertheless, a flux $\phi_B = \int \vec{B} d\vec{S}$ between the two rays leads to a relative phase shift

$$\delta(\arg(\psi_1) - \arg(\psi_2)) = \frac{e}{\hbar c}\phi_B \tag{9.129}$$

and hence to a shift in the interference pattern on the screen behind the slits.

In the Aharonov–Bohm experiment all electromagnetic fields are static. In a domain without magnetic flux $\vec{B} = \vec{\nabla} \times \vec{A} = 0$ the vector potential $\vec{A}$ is curl-free and hence can locally (in a contractible domain) be written as a divergence $\vec{A} = \vec{\nabla}\Lambda(\vec{x})$. This can be considered as a gauge

transform of the special solution $\vec{A} = 0$. The "potential" Λ for the vector potential $\vec{A} = \text{grad}\Lambda$ can be computed as a line integral $\Lambda(x) = \int_{x_0}^{x} \vec{A} \cdot d\vec{s}$, which is invariant under continuous deformations of the path as long as we stay in regions without magnetic flux. The computation of the phase shift (9.129) now uses this fact to relate the wave functions $\psi_i^{(0)}$ of the coherent electron beams in the double slit experiment without magnetic field, which are solutions to the Schrödinger equation with $\vec{A} = 0$, to the wave functions $\psi_i^{(B)}$ by gauge transformations. For each of the two beams the trajectories C_i belong to a contractible domain so that we can choose a gauge

$$\Lambda_i(\vec{x}) = \int_{C_i(x)} \vec{A} \cdot d\vec{s} \qquad \Rightarrow \qquad \psi_i^{(B)}(\vec{x}) = \psi_i^{(0)}(\vec{x})\, e^{\frac{ie}{\hbar c}\Lambda_i(\vec{x})}, \tag{9.130}$$

where the contour $C_i(x)$ starts at the electron source and extends along C_i to the position of the electron. Since the initial points of the paths C_i at the electron source and their final points at the screen where the interference is observed are the same for both paths the difference between the phase shifts is

$$\frac{e}{\hbar c}\left(\int_{C_2} \vec{A} \cdot d\vec{s} - \int_{C_1} \vec{A} \cdot d\vec{s}\right) = \frac{e}{\hbar c} \oint \vec{A} \cdot d\vec{s} = \frac{e}{\hbar c} \int (\nabla \times \vec{A})\, d\vec{S} = \frac{e}{\hbar c} \int \vec{B}\, d\vec{S} = \frac{e}{\hbar c}\phi_B, \tag{9.131}$$

where Stockes' theorem has been used to convert the circle integral extending along the closed path $C_2 - C_1$ into a surface integral over a surface enclosed by the beams. This surface integral $\int (\nabla \times \vec{A}) d\vec{S}$ measures the complete magnetic flux between the beams. This completes the derivation of the phase shift (9.129). Since the gauge transformation is the same for the Dirac equation and for the Schrödinger equation the relativistic and the nonrelativistic calculations are identical.

Chapter 10

Many–particle systems

In previous chapters we mostly considered the behaviour of a single quantum particle in a classical environment. We now turn to systems with many quantum particles, like electrons in solid matter or in atoms with $Z \gg 1$. The basic principle that identical particles cannot be distinguished in the quantum world will have important and far-reaching consequences.

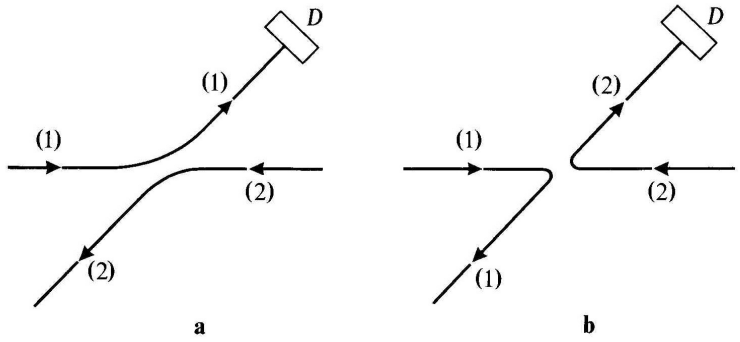


Figure 10.1: Two types of paths that the system of two colliding particles could have followed.

As an example we consider the scattering of two identical particles, as shown in figure 10.1. Oppositely polarized electrons essentially remain distinguishable because the magnetic spin interaction is negligible as compared to the electric repulsion. If the detector D counts electrons of any spin then the classical probabilities for both processes just add up

$$\frac{d\sigma}{d\Omega} = |f(\theta)|^2 + |f(\pi - \theta)|^2. \tag{10.1}$$

If beams (1) and (2) are both polarized spin up then the detector cannot distinguish between the incident particles and we will see that the rules of quantum mechanics tell us to superimpose amplitudes rather than probabilities. Particles with integral spin, like α-particles, follow the Bose–Einstein statistics, which means that their amplitudes have to be added

$$\left(\tfrac{d\sigma}{d\Omega}\right)_{BE} = |f(\theta) + f(\pi - \theta)|^2, \tag{10.2}$$

while particles with half-integral spin, like electrons, obey Fermi–Dirac statistics, which means that their amplitudes have to be subtracted

$$\left(\tfrac{d\sigma}{d\Omega}\right)_{FD} = |f(\theta) - f(\pi - \theta)|^2. \tag{10.3}$$

At an angle $\theta = \pi/2$ this implies

$$\left(\tfrac{d\sigma}{d\Omega}\right)_{BE}(\tfrac{\pi}{2}) = 4\,|f(\tfrac{\pi}{2})|^2, \qquad \left(\tfrac{d\sigma}{d\Omega}\right)_{FD}(\tfrac{\pi}{2}) = 0 \tag{10.4}$$

so that identical fermions never scatter at an angle of $\pi/2$, while the differential cross section for boson scattering becomes twice the classical value for this angle [Feynman].

In the present chapter we discuss the construction of many particle Hilbert spaces and their application to systems with many identical particles. As applications we discuss the derivation of (10.1–10.3) for particle scattering and the Hartree–Fock approximation for the computation of energy levels in atoms. We then introduce the occupation number representation and discuss the quantization of the radiation field, which will allow us to compute the amplitudes for electromagnetic transitions. As a last point we briefly discuss phonons and the concept of quasiparticles.

10.1 Identical particles and (anti)symmetrization

Particles are said to be identical if all their properties are exactly the same. In classical mechanics the dynamics of a system of N identical particles is described by a Hamilton function that is invariant under all $N!$ permutations $i \to \pi_i \equiv \pi(i)$ of the index set for the positions $\vec{x}_i$ and momenta $\vec{p}_i$,

$$H(\vec{x}_1, \vec{p}_1, \ldots, \vec{x}_N, \vec{p}_N) = H(\vec{x}_{\pi(1)}, \vec{p}_{\pi(1)}, \ldots, \vec{x}_{\pi(N)}, \vec{p}_{\pi(N)}), \qquad \pi \equiv \begin{pmatrix} 1 & 2 & \ldots & N \\ \pi_1 & \pi_2 & \ldots & \pi_N \end{pmatrix}. \tag{10.5}$$

The permutation group is generated by transpositions

$$\pi_{ij} = \begin{pmatrix} 1 \ldots i \ldots j \ldots N \\ 1 \ldots j \ldots i \ldots N \end{pmatrix}, \tag{10.6}$$

and the sign $(-1)^n$ of a permutation π is defined in terms of the number n modulo 2 of required transpositions,

$$\mathrm{sign}(\pi \circ \pi') = \mathrm{sign}(\pi) \cdot \mathrm{sign}(\pi'), \qquad \mathrm{sign}(\pi_{ij}) = -1. \tag{10.7}$$

Even and odd permutations have $sign_\pi = +1$ and $sign_\pi = -1$, respectively.

Although identical particles cannot be distinguished by their properties we can number them at some instant of time and identify them individually at later times by following their trajectories, as is illustrated for the example of a scattering experiment in figure 10.1. In

quantum mechanics, however, there are no well-defined trajectories and the wave functions start to overlap in the interaction region. It is hence no longer possible to tell which of the two particles went into the detector. The principle of indistinguishability of identical particles states that there is no observable that can distinguish between the state with particle 1 at position $|x_1\rangle$ and particle 2 at position $|x_2\rangle$ and the state with the positions of the particles exchanged. But then the state vectors in the Hilbert space $\mathcal{H}_2$ of two identical particles for the states with quantum numbers $|x_1, x_2\rangle_{id}$ and $|x_2, x_1\rangle_{id}$ must be the same up to a phase

$$|x_2, x_1\rangle_{id} = e^{i\rho}|x_1, x_2\rangle_{id} \tag{10.8}$$

because otherwise, for example, the projector $|x_1, x_2\rangle\langle x_1, x_2|$ could be used to find out whether the position x_1 is occupied by particle 1 or by particle 2.

In the case of N *distinguishable* particles the wave functions $\psi(\vec{x}_1, \ldots, \vec{x}_N)$ are *arbitrary* (normalizable) functions of the N positions. An N-particle state is thus in general not a product state $\varphi_{\lambda_1}(x_1) \cdot \ldots \cdot \varphi_{\lambda_N}(x_N)$ but a superposition of such states and hence an element of the tensor product

$$\psi(x_1, \ldots, x_N) \in \mathcal{H}^N \equiv \underbrace{\mathcal{H} \otimes \ldots \otimes \mathcal{H}}_{N} \tag{10.9}$$

of N copies of the 1-particle Hilbert space. The states $|\varphi_\lambda\rangle$ can be taken from an arbitrary complete orthonormal basis of $\mathcal{H}$.

Often the positions are combined with the magnetic quantum numbers describing the spin degrees of freedom (and possibly other quantum numbers) as $q_i = (\vec{x}_i, m_i)$. Permutations π of the particles correspond to unitary operations $\mathcal{P}_\pi$ in the product space

$$\mathcal{P}_\pi|q_1 \ldots q_N\rangle = |q_{\pi(1)} \ldots q_{\pi(N)}\rangle. \tag{10.10}$$

For identical particles $|q_1, q_2\rangle_{id}$ and $|q_2, q_1\rangle_{id}$ should correspond to indistinguishable states

$$|q_j, q_i\rangle_{id} = \mathcal{P}_{ij}|q_i, q_j\rangle_{id} = e^{i\rho}|q_i, q_j\rangle_{id} \in \mathcal{H}_2 \qquad \text{with} \qquad \mathcal{P}_{ij} \equiv \mathcal{P}(\pi_{ij}) \tag{10.11}$$

in the 2-particle Hilbert space $\mathcal{H}_2$. It is now usually argued that

$$\mathcal{P}_{ij}^2|q_i, q_j\rangle_{id} = e^{2i\rho}|q_i, q_j\rangle_{id} = |q_i, q_j\rangle_{id}, \tag{10.12}$$

hence

$$2\rho \in 2\pi\mathbb{Z} \qquad \Rightarrow \qquad |q_j, q_i\rangle_{id} = \mathcal{P}_{ij}|q_i, q_j\rangle_{id} = \pm|q_i, q_j\rangle_{id}, \tag{10.13}$$

because exchanging the position twice brings us back to the original state. Particles for which $|q_j, q_i\rangle_{id} = +|q_i, q_j\rangle_{id}$ are called bosons and particles for which $|q_j, q_i\rangle_{id} = -|q_i, q_j\rangle_{id}$ are called fermions. For more than two particles every permutation can be obtained as a product of transpositions and it is easy to see that N-boson states are invariant under $\mathcal{P}_\pi$ while N-fermion

states transform with a factor $\text{sign}(\pi)$.[1] Based on the axioms of relativistic quantum field theory Wolfgang Pauli (1940) proved the spin statistic theorem, which states that particles are bosons if they have integer spin $j \in \mathbb{Z}$ and fermions if they have half-integral spin $j \in \mathbb{Z} + \frac{1}{2}$.[2]

Symmetrization and antisymmetrization. For bosons and fermions the N-particle Hilbert spaces $\mathcal{H}_N^{(B)}$ and $\mathcal{H}_N^{(F)}$ can now be constructed as subspaces of the N-particle Hilbert space $\mathcal{H}^N$ of distinguishable particles. We introduce the *symmetrization* operator

$$\mathcal{S} = \frac{1}{N!} \sum_\pi P_\pi \tag{10.14}$$

and the *antisymmetrization* operator

$$\mathcal{A} = \frac{1}{N!} \sum_\pi \text{sign}(\pi) P_\pi. \tag{10.15}$$

The operators $\mathcal{S}$ and $\mathcal{A}$ are Hermitian because $(\mathcal{P}_\pi)^\dagger = \mathcal{P}_{\pi^{-1}}$ and the set of all permutations is equal to the set of all inverse permutations, for which $\text{sign}(\pi^{-1}) = \text{sign}(\pi)$. Similarly it can be shown that both operators are idempotent

$$\mathcal{S}^2 = \frac{1}{N!} \sum_\pi P_\pi \mathcal{S} = \frac{1}{N!} \sum_\pi \mathcal{S} = \mathcal{S} = \mathcal{S}^\dagger, \tag{10.16}$$

$$\mathcal{A}^2 = \frac{1}{N!} \sum_\pi \text{sign}(\pi) P_\pi \mathcal{A} = \frac{1}{N!} \sum_\pi \mathcal{A} = \mathcal{A} = \mathcal{A}^\dagger \tag{10.17}$$

and hence projection operators. States of the form $\mathcal{S}|\psi\rangle$ and $\mathcal{A}|\psi\rangle$ are eigenstates of the transposition operator $\mathcal{P}_{ij}$ with eigenvalues $+1$ and -1, respectively. Moreover, $\mathcal{S}$ and $\mathcal{A}$ project onto orthogonal eigenspaces, $\mathcal{S}\mathcal{A} = \mathcal{A}\mathcal{S} = 0$, and commute with all observables $\mathcal{O}$ for identical particles

$$[\mathcal{S}, \mathcal{O}] = [\mathcal{A}, \mathcal{O}] = 0 \tag{10.18}$$

because such observables must be invariant under every exchange of two identical particles $[\mathcal{P}_{ij}, \mathcal{O}] = 0$. The images of the projectors $\mathcal{S}$ and $\mathcal{A}$ hence can be used as Hilbert spaces for the

[1] The conclusion $\mathcal{P}_{ij}^2 = \mathbb{1}$ in eq. (10.12) is not stringent because, in principle, $\mathcal{P}_{ij}^2$ could differ from the identity by a phase. For 2-dimensional quantum systems that violate parity it is indeed conceivable that the phase $e^{i\rho}$ in (10.11) depends on the direction in which the particles are moved about one another so that the phase of $\mathcal{P}_{ij} = \mathcal{P}_{ji}^{-1}$ remains free. The particles are then neither bosons nor fermions and were therefore called *anyons* in the 1970s. For rational phases $\frac{1}{2\pi}\rho \in \mathbb{Q}$ such particles would have *fractional statistics* or *braid group statistics*. In the present context the braid group relates to the permutation group in the same way as the spin group $SU(2)$ relates to the rotation group $SO(3)$: A double exchange, like a rotation by 2π, is physically unobservable but still can lead to a non-trivial phase in quantum mechanics. The permutation of two particles in two dimensions can thus be regarded as a braiding process where the phase $e^{i\rho}$ depends on which strand is above and which strand is below.

Fractional statistics presumably has indeed been observed in the 1980s in the *fractional quantum Hall effect*, where charge carriers with fractional charges $Q = 1/3$ ... up to about $Q = 1/11$ have been observed. These are believed to be "quasi-particles" that obey a corresponding fractional statistics. Parity violation in these effectively two-dimensional thin layers is due to a strong magnetic field.

[2] In two dimensions the rotation group $SO(2)$ is abelian and therefore spin is not quantized. In accord with the spin-statistics connection fractional statistics, as discussed in the previous footnote, comes along with fractional spin of the quasi-particles on the fractional quantum Hall effect.

quantum mechanical description of identical particles

$$\mathcal{H}_N^{(B)} = \mathcal{S}(\underbrace{\mathcal{H}\otimes\ldots\otimes\mathcal{H}}_{N}), \qquad \mathcal{H}_N^{(F)} = \mathcal{A}(\underbrace{\mathcal{H}\otimes\ldots\otimes\mathcal{H}}_{N}). \tag{10.19}$$

Operators corresponding to permutation invariant observables automatically restrict to well-defined operators on $\mathcal{H}_N^{(B)}$ and on $\mathcal{H}_N^{(F)}$.

Given some basis $|q_j\rangle$ of $\mathcal{H}$ we now want to construct useful bases for the Hilbert spaces $\mathcal{H}_N$ of N identical particles. To get started we consider the examples of antisymmetrized two-particle and three-particle states

$$|q_1,q_2\rangle_A = \tfrac{1}{\sqrt{2}}(|q_1\rangle\otimes|q_2\rangle - |q_2\rangle\otimes|q_1\rangle) = \sqrt{2}\,\mathcal{A}|q_1,q_2\rangle, \tag{10.20}$$

$$|q_1,q_2,q_3\rangle_A = \tfrac{1}{\sqrt{3!}}\sum_\pi \mathrm{sign}(\pi)\,|q_{\pi_1}\rangle^{(1)}\,|q_{\pi_2}\rangle^{(2)}\,|q_{\pi_3}\rangle^{(3)} = \sqrt{3!}\,\mathcal{A}|q_1,q_2,q_3\rangle, \tag{10.21}$$

where the superscript i of $|q_j\rangle^{(i)}$ refers to the number of the particle and the subscript j refers to the quantum numbers labelling an orthonormal basis of 1-particle wave functions $\varphi_j(\vec{x}) \equiv |q_j\rangle \in \mathcal{H}$. It is easily verfied that ${}_A\langle q_1,q_2|q_1,q_2\rangle_A = {}_A\langle q_1,q_2,q_3|q_1,q_2,q_3\rangle_A = 1$. More generally, the antisymmetrized product states

$$|q_1,q_2,\ldots\rangle_A = \sqrt{N!}\;\mathcal{A}|q_1,q_2,\ldots\rangle \tag{10.22}$$

provide an orthonormal basis of $\mathcal{H}_N$,

$${}_A\langle q_1,q_2,\ldots|q_1,q_2,\ldots\rangle_A = 1, \qquad q_i \neq q_j \quad \text{for} \quad i\neq j, \tag{10.23}$$

where the normalization factor had to be chosen as $\sqrt{N!}$ because only the $N!$ scalar products in the double sum over all permutations of the *bra* and the *ket* vectors for which the permutations match contribute to the norm. The antisymmetrization of a product state can also be written as a determinant

$$\psi_A(q_1,\ldots,q_N) = \sqrt{N!}\,\mathcal{A}\,|q_1,\ldots,q_N\rangle = \frac{1}{\sqrt{N!}}\begin{vmatrix} |q_1\rangle^{(1)} & |q_1\rangle^{(2)} & \ldots & |q_1\rangle^{(N)} \\ \vdots & \vdots & & \vdots \\ |q_N\rangle^{(1)} & |q_N\rangle^{(2)} & \ldots & |q_N\rangle^{(N)} \end{vmatrix}, \tag{10.24}$$

called *Slater determinant*, which vanishes if two quantum numbers agree. Antisymmetrization hence implies Pauli's exclusion principle.

For bosons we can similarly construct an orthonormal basis as

$$|q_1\ldots q_N\rangle_S = \sqrt{\frac{N!}{n_1!\ldots n_r!}}\;\mathcal{S}|q_1\ldots q_N\rangle, \qquad \sum_{j=1}^{r} n_j = N, \tag{10.25}$$

where the normalization ${}_S\langle q_1\ldots q_N|q_1\ldots q_N\rangle_S = 1$ has required additional factors $1/\sqrt{n_j!}$ if groups of n_j of the quantum numbers q_i agree because then all terms where the order of identical

quantum numbers is exchanged also contribute in the double sum over all permutations of the quantum numbers of the *bra* and the *ket* vectors. (If, for example, all quantum numbers agree, then $|q,\ldots,q\rangle$ is already symmetric and the prefactor becomes $\sqrt{N!/N!}=1$). In analogy to the Slater determinant the symmetrization of product states is sometimes written in terms of the *permutant* $\left|\,|q_i\rangle^{(j)}\right|_+$,

$$\mathcal{S}\,|q_1,\ldots,q_N\rangle = \frac{1}{N!}\begin{vmatrix} |q_1\rangle^{(1)} & |q_1\rangle^{(2)} & \ldots & |q_1\rangle^{(N)} \\ \vdots & \vdots & & \vdots \\ |q_N\rangle^{(1)} & |q_N\rangle^{(2)} & \ldots & |q_N\rangle^{(N)} \end{vmatrix}_+ , \tag{10.26}$$

which is defined similarly to the determinant except that all signs of the $N!$ terms are positive.

10.2 Electron-electron scattering

The above considerations imply that our ansatz (8.20) for the asymptotic scattering wave function

$$u_{as} = (e^{i\vec{k}\cdot\vec{x}})_{as} + f(k,\theta)\frac{e^{ikr}}{r} \tag{10.27}$$

has to be modified for identical particles. With $u_S=\sqrt{2}\mathcal{S}u_{as}$ and $u_A=\sqrt{2}\mathcal{A}u_{as}$ it becomes

$$u_{\{_A^S} = \frac{1}{\sqrt{2}}\left((e^{i\vec{k}\vec{x}}\pm e^{-i\vec{k}\vec{x}}) + (f(\theta)\pm f(\pi-\theta))\frac{e^{ikr}}{r} + \mathcal{O}(\frac{1}{r^2})\right) \tag{10.28}$$

in the center of mass system, which leads to the differential cross section

$$\frac{d\sigma}{d\Omega} = |f(\theta)\pm f(\pi-\theta)|^2 \tag{10.29}$$

as we anticipated in the introduction of the present chapter.

For non-scalar wave functions we have to be more precise, however, because antisymmetrization non only affects the positions but also the other quantum numbers. For scattering of identical spin 1/2 particles like electrons the relevant quantum numbers are the relative coordinate $\vec{x}$ and the magnetic quantum numbers m_1, m_2 of the two particles. In the total spin basis the spin part of the wave function is either in the singlet state

$$|u\rangle^{(singlet)} = u_S(\vec{x})\chi_S, \qquad \chi_S = \tfrac{1}{\sqrt{2}}(\,|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \tag{10.30}$$

or in the triplet state

$$|u\rangle^{(triplet)} = u_T(\vec{x})\chi_T, \qquad \chi_T = \begin{cases} |\uparrow\uparrow\rangle \\ \frac{1}{\sqrt{2}}(\,|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \\ |\downarrow\downarrow\rangle \end{cases}. \tag{10.31}$$

Since the spin part of the singlet is antisymmetric the total antisymmetrization leads to a symmetrization of the position space wave function and hence

$$\left(\frac{d\sigma}{d\Omega}\right)_S = |f(\theta)+f(\pi-\theta)|^2\,, \tag{10.32}$$

while the triplet is symmetric under exchange of the two electrons so that the position part has to be antisymmetrized

$$\left(\frac{d\sigma}{d\Omega}\right)_T = |f(\theta) - f(\pi-\theta)|^2 \,. \tag{10.33}$$

For unpolarized electrons the triplet state is 3 times more likely than the singlet state and since

$$|f(\theta) \pm f(\pi-\theta)|^2 = |f(\theta)|^2 + |f(\pi-\theta)|^2 \pm \Big(f(\theta)f^*(\pi-\theta) + f^*(\theta)f(\pi-\theta)\Big) \tag{10.34}$$

the classical probabilities $\rho_T = 3/4$ and $\rho_T = 1/4$ imply

$$\frac{d\sigma}{d\Omega} = \frac{3}{4}\left(\frac{d\sigma}{d\Omega}\right)_T + \frac{1}{4}\left(\frac{d\sigma}{d\Omega}\right)_S = |f(\theta)|^2| + f(\pi-\theta)|^2 - \mathrm{Re}\Big(f(\theta)f^*(\pi-\theta)\Big)\,. \tag{10.35}$$

This is in accord with our anticipation that electrons with different spin orientation experience no quantum interference while the exclusion principle affects the 50% of the scattering events where both spins are up or both spins are down.

For Coulomb scattering of electrons we recall formula (8.145) for the amplitude,

$$f(\theta) = -\frac{\gamma}{2k\sin^2(\theta/2)} e^{i(2\sigma_0 - \gamma\log\sin^2(\theta/2)} \tag{10.36}$$

with $\sigma_0 = \mathrm{Im}\log\Gamma(1+i\gamma)$, and hence

$$f(\pi-\theta) = -\frac{\gamma}{2k\cos^2(\theta/2)} e^{i(2\sigma_0 - \gamma\log\cos^2(\theta/2)}. \tag{10.37}$$

In $f(\theta)f^*(\pi-\theta)$ the constant σ_0 drops out and the logarithms combine to $\log\tan^2(\theta/2)$. We thus arrive at *Mott's scattering formula*

$$\frac{d\sigma}{d\Omega} = \frac{\gamma^2}{4k^2}\left(\frac{1}{\sin^4(\theta/2)} + \frac{1}{\cos^4(\theta/2)} - \frac{\cos(\gamma\log\tan^2(\theta/2))}{\sin^2(\theta/2)\cos^2(\theta/2)}\right), \tag{10.38}$$

which shows that the quantum mechanical interference term and its modification by the phase correction to the classical formula for Coulomb scattering can be observed already for unpolarized electrons.

10.3 Selfconsistent fields and Hartree-Fock

The Hamilton function for an atom with N electrons and nuclear charge Z consists of the kinetic and potentials energies $T_i + V_i$ of the electrons in the electric field of the nucleus and the repulsive interaction terms W_{ij} among the electrons,

$$H \;=\; \sum_{i=1}^{N}(T_i + V_i) + \frac{1}{2}\sum_{i\neq j} W_{ij} \;=\; \sum_{i=1}^{N}\left(\frac{\vec{p}_i{}^2}{2m} - \frac{Ze^2}{r_i}\right) + \frac{1}{2}\sum_{i\neq j}\frac{e^2}{|\vec{x}_i - \vec{x}_j|} \tag{10.39}$$

or

$$H \;=\; H^{(1)} + H^{(2)} \qquad \text{with} \qquad H^{(1)} = \sum_{i=1}^{N} (T_i + V_i), \quad H^{(2)} = \frac{1}{2} \sum_{i \neq j} W_{ij}. \tag{10.40}$$

If N is large then it is plausible to assume that the potential that is felt by an individual electron is approximately independent of its own motion. We can hence think of each electron as moving in a *mean field* that is determined a posteriory in a *self consistent* way, i.e. we compute the electron states for a given potential $\widetilde{V}_i$ and then make sure that the electron states indeed produce exactly (or at least approximately) that potential.

In the present section we discuss selfconsistent methods for a *central potential* $V(r)$, i.e. in the context of atomic physics. Similar methods can also be used for solids, where the electrons move in the periodic potential of the nuclei that are located on a crystal lattice.

The Hartree method. Under the assumptions of the selfconsistent approach we can determine energy eigenstates $|\varphi_\alpha(\vec{x})\rangle$ by solving the Schrödinger equation for $\tilde{H}_i = T_i + \tilde{V}_i$ and then fill up the available orbits with increasing energies. This is motivation for the Hartree approximation, which assumes that the wave functiton is of the product form

$$\psi(q_1, \ldots, q_N) = \varphi_{\alpha_1}(q_1) \ldots \varphi_{\alpha_N}(q_N) \tag{10.41}$$

where $\varphi_i(q_i) = \varphi_i(\vec{x}_i)\chi_i$ are energy eigenstates

$$\tilde{H}_i|\varphi_i\rangle = (T_i + \tilde{V}_i)|\varphi_i\rangle = E_i|\varphi_i\rangle \tag{10.42}$$

and $\chi_i = |\frac{1}{2}, \pm\frac{1}{2}\rangle$ describes the spin degree of freedom. Within this class of wave fuctions the φ_i are determined with the help of the variational principle. The Pauli exclusion principle is implemented in the naiv way of assuming that any two eigenfunctions $\varphi_{\alpha_i}(\vec{x})$ and $\varphi_{\alpha_j}(\vec{x})$ are different except for a possible two-fold degeneracy for electrons that differ by their spin degrees of freedom $\chi_i \neq \chi_j$.

The variational method has been introduced in chapter 6, where we have shown that the Schrödinger equation is equivalent to the variational equation $\delta E[\psi] = 0$ for the energy functional $E[\psi] = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}$, which assumes its minimum exactly if ψ is the ground state wave function. If ψ is restricted to belong to a family of trial wave functions then an approximation to the ground state is obtained by minimizing $E[\psi]$ within that family. The quality of this approximation depends on the quality of chosen candidate family. In the Hartree approximation the family consists of all N-partile wave functions of the product form (10.41).

The main trick of this approach is to implement the orthonormalization of the one-particle wave functions φ_i by a collection of Lagrange multipliers ε_{ij}. We hence extremize the extended functional

$$E[\psi, \varepsilon_{ij}] = \langle\psi|H\psi\rangle + \sum_{i,j} \varepsilon_{ij} \Big(\delta_{ij} - \langle\varphi_i|\varphi_j\rangle\Big) \tag{10.43}$$

for arbitrary variations of $|\varphi_i\rangle$ and ε_{ij}. In order to simplify out notation we ignore for a moment a possibe degeneracy of the configuration space wave functions $\varphi_{\alpha_i}(\vec{x}) = \varphi_{\alpha_j}(\vec{x})$ in case of different spins $s_i \neq s_j$, which would reduce the number of Lagrange multipliers. For the evaluation of the expectation value $\langle\psi|H\psi\rangle$ it is important to note that the one-particle part $H^{(1)}$ of the Hamiltonian (10.39) is a sum of terms that only act on one of the factors of the product wave function (10.41) while all others are unaffected so that the respective *bra*'s and *ket*'s multiply to 1,

$$\langle\psi|\ H^{(1)}\ |\psi\rangle = \sum_{i=1}^{N} \langle\varphi_i|\ (T_i + V_i)\ |\varphi_i\rangle. \tag{10.44}$$

Similarly, the two-particle Hamiltonian $H^{(2)}$, which describes the interaction of two electrons, consists of a sum of terms that only act on two factors of the wave function, while the remaining factors can again be ignored,

$$\langle\psi|\ H^{(2)}\ |\psi\rangle = \frac{1}{2}\sum_{i\neq j} \langle\varphi_i, \varphi_j|\ W_{ij}\ |\varphi_i, \varphi_j\rangle. \tag{10.45}$$

Since orthonormality of the $|\varphi_i\rangle$ is implemented by the Lagrange multipliers we can freely vary all factors $\varphi_i(x_i)$ of the wave function $|\psi\rangle$. In the treatment of the variational method in chapter 6 we have shown that extremality under real and imaginary variations of $\varphi_i(\vec{x}_i)$ is equivalent to a formally independent variation of the bra-vector $\langle\delta\varphi_i|$ with fixed ket $|\varphi_i\rangle$. The variational equation hence becomes

$$\begin{aligned}\delta E[\psi, \varepsilon_{ij}] \;=\; & \sum_{i=1}^{N}\langle\delta\varphi_i|(T_i + V_i)|\varphi_i\rangle + \frac{1}{2}\sum_{i\neq j}\ \Big(\langle\delta\varphi_i, \varphi_j| + \langle\varphi_i, \delta\varphi_j|\Big)\ W_{ij}\ |\varphi_i, \varphi_j\rangle \\ & - \sum_{ij}\langle\delta\varphi_i|\varepsilon_{ij}|\varphi_j\rangle = 0\end{aligned} \tag{10.46}$$

which implies

$$(T_i + V_i)|\varphi_i\rangle + \sum_{j\neq i}\langle\varphi_j|W_{ij}|\varphi_j\rangle\,|\varphi_i\rangle = \sum_j \varepsilon_{ij}|\varphi_j\rangle. \tag{10.47}$$

Since the Lagrange multipliers ε_{ij} in (10.43) form a Hermitian matrix this matrix can be diagonalized by a unitary transformation of the $|\varphi_i\rangle$. We thus obtain the *Hartree equation*, which can be written in a more explicit notation as

$$\left(-\frac{\hbar^2}{2m}\Delta - \frac{Ze^2}{r}\right)\varphi_i(\vec{x}) + e^2\sum_{j\neq i}\int\frac{|\varphi_j(\vec{x}\,')|^2}{|\vec{x} - \vec{x}\,'|}d^3x'\ \varphi_i(\vec{x}) = \varepsilon_i|\varphi_i\rangle. \tag{10.48}$$

In addition to the one-particle potential V_i it contains the *Hartree potential*

$$V_i^H(\vec{x}) = \langle\varphi_j|W_{ij}|\varphi_j\rangle = e^2\sum_{j\neq i}\int\frac{|\varphi_j(\vec{x}')|^2}{|\vec{x} - \vec{x}'|}d^3x', \tag{10.49}$$

which describes the combinded repulsion by the other electrons. The entries ε_i of the diagonalized matrix of Lagrange multipliers thus obtain the meaning of energy eigenvalues of an

auxiliary one-particle Schrödinger equation (10.42) with potential $\widetilde{V}_i = V_i + V_i^H$. The complete binding energy of the system can now be written as

$$E = \langle H \rangle = \sum_{i=1}^{N} \varepsilon_i - \frac{e^2}{2} \sum_{i \neq j} \iint |\varphi_i(\vec{x})|^2 \frac{1}{|\vec{x} - \vec{x}'|} |\varphi_j(\vec{x}')|^2 \, d^3x d^3x', \tag{10.50}$$

where the energy of the electron-electron interaction has to be subtracted because is counted twice in the sum over the one-particle energies ε_i.

Hartree–Fock. We now improve the product ansatz for the wave function by antisymmetrization, as we should for fermionic N-particle states, and replace (10.41) by

$$\psi_A(q_1, \ldots, q_N) = \sqrt{N!}\, \mathcal{A}\, \varphi_{\alpha_1}(q_1) \ldots \varphi_{\alpha_N}(q_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_{\alpha_1}(q_1) & \ldots & \varphi_{\alpha_1}(q_N) \\ \vdots & & \vdots \\ \varphi_{\alpha_N}(q_1) & \ldots & \varphi_{\alpha_N}(q_N) \end{vmatrix}. \tag{10.51}$$

Evaluating the functional (10.43) for the Hartree–Fock family $\{\psi_A\}$ of wave fuctions it is straightforward to verify that the expectation value of the one-particle Hamiltonian $H^{(1)}$ remains unchanged, while the two-particle interaction term (10.45) obtains an additional contribution,

$$\langle \psi_A |\, H^{(2)} \, | \psi_A \rangle = \frac{1}{2} \sum_{i \neq j} \Big(\langle \varphi_i, \varphi_j |\, W_{ij} \, | \varphi_i, \varphi_j \rangle - \langle \varphi_j, \varphi_i |\, W_{ij} \, | \varphi_i, \varphi_j \rangle \Big), \tag{10.52}$$

because now also permutations for which the two interacting particles are exchanged can have a non-vanishing expectation value. The detailed calculation can be done by cancelling the normalization factor $(1/\sqrt{N!})^2$ of the Stater determinants against the sum over all permutations of the positions in the *ket*-vectors. This leaves us with a signed sum over all orderings of the *bra* quantum numbers for a fixed *ket*. But then the contribution of nontrivial permutations of the *bra* vectors vanishes because of the orthogonality of the factors of ψ unless all displaced factors are modified by the action of a nontrivial operator. For the two-particle operator W_{ij} this keeps the identity and the transposition $\mathcal{P}_{ij}$. For the one-particle operator $H^{(1)}$ only the trivial permutation survives.

The second term in (10.52) is called *exchange energy*. Since it is negative it amounts to an attractive force that reduces the mutual repulsion of the electron. Variation of the *bra*-vectors in (10.52) adds an exchange contribution

$$- \sum_{i \neq j} \langle \varphi_j, \delta\varphi_i |\, W_{ij} \, | \varphi_i, \varphi_j \rangle \tag{10.53}$$

to the variational equation and we obtain the *Hartree–Fock equation*

$$\boxed{T_i|\varphi_i\rangle + V_i|\varphi_i\rangle + \sum_{j \neq i} (\langle \varphi_j | W_{ij} | \varphi_j \rangle) | \varphi_i \rangle - \sum_{j \neq i} (\langle \varphi_j | W_{ij} | \varphi_i \rangle) | \varphi_j \rangle = \varepsilon_i |\varphi_i\rangle,} \tag{10.54}$$

which is an integro-differential equation for $\varphi_i(x)$ because T_i acts as a differential operator while $\varphi_i(x')$ is integrated over in the exchage term. Including the spin degrees of freedom into our discussion we note that the sum over $j \neq i$ in the exchange terms is restricted to equal spins because the product $\langle\varphi_j|W_{ij}|\varphi_i\rangle$ is proportional to $\delta_{s_i s_j}$. This is in accord with our experience from electron scattering, where quantum interference also occured only for equal spin directions. The expectation value of the total energy for the Hartree–Fock wave function thus becomes

$$\begin{aligned}\langle\psi_A|H|\psi_A\rangle &= \sum_{i=1}^{N}\langle\varphi_i|(T_i+V_i)|\varphi_i\rangle + \frac{e^2}{2}\sum_{i\neq j}\iint |\varphi_i(\vec{x})|^2\frac{1}{|\vec{x}-\vec{x}\,'|}|\varphi_j(\vec{x}\,')|^2\, d^3x d^3x' \\ &\qquad -\frac{e^2}{2}\sum_{i\neq j}\delta_{s_i s_j}\iint \frac{\varphi_i^*(\vec{x})\varphi_j(\vec{x})\;\varphi_j^*(\vec{x}\,')\varphi_i(\vec{x}\,')}{|\vec{x}-\vec{x}\,'|}\, d^3x d^3x' \qquad (10.55)\\ &= \sum_{i=1}^{N}\varepsilon_i - \frac{e^2}{2}\sum_{i\neq j}\iint |\varphi_i(\vec{x})|^2\frac{1}{|\vec{x}-\vec{x}\,'|}|\varphi_j(\vec{x}\,')|^2\, d^3x d^3x' \\ &\qquad +\frac{e^2}{2}\sum_{i\neq j}\delta_{s_i s_j}\iint \frac{\varphi_i^*(\vec{x})\varphi_j(\vec{x})\;\varphi_j^*(\vec{x}\,')\varphi_i(\vec{x}\,')}{|\vec{x}-\vec{x}\,'|}\, d^3x d^3x'. \qquad (10.56)\end{aligned}$$

for solutions to the Hartree–Fock equation.

Summarizing, the Hartree–Fock approximation is a variational procedure so that the result can only be as close to the correct ground state wave fuction as one can get with an antisymmetrized product wave function in the much larger N-particle Hilbert space $\mathcal{H}_N$. On top of this, the Hartree–Fock equation (10.54) can only be solved approximately, which is usually done by a numerical interation. A resonable starting point for this interation can be constructed by following the semiclassical ideas of L.H. Thomas (1926) and E. Fermi (1928).

The Thomas–Fermi method. This method is based on the idea that the uncertainty principle implies that each particle occupies at least a volume $d^3x d^3p \approx \hbar^3$ in phase space. According to Pauli's exclusion principle N electrons hence occupy a volume of at least $\hbar^3\, N/2$ where the factor $1/2$ accounts for the two allowed spin projections. For a classical Hamilton function $H(\vec{x},\vec{p}) = \frac{\vec{p}^2}{2m} + V(\vec{x})$ we can then assume that the density of states $f(\vec{x},\vec{p})$ in phase space has the constant value $2/\hbar^3$ up to the energy level E_F that is required for accomodating N particles and vanishes for higher energies $H(\vec{x},\vec{p}) > E_F$. This energy level is refered to as *Fermi surface* because if bounds the volume in phase space that is occupied by the N particles. The momentum $p_F(r) = \sqrt{2m(E_F - V(r))}$ is called *Fermi momentum*, which obviously is position dependent. These ideas have a wide range of application including nuclear physics and an easy derivation of an estimate for the size of neutron stars.

Density functional theory. Modern computations in chemistry and in solid state physics are mostly based on this improvement of the Hartree–Fock method, which is based on the theorem of P. Hohenberg and W. Kohn theorem (1964) stating that the ground state energy can

be expressed as a functional $E[\rho]$ of the electron density $\rho(\vec{x})$. This functional is a sum of the Hartree energy $H_{Hartree}$, the exchange energy $H_{exchange}$ and the correlation energy $H_{correlation}$, which accounts for the fact that the correct ground state wave function is not of the anti-symmetrized product form. In less fancy terms, the correlation energy is simply all the rest. Unfortunately, the correlation functional is not known explicitly, but it is determined by the Kohn–Sham equation (1965), which serves as the analog of the Hartree equation. Popular approaches to the solution of that equation go under the names LDA (local density approximation) and LSD (local spin density approximation).

10.4 Occupation number representation

We have seen that it is often useful to describe states $|\psi\rangle \in \mathcal{H}_N$ of systems with a large number N of identical particles in terms of a basis of symmetrized or antisymmetrized products states whose factors belong to a fixed basis $|\varphi_i\rangle$ of the one-particle Hilbert space $\mathcal{H}$. Such a product state is then uniquely described by the occupation numbers n_i of the states $|\varphi_i\rangle$, where n_i counts how often the vector $|\varphi_i\rangle$ occurs as a factor in the product state. Since $N = \sum n_i$ only a finite number of occupation numbers is nonzero and for fermions n_i is restricted to the values 0 and 1. A basis vector of $\mathcal{H}_N$ can hence be simply be characterized by the collection of nonzero occupation numbers

$$|n_{i_1} \dots n_{i_L}\rangle \qquad \text{with} \qquad \textstyle\sum_{l=1}^{L} n_{i_l} = N, \tag{10.57}$$

where it is usually clear from the context whether we are talking about bosons or fermions.

The Fock space. It is now a small step to drop the condition of having a fixed particle number N. Nonconstant particle numbers are needed for many purpuses, like the description of grand canonical ensembles in statistical mechanics, particle creation in relativistic quantum field theory, but also for the description of photons or phonons, which can be created with little energy in quantum optics of solid state physics. The appropriate Hilbert space is now the infinite direct sum of the N-particle spaces for $N = 0, 1, 2, \dots$, which is called *Fock space*

$$\mathcal{F} = \mathcal{H}_0 \oplus \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3 \oplus \dots \tag{10.58}$$

where $\mathcal{H}_1 = \mathcal{H}$ and $\mathcal{H}_0$ is the 0-dimensional Hilbert space $\mathbb{C}$. In a sense, the bosonic Fock space $\mathcal{F}^{(B)}$ is much larger than the fermion one $\mathcal{F}^{(F)}$, because the occupation numbers are not restricted to $n_{i_l} \leq 1$.[3]

Creation and annihilation operators. All operators that we inherited from the single-particle Hilbert space can change occupation numbers but do not change N. Once we have introduced the Fock space we should also consider operators that allow us to change the total

[3]One the other hand, they are of equal size in the sense that both remain seperable if $\mathcal{H}$ is separable.

number of particles. We first discuss the case of bosons. In order to increase a particular occupation number n_i by one we tensor with $\sqrt{N+1}\,|\varphi_i\rangle$ and symmetrize. The resulting map from $\mathcal{H}_N^{(B)}$ to $\mathcal{H}_{N+1}^{(B)}$ is called creation operator $a_i^\dagger$ and it acts as

$$a_i^\dagger\,|n_{i_1}\dots n_i\dots n_{i_L}\rangle = \sqrt{N+1}\,\mathcal{S}(|\varphi_i\rangle\otimes|n_{i_1}\dots n_i\dots n_{i_L}\rangle) \tag{10.59}$$

$$= \sqrt{n_i+1}\,|n_{i_1}\dots(n_i+1)\dots n_{i_L}\rangle \tag{10.60}$$

on the normalized basis vectors, where the factor $\sqrt{n_i+1}$ in the second line follows from

$$\mathcal{S}\left(|q_0\rangle\otimes\sqrt{\tfrac{N!}{\prod n_i!}}\,\mathcal{S}\,|q_1\dots q_N\rangle\right) = \sum_1^{(N+1)!}\tfrac{\mathcal{P}_\pi}{(N+1)!}\sqrt{\tfrac{N!}{\prod n_i!}}\sum_1^{N!}\tfrac{\mathcal{P}_{\pi'}}{N!}|q_0q_1\dots q_N\rangle \tag{10.61}$$

$$= \sqrt{\tfrac{N!}{\prod n_i!}}\sum_1^{(N+1)!}\tfrac{\mathcal{P}_\pi}{(N+1)!}|q_0q_1\dots q_N\rangle = \sqrt{\tfrac{n_i+1}{N+1}}\,|q_0q_1\dots q_N\rangle_S \tag{10.62}$$

in the notation of formula (10.25). The adjoint of the creation operator in Fock space is called annihilation operator and its action on the basis vectors is

$$a_i\,|n_{i_1}\dots n_i\dots n_{i_L}\rangle = \sqrt{n_i}\,|n_{i_1}\dots(n_i-1)\dots n_{i_L}\rangle. \tag{10.63}$$

With (10.60) and (10.63) it is now easily verified that creation operators $a_i^\dagger$ and annihilation operators a_j commute for $i\neq j$, while they satisfy the same algebraic relations as those of the harmonic oscillator if $i=j$. All commutation relations are summarized in the formulas

$$\boxed{[a_i,a_j]=0,\qquad [a_i,a_j^\dagger]=\delta_{ij},\qquad [a_i^\dagger,a_j^\dagger]=0.} \tag{10.64}$$

Some authors use b_i for bosonic and a_i for fermionic annihilation operators but we prefer to keep the a of the harmonic oscillator for the bosonic case.

Fermions. The same construction can now be applied to fermions. Since now all nonzero occupation numbers are 1 we can drop the redunant n from the notation and denote the states by $|i_1\dots i_L\rangle$. Accordingly the normalizations simplify. But on the other hand orderings and signs have to be treated more carefully because the sign of a state changes for every transposition,

$$|i_1\dots i_k\dots i_l\dots i_L\rangle = -|i_1\dots i_l\dots i_k\dots i_L\rangle. \tag{10.65}$$

The fermionic creation operators $b_i^\dagger$ are again defined by tensoring with $\sqrt{N+1}\,|\varphi_i\rangle$, but now with a subsequent antisymmetrization so that we obtain

$$b_i\,|ii_1i_2\dots i_L\rangle = |i_1i_2\dots i_L\rangle,\qquad b_i^\dagger\,|i_1i_2\dots i_L\rangle = \begin{cases}|i\,i_1i_2\dots i_L\rangle & \text{if } n_i=0\\ 0 & \text{if } n_i=1\end{cases}, \tag{10.66}$$

and b_i vanishes if $n_i=0$. For fermions the sign changes whenever we transpose two positions. We hence obtain the same algebra as for bosons, except that now all commutators are replaced by anticommutators

$$\boxed{\{b_i,b_j\}=0,\qquad \{b_i,b_j^\dagger\}=\delta_{ij},\qquad \{b_i^\dagger,b_j^\dagger\}=0.} \tag{10.67}$$

These formulas are easily verified for our basis of the Fock space by using the definitions (10.66).

Operators and occupation numbers. For bosons, as well as for fermions, every basis vector of the Fock space can now be obtained by repeated application of creation operators from the Fock vacuum $|0\rangle$, for which all occupation numbers are zero and which is hence a normalized state in $\mathcal{H}_0$. As for the harmonic oscillator we can define the occupation number operators

$$N^{(B)} = \sum_k a_k^\dagger a_k, \qquad N^{(F)} = \sum_k b_k^\dagger b_k, \tag{10.68}$$

which count the occupation numbers for bosons and fermions, respectively. It is a beautiful feature of this formalism that we can rewrite all our previous operators for identical particles in terms of creation and annihilation operators [Hittmair]. For a single-particle operator like $V = \sum_{i=1}^N V(x_i)$ the formula is

$$V = \sum_{ij} \langle i|V|j\rangle \; a_i^\dagger a_j. \tag{10.69}$$

For two-particle operators like the electron-electon interaction $W = \frac{1}{2}\sum_{i\neq j} W_{ij}$ one can show that

$$W = \frac{1}{2}\sum_{ijkl} \langle ij|W|kl\rangle \; a_i^\dagger a_j^\dagger a_l a_k. \tag{10.70}$$

An important special case of this is the Hamilton operator of non-interacting particles, which as a one-particle operator so that

$$H^{(1)} = \sum_i \hbar\omega_i \; a_i^\dagger a_i \tag{10.71}$$

where $H|\varphi_i\rangle = \hbar\omega_i|\varphi_i\rangle$. From the analogy with the harmonic oscillator we expect an additional contribution from the zero point energies. A constant is, however, not a one-particle but rather a zero-particle operator, whose contribution could be recovered as $H^{(0)} = \langle 0|H|0\rangle$. In any case, a constant contribution to the Hamilton function is unobservable in quantum mechanics.

10.4.1 Quantization of the radiation field

Our first application of the occupation number formalism is the quantization of the electromagnetic field in an intuitive and simplified form where we are only interested in radiation and exclude Coulomb interactions. In the *absence of charges* the Maxwell equations for the electromagnetic potentials are

$$\Box A_\mu - \partial_\mu \partial_\nu A^\nu = 0. \tag{10.72}$$

Imposing the *Coulomb gauge*

$$\nabla \vec{A}(\vec{x}, t) = 0 \tag{10.73}$$

the equation for the scalar potential $\phi = A^0$ becomes

$$\Box\,\phi - \frac{1}{c^2}\partial_t^2\phi \;=\; -\Delta\phi \;=\; 0, \tag{10.74}$$

which contains no time derivative so that ϕ is not dynamical and can be set to $\phi = 0$. This gauge is also called *radiation gauge* and the vector potential equation becomes

$$\Box\vec{A}(\vec{x},t) = \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\vec{A}(\vec{x},t) - \Delta\vec{A}(\vec{x},t) = 0 \tag{10.75}$$

with

$$\vec{E}(\vec{x},t) = -\frac{1}{c}\frac{\partial}{\partial t}\vec{A}(\vec{x},t), \qquad \vec{B}(\vec{x},t) = \nabla\times\vec{A}(\vec{x},t). \tag{10.76}$$

Its solutions $\vec{A}(\vec{x},t)$ can be written as superpositions of plane waves $e^{i(\vec{k}\vec{x}-\omega t)}$ with $\omega = c\,|\vec{k}|$.

To simplify the subsequent discussion we put our system into a large box with volume $V = L^3$ and impose periodic boundary conditions (physical boundary conditions would be inconsistent with $\phi = 0$ because of the presence of surface charges). For later convenience the coefficients in the Fourier series of the vector potential $\vec{A}(\vec{x},t)$ are normalized such that

$$\vec{A}(\vec{x},t) = \sum_{\vec{n}\in\mathbb{Z}^3}\sqrt{\tfrac{2\pi\hbar c^2}{L^3\omega}}\left(\vec{a}_k\; e^{i(\vec{k}\vec{x}-\omega t)} + \vec{a}_k^{\,*}\; e^{-i(\vec{k}\vec{x}-\omega t)}\right) \qquad \text{with} \qquad \vec{k} \equiv \tfrac{2\pi}{L}\vec{n}. \tag{10.77}$$

The constant contribution for $\vec{n} = \vec{0}$ can be omitted because it does not contribute to $\vec{E}$ or $\vec{B}$. For $\vec{n} \neq \vec{0}$ the coefficient vectors $\vec{a}(k)$ have to be transversal $\vec{k}\cdot\vec{a}_k = 0$. This condition is solved by linear combinations of two transversal polarization vectors $\vec{e}_{k\alpha}$ with $\alpha = 1,2$ which we choose to be orthonormal,

$$\vec{k}\cdot\vec{e}_{k\alpha} = 0, \qquad \vec{e}_{k\alpha}\cdot\vec{e}_{k\alpha'} = \delta_{\alpha\alpha'} \qquad \Rightarrow \qquad \sum_{\alpha=1,2}(e_{k\alpha})_i(e_{k\alpha})_j = \delta_{ij} - \tfrac{k_i k_j}{k^2} \equiv \delta_{ij}^T. \tag{10.78}$$

With the expansion $\vec{a}_k = \sum_{\alpha=1}^{2} a_{k\alpha}\,\vec{e}_{k\alpha}$ our ansatz becomes

$$\vec{A}(\vec{x},t) = \sum_{\substack{\vec{n}\in\mathbb{Z}^3\\ \vec{n}\neq\vec{0}}}\sum_{\alpha=1,2}\sqrt{\frac{2\pi\hbar c^2}{V\omega}}\left(a_{k\alpha}e^{i(\vec{k}\vec{x}-\omega t)} + a_{k\alpha}^* e^{-i(\vec{k}\vec{x}-\omega t)}\right)\vec{e}_{k\alpha} \tag{10.79}$$

In terms of the vectorpotential the energy of our radiation field is

$$H = \frac{1}{8\pi}\int_V d^3x\,\left(\vec{E}^2 + \vec{B}^2\right) = \frac{1}{8\pi}\int_V d^3x\,\left(\frac{1}{c^2}\left(\frac{\partial\vec{A}}{\partial t}\right)^2 + \left(\nabla\times\vec{A}\right)^2\right) \tag{10.80}$$

Inserting the ansatz (10.79) into this expression the integral can be evaluated and we obtain

$$H = \frac{1}{2}\sum_{\vec{n},\alpha}\hbar\omega\,(a_{k\alpha}a_{k\alpha}^* + a_{k\alpha}^* a_{k\alpha}) \qquad \text{with} \qquad \vec{k} = \frac{2\pi}{L}\,\vec{n}, \quad \omega = c\,|\vec{k}|. \tag{10.81}$$

This form of the Hamilton function reminds us of the harmonic oscillator and also of the Hamiltonian (10.71) of free particles in the occupation number representation. It is hence

natural to interpret the Fourier coefficients $a_{k\alpha}$ as anihilation operators and to replace their complex conjugates by the corresponding creation operators

$$a^*_{k\alpha} \;\rightarrow\; a^\dagger_{k\alpha}, \qquad [a_{k\alpha}, a^\dagger_{k'\alpha'}] = \delta_{\alpha,\alpha'}\delta_{\vec{k},\vec{k}'}. \tag{10.82}$$

This procedure is called *field quantization*, or *second quantization* in contrast to the first quatization in which particle trajectories were replaced by wave functions. Electromagnetism is described by a field already at the classical level, and its quantization is performed by replacing (the Fourier modes of) the classical field by operators.[4]

With the identification of the Fourier coefficients of the electromagnetic field in a box with creation and annihilation operators we interpret each oscillation mode as a harmonic oscillator. According to the Hamilton function (10.81) $\vec{E}^2$ provides the kinetic term and $\vec{B}^2$ plays the role of a harmonic potential. For finite volume we have a discretely infinite sum over polarizations α and wave vectors $\vec{k} \in \frac{2\pi}{L}\mathbb{Z}^3$ of harmonic oscillators. For $L \to \infty$ the Fourier series turns into a Fourier integral and we obtain a continuum of such oscillators.

10.4.2 Interaction of matter and radiation

We are now in the position to compute electromagnetic transitions in atomic physics. Our starting point is Fermi's golden rule

$$P_{i\to f} \;=\; \frac{2\pi}{\hbar}\,\delta(E_i - E_f)\,|\langle f|H_I|i\rangle|^2, \tag{10.86}$$

which we derived in chapter 6. It expresses the transition probability $P_{i\to f}$ per unit time from an initial state $|i\rangle$ to a final state $|f\rangle$ in terms of the matrix element of the interaction Hamiltonian H_I in the interaction picture. We will only work out the leading approximation and neglect

[4] In a more deductive approach, the starting point for the quantization of the electromagnetic field $\vec{A}(x,t)$ is the Lagrange function $L = \int d^3x\mathcal{L}$ with Lagrange density

$$\mathcal{L} = \frac{1}{8\pi}(\vec{E}^2 - \vec{B}^2) = \frac{1}{8\pi}\left(\frac{1}{c^2}\dot{\vec{A}}^2 - (\nabla\times\vec{A})^2\right). \tag{10.83}$$

The canonical momenta π_i conjugate to the dynamical variables A_i are

$$\pi_i = \frac{\partial\mathcal{L}}{\partial\dot{A}_i} = \frac{1}{4\pi c^2}\dot{A}_i = -\frac{1}{4\pi c}E_i. \tag{10.84}$$

By inserting the expansion (10.79) one can show that the commutation relations (10.82) for the Fourier coefficients are equivalent to

$$[A_i(\vec{x},t), \pi_j(\vec{y},t)] = i\hbar\,\delta^T_{ij}(\vec{x}-\vec{y}), \tag{10.85}$$

where $\delta^T_{ij}(\vec{x}-\vec{y})$ is the Fourier transform of the transversal δ-function that was defined in eq. (10.78). The restriction of the δ-function to transversal degrees of freedom is required by the Coulomb gauge condition (10.73). This can be shown to follow from the quantization prescription in the presence of constraints that was developed by Dirac and Bergmann [Dirac]. A theoretical framework that enabled a Lorentz covariant canonical quantization of the full electromagnetic field was only developed in the 1970s.

magnetic interactions with the electron's spin and the order $(eA)^2$ term in the Hamiltonian

$$H = \frac{1}{2m}(\vec{p} - \frac{e}{c}\vec{A})^2 + \ldots \qquad \Rightarrow \qquad H_I = -\frac{e}{2mc}(\vec{p}\vec{A} + \vec{A}\vec{p}) + \ldots, \tag{10.87}$$

where $\vec{p}\vec{A} = \vec{A}\vec{p}$ in the Coulomb gauge.

In the interaction picture the initial and final states are energy eigenstates of the Hamiltonian $H_{mat} + H_{em}$, which factorize into a matter part with energy eigenvalue ε and a photon state in the occupation number representation,

$$|i\rangle = |\lambda_i\rangle \otimes |n_i\rangle, \qquad |f\rangle = |\lambda_f\rangle \otimes |n_f\rangle, \qquad H_{mat}\,|\lambda\rangle = \varepsilon|\lambda\rangle \tag{10.88}$$

where $\lambda = (\varepsilon, \ldots)$ specifies the energy eigenstate of the electron. For the emission or absorption of a single photon it is sufficient to specify the occupation number $n_f = n_i \pm 1$ for the momentum $\vec{k}$ and the polarization α for which we want to compute the probability (10.86). Inserting $-\frac{e}{mc}\vec{p}\vec{A}$ with

$$\vec{A}(\vec{x}, 0) = \sum_{\substack{\vec{n}\in\mathbb{Z}^3 \\ \vec{n}\neq\vec{0}}} \sum_{\alpha=1,2} \sqrt{\frac{2\pi\hbar c^2}{V\omega}} \left(a_{k\alpha} e^{i\vec{k}\vec{x}} + a^\dagger_{k\alpha} e^{-i\vec{k}\vec{x}} \right) \vec{e}_{k\alpha} \tag{10.89}$$

and the matrix elements

$$\langle n_i - 1|a|n_i\rangle = \sqrt{n_i}. \qquad \langle n_i + 1|a^\dagger|n_i\rangle = \sqrt{n_i + 1} \tag{10.90}$$

we obtain the *absorption* probability

$$P_{i\to f} \;=\; \frac{4\pi^2 e^2}{m^2 V\omega}\, \delta(\varepsilon_i + \hbar\omega - \varepsilon_f)\; n_{i,k\alpha}\; \left| \langle \varepsilon_f |\; \vec{p}\,\vec{e}_{k\alpha}\; e^{i\vec{k}\vec{x}}\; |\varepsilon_i\rangle \right|^2 \tag{10.91}$$

and the *emission* probability

$$P_{i\to f} \;=\; \frac{4\pi^2 e^2}{m^2 V\omega}\, \delta(\varepsilon_i - \hbar\omega - \varepsilon_f)\; (n_{i,k\alpha} + 1)\; \left| \langle \varepsilon_f |\; \vec{p}\,\vec{e}_{k\alpha}\; e^{-i\vec{k}\vec{x}}\; |\varepsilon_i\rangle \right|^2 . \tag{10.92}$$

For $n_i = 0$ we obtain the probability for spontaneous emission, while the probabilities for absorption and induced emission are proportional to the occupation number.

In the dipol approximation the exponentials $e^{i\vec{k}\vec{x}}$ are replaced by 1. Since $[\vec{x}, H] = \frac{i\hbar}{m}\vec{p}$ the expectation value of $\vec{p}$ is related to the dipol moment by

$$\langle \varepsilon_f|\vec{p}|\varepsilon_i\rangle = \frac{im}{\hbar} \frac{\langle \varepsilon_f|\vec{x}|\varepsilon_i\rangle}{\varepsilon_f - \varepsilon_i} \tag{10.93}$$

If we want to compute the life time for an excited state we need to integrate over all momenta and to sum over all polarizations. Since $\vec{k} = \frac{2\pi}{L}\vec{\mu}$ the limit $L \to \infty$ amonts to

$$\frac{1}{V} \sum_{\vec{\mu}\,\in\,\mathbb{Z}^3} \qquad \to \qquad \int \frac{d^3k}{(2\pi)^3} \tag{10.94}$$

The energy conserving δ-function leads to a finite integral over a sphere in momentum space.

10.4.3 Phonons and quasiparticles

As one can hear by knocking on a door the lowest vibration frequencies of a solid are far below the frequencies ω that correspond to excitations of single atoms or molecules. This implies that the smallest energy quanta $\hbar\omega$ that are available in a solid for quantum mechanical processes correspond to excitations that involve the collective motion of a large number of atoms. The particle-like degrees of freedom that enter such a process are called *quasi-particles*. In the case of lattice vibrations of a solid the quasi-particles are called phonons.

For small temperatures anharmonic effects may be neglected and the Hamiltonian has the form

$$H = \sum_{l=1}^{L} \frac{p_l^2}{2m_l} + \frac{1}{2} \sum_{k,l=1}^{L} K_{kl} x_k x_l, \tag{10.95}$$

where m_l are the effective masses of the elementary degrees of freedom and the matrix K_{kl} describes the harmonic forces. Diagonalization yields the normal modes, which correspond to decoupled harmonic oscillators. In terms of the respective creation and annihilation operators the quantum system is hence described by a Hamilton function

$$H = \sum \hbar\omega_i (a_i^\dagger a_i + \frac{1}{2}). \tag{10.96}$$

The number operators $N_i = a_i^\dagger a_i$ count the occupation numbers of the phonon states with wave vector $\vec{k}$ and their dispersion relation $\omega(\vec{k})$ is given by the speed of sound. Many physical properties of solids such as specific heat and thermal conductivity can be describe in terms of phonons, which are bosons with spin zero.

Phonons in a solid are quite analogous to electromagnetic modes in a cavity and we can read the previous quantization procedure backwards and construct a phonon field $\phi(\vec{x}, t) \sim \sum (a_k e^{i(\vec{k}\vec{x}-\omega t)} + a_k^\dagger e^{-i(\vec{k}\vec{x}-\omega t)})$. For certain quantities non-linear interaction terms become important. Creation and annihilation of phonons requires cubic terms ϕ^3, which automatically have a excess of a creation operator or an annihilation operator and hence change the phonon number. Phonon-phonon scattering is described by interaction terms ϕ^4, whose expansion in creation and annihilation operators assumes the form of two-particle operator.

Chapter 11

WKB and the path integral

In this chapter we discuss two reformulations of the Schrödinger equations that can be used to study the transition from quantum mechanics to classical mechanics. They lead to new approximation techniques called WKB and stationary phase approximation, respectively. These methods are called *semiclassical* [Brack-Bhaduri] as they are based on the data of classical trajectories and add some phase information in order to account for interference phenomena.

Recollections from classical mechanics. According to the variational principle the equations of motion $\delta L/\delta q = 0$ are obtained by extremizing the *action functional* $S[\gamma] = \int_\gamma dt\ L(q^i, \dot{q}^i)$. *Hamilton's principal function*

$$S(q'', q', t'' - t') = \int_{t'}^{t''} L(q, \dot{q})\, dt \tag{11.1}$$

is the value of $S[\gamma]$ for a classical trajectory γ from q' to q'' regarded as a function of the initial and the final coordinates and of the time $t = t'' - t'$ needed for the travelling. The dependency of $S(q'', q', t)$ on its arguments can be shown to be given by

$$\frac{\partial S}{\partial q''} = p'', \qquad \frac{\partial S}{\partial q'} = -p', \qquad \frac{\partial S}{\partial t} = -E \tag{11.2}$$

or, equivalently, $\delta S = p''\delta q'' - p'\delta q' - E\delta t$ and it satisfies the Hamilton–Jacobi equation

$$H(q''_1, \ldots, q''_N, \frac{\partial S}{\partial q''_1}, \ldots, \frac{\partial S}{\partial q''_N}) + \frac{\partial S}{\partial t} = 0. \tag{11.3}$$

The dependency of S on time can be traded for a dependency on the energy E that is available for the trip from q' to q'' by a Legendre transformation

$$\tilde{S}(q'', q', E) = S + Et = S + \int_{t'}^{t''} dt\, H = \int_{t'}^{t''} dt(L + H) = \int_{t'}^{t''} p_i \dot{q}^i\, dt = \int_{q'}^{q''} p_i\, dq^i. \tag{11.4}$$

$\tilde{S}(q'', q', E) = \int p_i\, dq^i$, whose variations are $\delta\tilde{S}(q'', q', E) = p''\,\delta q'' - p'\,\delta q' + t\,\delta E$, is sometimes called *simplified action* to destinguish it from the action (11.1), while textbooks on semiclassics

often use the name Hamilton's principal function and the letter R for (11.1), and reserve the word *action* and the letter S for (11.4).

While S and $\tilde{S}$ are well-defined as long as q'' and the trajectory γ stay in a neighborhood of q', ambiguities can arise globally because several classical trajectories may lead from q' to q'' with the same energy (consider, for example, the free motion on the surface of a sphere for which there are generically two extremal paths connecting two points). Moreover, a whole family of different trajectories originating from the same point q' can meet along a *caustic*. Itersections of classical trajectories from q' to q'' with caustics are called *conjugate points* (on a sphere the south pole is a conjugate point for trajectories originating from the north pole).

In one dimension the dynamics of an autonomous system is quite simple because energy conservation fixes the momentum $p(x) = \pm\sqrt{(E - V(x))/2m}$ up to a sign. For a two-dimensional *rectangular billard*, i.e. a particle moving in a rectangular box with perfectly reflecting walls, there are already infinitely many trajectories from q' to q'', but additional constants of motion (the momentum components squared) make the system integrable. A dynamical system is called *integrable* if there are d constants of motion $I_1, \ldots, I_d$ with vanishing Poisson brackets $\{I_i, I_j\}_{PB}$, where d is the dimension of the configuration space. For such systems there exists a canonical transformation to *action-angle* variables (or *torus* variables) that brings the Hamiltonian to the form $H(q^i, p_i) \to \tilde{H}(\phi, I) = \tilde{H}(I)$. In a rectangular billard the two conserved quantities are the squares of the momentum components $I_x = p_x^2$ and $I_y = p_y^2$, and for fixed I_x and I_y there are 4 possible directions of the momentum. In a stadium-shaped billard, however, the motion becomes chaotic (non-integrable and extremely sensitive to initial conditions), which makes the WKB approach inadequate. Path integral techniques have a wider range of applicability.

11.1 WKB approximation

This semi-classical method was named after G. Wentzel, A. Kramers and L. Brillouin, who developed it independently in 1926. It is the quantum analog of the Sommerfeld-Runge procedure for the transition from wave optics to ray optics and hence also called eiconal approximation. We parametrize the wave function

$$\psi(\vec{x}, t) = A(\vec{x}, t) e^{\frac{i}{\hbar} S(\vec{x}, t)} \tag{11.5}$$

by its the real amplitude A and its real phase S and search solution to the Schrödinger equation $i\hbar\dot{\psi} = H\psi$ with $H = -\frac{\hbar^2}{2m}\Delta + V(\vec{x})$. Since $\hbar^2 \Delta e^{\frac{i}{\hbar}S} = (i\hbar\Delta S - (\nabla S)^2) e^{\frac{i}{\hbar}S}$ we find

$$-\hbar^2\Delta\psi = \left(-\hbar^2\Delta A - 2i\hbar\nabla A\nabla S - i\hbar A\Delta S + A(\nabla S)^2\right) e^{\frac{i}{\hbar}S}, \tag{11.6}$$

$$i\hbar\partial_t\psi = (i\hbar\dot{A} - A\dot{S}) e^{\frac{i}{\hbar}S}. \tag{11.7}$$

Dropping the overall factor $\psi = Ae^{\frac{i}{\hbar}S}$ the real part of the Schrödinger equation thus becomes

$$\frac{(\nabla S)^2}{2m} + V + \frac{\partial S}{\partial t} = \hbar^2 \frac{\Delta A}{2mA}. \tag{11.8}$$

The left-hand-side exactly corresponds to the Hamilton-Jacobi equation if we identify S with the classical action, while the right-hand-side is a quantum correction of order $\mathcal{O}(\hbar^2)$ that is neglected in the WKB approximation.

For time-independent potentials stationary solutions are obtained with the separation ansatz

$$\psi(\vec{x},t) = u(\vec{x})e^{-\frac{i}{\hbar}Et} \qquad \Rightarrow \qquad u(\vec{x}) = A(\vec{x})e^{\frac{i}{\hbar}(S(\vec{x},t)+Et)} = A(\vec{x})e^{\frac{i}{\hbar}\tilde{S}(\vec{x})} \tag{11.9}$$

with $\tilde{S} = S + Et$. The stationary Schrödinger equation thus becomes equivalent to

$$(\nabla\tilde{S})^2 = 2m(E-V) + \hbar^2\frac{\Delta A}{A}, \tag{11.10}$$

$$0 = \Delta\tilde{S} + 2\nabla\tilde{S}\frac{\nabla A}{A}, \tag{11.11}$$

where the second equation is just the imaginary part of (11.6) because $\dot{A} = 0$. We now restrict our discussion to one-dimensional problems. Then the equation for the imaginary part becomes

$$\frac{1}{2}\frac{\tilde{S}''}{\tilde{S}'} + \frac{A'}{A} = 0 \qquad \Rightarrow \qquad \frac{d}{dx}\left(\frac{1}{2}\log\frac{d\tilde{S}}{dx} + \log A\right) = 0 \qquad \Rightarrow \quad A = c\left(\frac{d\tilde{S}}{dx}\right)^{-\frac{1}{2}} \tag{11.12}$$

with an integration constant c. For the real part (11.10) of the Schrödinger equation we use the WKB approximation and drop the term of order $\mathcal{O}(\hbar^2)$ so that

$$\left(\frac{d\tilde{S}}{dx}\right)^2 = 2m(E-V). \tag{11.13}$$

This equation is easily integrated to

$$\tilde{S}(x) = \pm\int^x dx'\sqrt{2m(E-V(x'))} = \pm\int^x dx'\; p(x') \tag{11.14}$$

where $p(x) = \sqrt{2m(E-V(x))}$ is the classical expression for the momentum. The WKB wave function thus becomes a sum of a left-moving and a right-moving wave

$$u(x) = \frac{c_+}{\sqrt{p(x)}}\exp\left\{+\frac{i}{\hbar}\int^x dx'\,p(x')\right\} + \frac{c_-}{\sqrt{p(x)}}\exp\left\{-\frac{i}{\hbar}\int^x dx'\,p(x')\right\} \tag{11.15}$$

with amplitudes $A \sim c_\pm/\sqrt{p(x)}$. This is easy to interpret because the probability density A^2 is inverse proportional to the velocity for a conserved particle flux. For bound state solutions we can always choose $u(x)$ to be real so that, by an approriate choice of the constants $c_\pm$,

$$u(x) = \frac{c}{\sqrt{p(x)}}\cos\left(\frac{1}{\hbar}\int_a^x dx'\,p(x') + \varphi(a)\right) \tag{11.16}$$

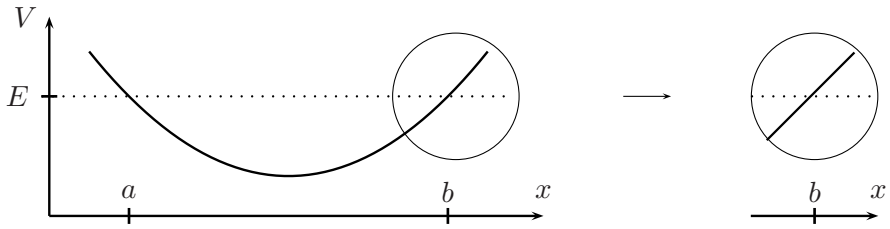


Figure 11.1: Soft wall approximation for a potential with turning points $x = a$ and $x = b$.

with a phase $\varphi(a)$ depending on the choice of the lower limit $x = a$ of the integration domain.

Validity of the WKB approximation. Intuitively we can expect that the WKB approximation is good if the variation of the amplitude is small over distances of the order of the wave length $\lambda(x) \approx 2\pi\hbar/p(x)$. More precisely the condition is

$$\hbar^2 \frac{1}{A}\frac{d^2A}{dx^2} \ll \left(\frac{d\tilde{S}}{dx}\right)^2. \tag{11.17}$$

Using (11.12) this can be written as

$$\left|\frac{dp}{dx}\right| \ll \frac{1}{\hbar}p^2 = 2\pi\frac{p}{\lambda}. \tag{11.18}$$

In regions where this condition is valid we may trust the WKB wave function. In particular, we can expect good results for high energies and short wave lengths.

Soft reflection and Airy functions. The WBK approximation certainly breaks down at classical turning points where $V(x) = E$ so that $p(x) \to 0$ and the amplitude $A(x)$ diverges. While this local effect will be negligable for high energies we can improve our results by solving the Schrödinger equation exactly at the zeros of $E - V(x)$. For smooth potentials we can use a linear approximation as shown in figure 11.1. The exact solution for the linearized potential is then compared with the WKB solution. At the right turning point $x = b$ the Schrödinger equation becomes

$$u'' = \tfrac{2m}{\hbar^2}(V - E)u \approx \tfrac{2mV'(b)}{\hbar^2}(x - b)u. \tag{11.19}$$

After a change of variables of the form $z = c(x - b)$ we hence have to solve the equation

$$w'' - zw = 0. \tag{11.20}$$

The solutions are linear combinations of the Airy functions,

$$w(z) = \alpha Ai(z) + \beta Bi(z), \tag{11.21}$$

which can be defined in terms of Bessel functions as

$$\begin{aligned} Ai(-z) &= \tfrac{1}{3}\sqrt{z}\left(J_{-1/3}(\tfrac{2}{3}z^{3/2}) + J_{1/3}(\tfrac{2}{3}z^{3/2})\right), \\ Bi(-z) &= \sqrt{\tfrac{z}{3}}\left(J_{-1/3}(\tfrac{2}{3}z^{3/2}) - J_{1/3}(\tfrac{2}{3}z^{3/2})\right). \end{aligned} \tag{11.22}$$

For real $z \to \infty$ the asymptotics is given by

$$Ai(-z) \to \frac{\cos(\frac{2}{3}z^{3/2} - \frac{1}{4}\pi)}{\sqrt{\pi}z^{1/4}}, \qquad Ai(z) \to \frac{\exp(-\frac{2}{3}z^{3/2})}{2\sqrt{\pi}z^{1/4}}, \tag{11.23}$$

$$Bi(-z) \to -\frac{\sin(\frac{2}{3}z^{3/2} - \frac{1}{4}\pi)}{\sqrt{\pi}z^{1/4}}, \qquad Bi(z) \to \frac{\exp(\frac{2}{3}z^{3/2})}{\sqrt{\pi}z^{1/4}}. \tag{11.24}$$

Since the second Airy function $Bi(z)$ blows up at large z only $Ai(z)$ is relevant for our purposes. For real z it has the integral representation

$$Ai(z) = \tfrac{1}{\pi}\int_0^\infty \cos(t^3/3 + zt)dt, \tag{11.25}$$

which can directly be checked to satisfy the differential equation (11.20).

Since $\partial_x^2 Ai(c(x-b)) = c^2 Ai''(c(x-b)) = c^3(x-b)Ai(c(x-b))$ the Schrödinger equation with linearized potential near $x = b$ is solved by

$$u_b(x) = Ai\big(c_b(x-b)\big) \qquad \text{with} \qquad c_b = \sqrt[3]{\tfrac{2mV'(b)}{\hbar^2}}, \tag{11.26}$$

and analogously close to the left classical turning point $x = a$ by

$$u_a(x) = Ai\big(c_a(a-x)\big) \qquad \text{with} \qquad c_a = \sqrt[3]{-\tfrac{2mV'(a)}{\hbar^2}}. \tag{11.27}$$

Comparing the asymptotic form (11.23) of $u_a(x)$ to the WKB solution (11.16)

$$u_a^{WKB} = c\,\frac{\cos\left(\frac{1}{\hbar}\frac{2}{3}\sqrt{-2mV'(a)}(x-a)^{3/2} - \varphi_a\right)}{\sqrt{(a-x)2mV'(a)}} \tag{11.28}$$

for the linearized potential $V - E \approx V'(a)(x-a)$ we find a phase correction $\varphi_a = -\pi/4$, where we have choosen the lower limit for the momentum integration in (11.16) at the classical turning point. This phase shift can be interpreted as a quantum mechanical tunneling into the classically forbidden region with an effective penetration depth of one eights of the wavelengths. We will see that this leads to a correction of the Bohr–Sommerfeld quantization condition.

11.1.1 Bound states, tunneling, scattering and EKB

Bound states. The original application of the WKB approximation is the derivation of the Bohr–Sommerfeld quantization condition. If we begin our considerations, for simplicity, with Neumann boundary conditions at the classical turning points a bound state wave function of the form (11.16) with n nodes has $\frac{1}{\hbar}\int_a^b p(x)dx = n\pi$. Interpreting this standing wave solution as superposition (11.15) of a left-moving and a right-moving wave the complete action integral for the round trip is $\oint = \int_a^b + \int_b^a = 2\int_a^b$ and we obtained the Bohr–Sommerfeld quantization condition

$$\tfrac{1}{2\pi\hbar}\oint p(x)dx = n \tag{11.29}$$

for the ring integral of the momentum along a closed trajectory. Note that this integral can be interpreted as the area enclosed by the periodic orbit of the particle in its two-dimensional

phase space, i.e. in the $x-p$ plane. If the potential is smooth at both turning points, like for the harmonic oscillator, then the phase shifts φ_a and φ_b add up to an effective shift $n \to n + \frac{1}{2}$ which exactly reproduces the ground state energy of the harmonic oscillator. For Dirichlet boundary conditions (i.e. for a hard reflection at an infinitely high potential step) the wave function has a node at the classical turning point which leads to a phase shift $\varphi_a = \pm\pi/2$. In general the improved Bohr–Sommerfeld quantization formula can hence be written as

$$\oint p(x)dx = 2\pi\hbar(n + \mu/4) \qquad \text{with} \qquad \mu = N_{soft} + 2N_{hard}. \tag{11.30}$$

μ is called *Maslov index* and counts the number of classical turning points with smooth potential (soft reflection) plus twice the number of classical turning points with Dirichlet boundary conditions (hard reflections).

Tunneling. A semiclassical interpretation of the tunneling effect might be a bit far fetched because there are no classical trajectories available. In any case, however, we can use our approximate WKB solution (11.15) of the Schrödinger equation to obtain a formula for the tunneling rate. Turning the bound state potential in figure 11.1 upside down amounts, by analytic continuation, to an imaginary momentum $p(x) \to ip(x) = \sqrt{2m(V-E)}$ in the classically forbidden region of a potential hill. The appropriate solution for a tunneling process from $x < a$ to $x > b$ is hence the second term in eq. (11.15). The resulting ratio $\exp(-\frac{1}{\hbar}\int_a^b \sqrt{2m(V-E)}$ of the amplitudes at $x \to a$ and $x \to b$ squares to a tunneling rate of

$$T = \exp\left(-2\int_a^b dx\, \sqrt{\tfrac{2m}{\hbar^2}\big(V(x)-E\big)}\right) \tag{11.31}$$

in the WKB approximation, where boundary effects and back-tunneling have been neglected.

Semiclassical scattering. For central potentials we can use spherical symmetry to reduce the Schrödinger equation to a one-dimensional problem by the separation ansatz $u(\vec{x}) = R(r)Y_{lm}(\theta,\varphi)$. We recall the radial equation (8.34)

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dr^2} + V_{eff}(r)\right)\tilde{R}(r) = E\tilde{R}(r) \tag{11.32}$$

with the effective potential

$$V_{eff} = V(r) + \frac{l(l+1)\hbar^2}{2mr^2} \tag{11.33}$$

for the radial wave fuction $\tilde{R}(r) = rR(r)$, which has to vanish at the origin $\tilde{R}(0) = 0$ for normalizale $u(\vec{x})$. In order to derive a semiclassical approximation for the phase shift δ_l we compare the asymptotic ansatz (8.59)

$$R_l^{as}(k,r) = A_l(k)\frac{1}{kr}\sin\left(kr - \frac{l\pi}{2} + \delta_l(k)\right), \tag{11.34}$$

which defined the phase shift, to the radial WKB solution

$$\tilde{R}_{WKB} \sim \cos\left(\frac{1}{\hbar}\int_{r_0}^r d\rho\, \sqrt{2m(E - V_{eff}(\rho))} - \pi\frac{\mu}{4}\right) = \sin\left(\frac{1}{\hbar}\int_{r_0}^r d\rho\, p(\rho) + \frac{\pi}{2} - \pi\frac{\mu}{4}\right) \tag{11.35}$$

with classical turning point r_0. Since $\tilde{R}_{WKB}$ has to become proportional to $\tilde{R}_l^{as} = rR_l^{as}$ for $r \to \infty$ we obtain

$$\delta_l = \lim_{r\to\infty}\left(l\frac{\pi}{2} - kr + \frac{1}{\hbar}\int_{r_0}^{r} d\rho\, p(\rho) + \frac{\pi}{2} - \pi\frac{\mu_l}{4}\right) \tag{11.36}$$

$$= (l+1)\frac{\pi}{2} + \frac{1}{\hbar}\int_{r_0}^{r} d\rho\, \big(p(\rho) - p(\infty)\big) - \frac{r_0}{\hbar}p(\infty) - \pi\frac{\mu_l}{4}. \tag{11.37}$$

with $p(\infty) = \hbar k = \sqrt{2mE}$. For $l = 0$ and an attractive potential the classical turning point is $r_0 = 0$ with Dirichlet boundary conditions so that the Maslov index is $\mu_0 = 2$. For $l > 0$ the centrifugal barrier dominates stable potentials at the origin so that we have $r_0 > 0$ and soft boundary conditions with Maslov index $\mu_l = 1$. Except for $l = 0$ the WKB approximation turns out to yield good results only for large l.[1]

EKB approximation. The generalization of the WKB approach to higher-dimensional dynamical systems was named after A. Einstein (1917), L. Brillouin (1926) and J. B. Keller (1958). Already in 1917 Einstein realized that the Bohr–Sommerfeld quantization rules can only work for integrable dynamical systems because nonperiodic orbits only form a subset of measure zero in the non-integrable case, so that a quantization condition like (11.29) does not make sense if the classical trajectory of a bound particle does not form a closed curve. Einstein also gave the coordinate independent formula $\frac{1}{2\pi}\oint_{C_\alpha} p_i\, dq^i = \hbar n_\alpha$ for the Bohr–Sommerfeld quantization condition, where the d closed orbits C_α form a basis for the cycles in the phase space of the integrable system. This formula was later improved to

$$\frac{1}{2\pi}\oint_{C_\alpha} p_i\, dq^i = \hbar(n_\alpha + \mu_\alpha/4), \tag{11.38}$$

which takes into accout the Maslov indices μ_α along the orbits C_α.

11.2 The path integral

The first attempt to formulate quantum mechanics in terms of the Lagrangian goes back to P.A.M. Dirac (1933), who discovered that the overlap of position state vectors "corresponds" to the classically computed exponential $\exp(\frac{i}{\hbar}\int \mathcal{L}dt)$. But it took more than a decade until R. Feynman (1949) took up the idea and turned it into a powerful computational scheme. The central object of our interest is the propagator

$$K(x'', x'; t'' - t') = \langle x'', t''|x', t'\rangle = \langle x''|e^{-\frac{i}{\hbar}H(t''-t')}|x'\rangle \tag{11.39}$$

[1]This problem was overcome by Langer (1937), who applied the change of variables $r = e^{-x}$ that magnifies the critical region near the origin $r \to 0$ and makes WKB applicable also for small l. It turned out that the net effect of this change of coordinates amounts to the replacement $l(l+1) \to (l+\frac{1}{2})^2$ in the effective potential. The bound state problem for centrally symmetric potentials can, of course, be analyzed similarly.

which corresponds to the matrix elements of the time evolution operator in the position space basis $X|x\rangle = x|x\rangle$ (for simplicity we consider the one-dimensional situation). In the double slit experiment our intuition from ray optics tells us to superimpose the contributions of the two slits to the complete transition amplitude. More generally, the superposition principle of quantum mechanics and completness of the basis $|x\rangle$ implies

$$\langle x'', t''|x', t'\rangle = \int dx\ \langle x', t'|x, t\rangle\langle x, t|x', t'\rangle \tag{11.40}$$

for some intermediate time t with $t'' > t > t'$. By the same token we can decompose the time intervall into n small time steps and write the transition amplitude as an $(n-1)$-fold integral over all intermediate positions. For large n this integral can also be considered as an integral "over all trajectories" connecting the intermediate positions. The path integral is formally constructed as the limit $n \to \infty$ of this expression and hence can be considered as an "integral over all paths" from x' at time t' to x'' at time t''.

The building blocks of the path integral are the transition amplitudes for small time steps. Since the Hamilton operator generates time evolution

$$\langle x_2, t_2|x_1, t_1\rangle = \langle x_2|e^{-\frac{i}{\hbar}H(t_2-t_1)}|x_1\rangle \tag{11.41}$$

and since the momentum P generates translations

$$|x_2\rangle = e^{-\frac{i}{\hbar}P(x_2-x_1)}|x_1\rangle, \tag{11.42}$$

where states without explicit time dependence refer to the Heisenberg picture and time independence of H is assumed. Putting this together we find

$$\langle x_2, t_2|x_1, t_1\rangle = \langle x_1|e^{\frac{i}{\hbar}P(x_2-x_1)}e^{-\frac{i}{\hbar}H(t_2-t_1)}|x_1\rangle. \tag{11.43}$$

The momentum operator can now be evaluated if we insert a complete set $|p_1\rangle$ of momentum eigenstates between the exponentials

$$\langle x_2, t_2|x_1, t_1\rangle = \int dp_1\ \langle x_1|e^{\frac{i}{\hbar}p_1(x_2-x_1)}|p_1\rangle\langle p_1|e^{-\frac{i}{\hbar}H(t_2-t_1)}|x_1\rangle \tag{11.44}$$

In order to replace all operators by classical functions we would also like to evaluate the position and the momentum in the Hamilton operator $H(X, P)$. For this we assume that H can writen H as a sum of terms with all momentum operators on the left of all position operators. This is certainly the case for the Hamiltonian $H = \frac{P^2}{2m} + V(X)$ of a particle in a potential V. If we consider short time intervals $\delta t = t_2 - t_1$ and neglect terms of order $\mathcal{O}(\delta t^2)$ then $\exp(-\frac{i}{\hbar}H\delta t) \approx \mathbb{1} - \frac{i}{\hbar}H\delta t$ also has this property and we can evaluate X on the right and P on the left to obtain

$$\langle x_2, t_2|x_1, t_1\rangle = \int dp_1\ \langle x_1|e^{\frac{i}{\hbar}p_1(x_2-x_1)}|p_1\rangle\langle p_1|e^{-\frac{i}{\hbar}\delta t\, H(x_1,p_1)}|x_1\rangle + \mathcal{O}(\delta t^2). \tag{11.45}$$

For small δt we can write $x_2 = x_1 + \delta t\, \dot{x}_1$, and since all terms in the exponentials are mere functions we arrive at

$$\langle x_2, t_2 | x_1, t_1 \rangle = \int dp_1 \, \langle x_1 | p_1 \rangle e^{-\frac{i}{\hbar}\delta t (p_1, \dot{x}_1 - H(x_1, p_1))} \langle p_1 | x_1 \rangle + \mathcal{O}(\delta t^2) \tag{11.46}$$

$$= \int dp_1 \, e^{-\frac{i}{\hbar}\delta t (p_1, \dot{x}_1 - H(x_1, p_1)))} + \mathcal{O}(\delta t^2). \tag{11.47}$$

because $\langle x_1 | p_1 \rangle \langle p_1 | x_1 \rangle = 1$. We hence got rid of all operators and states and found a purely classical expression for the propagator for small time steps. As promissed, it conains the Lagrange function $L(q, \dot{q})$. For a Hamilton function of the form $H(x_1, p_1) = p_1^2/2m + V(x_1)$ that is quadratic in the momentum we can, moreover, perform the Gaussian momentum integration $\int dp_1$ and arrive at the final expression

$$\langle x_2, t_2 | x_1, t_1 \rangle = \sqrt{\frac{m}{2\pi i \hbar \delta t_1}} e^{i \delta t \, L(x_1, \dot{x}_1)/\hbar} + \mathcal{O}(\delta t^2). \tag{11.48}$$

for the transition amplitude in terms of the Lagrange function $L(x_1, \dot{x}_1) = \frac{1}{2} m \dot{x}_1^2 - V(x_1)$. For the path integral representation of the propagators we hence arrive at the formal expression

$$\langle x'', t'' | x', t' \rangle = \int dx_1 \ldots \int dx_{n-1} \langle x'', t'' | x_{n-1}, t_{n-1} \rangle \ldots \langle x_2, t_2 | x_1, t_1 \rangle \langle x_1, t_1 | x', t' \rangle \tag{11.49}$$

$$= \int \mathcal{D}x \; e^{\frac{i}{\hbar} \int_{t'}^{t''} L(x, \dot{x})} \tag{11.50}$$

where the measure $\mathcal{D}x$ is defined as

$$\mathcal{D}x \equiv \lim_{n \to \infty} \left(\frac{m}{2\pi i \hbar \delta t} \right)^{n/2} dx_1 dx_2 \ldots dx_{n-1}. \tag{11.51}$$

For an interpretation of the path integral we note that the integrand is a pure phase so that most contribution average themselves away due to rapidly changing phases for neighbouring paths. An exception occurs of exactly for the classical trajectories for which the phase, which is given by the action, is stationary so that we get constructive interference. In this way we get a very intuitive picture of the classical action principle because it is exactly the paths for which action is (near) extremal that contribute to the transition amplitude.

The stationary phase approximation is based on this interpretation and only keeps the leading quadratic variations of the action for evaluating the semiclassical contribution of a classical trajectory to the transition amplitude. The remaining path integral is a Gaussian integral that can be evaluated exactly. This is possible, in particular, for the Lagrangian of a free particle for which the evaluation of the Gaussian path integral yields the free propagator

$$K_{free}(x'', x'; t) = \left(\frac{m}{2i\pi\hbar} \frac{1}{t} \right)^{\frac{d}{2}} \exp \left\{ \frac{im}{2\hbar} \frac{(x'' - x')^2}{t} \right\} \tag{11.52}$$

in d dimensions. The argument of the exponent is indeed $i/\hbar$ times the action of the classical trajectory and the prefactor is due to the collective contribution of all near extremal paths.

Bibliography

[Bell] J.S. Bell, *Speakable and unspeakable in quantum mechanics* (Cambridge Univ. Press, 1987)

[Bjorken-Drell] J.D. Bjorken, S.D. Drell, *Relativistische Quantentheorie* (BI-Wiss.-Verl., Mannheim,Wien,Zürich, 1993)

[Brack-Bhaduri] M. Brack, R. K. Bhaduri, *Semiclassical physics* (Addison-Wesley, 1997)

[Bransden] B.H. Bransden, C.J. Joachain, *Quantum mechanics* (Pearson Education Limited, Edinburgh Gate, England, 2000)

[Chadan-Sabatier] K. Chadan, P. C. Sabatier, *Inverse problems in quantum scattering theory,* Texts and monographs in physics (Springer, New York 1989)

[Cohen-Tannoudji] C. Cohen–Tannoudji, B. Diu, F. Laloë, *Quantum Mechanics Vol.1&2* (Hermann, Paris, France, 1977)

[Dirac] P.A.M. Dirac, *Lectures on Quantum Mechanics* (Yeshiva Univ. Press, New York 1964)

[Dirschmid,Kummer,Schweda] Hansjörg Dirschmid, Wolfgang Kummer, Manfred Schweda, *Einführung in die mathemathischen Methoden der theoretischen Physik* (Vieweg, Braunschweig, 1976)

[Feynman] R.P. Feynman, R.B. Leighton, M. Sands, *Feynman Vorlesungen über Physik* (Oldenbourg Verlag, München, 1988)

[Grau] Dietrich Grau, *Übungsaufgaben zur Quantentheorie,* http://www.dietrich-grau.at/

[Hannabuss] Keith Hannabuss, *An introduction to quantum theory* (Clarendon Press, Oxford, 1997)

[Hittmair] Otto Hittmair, *Lehrbuch der Quantentheorie* (Thiemig, München 1972)

[Itzykson,Zuber] C. Itzykson, J-B. Zuber, *Quantum field theory* (McGraw-Hill Inc., USA, 1980)

[Kreyszig] Erwin Kreyszig, *Introductory Functional Analysis with Applications* (John Wiley & Sons, New York, 1978)

[Landau-Lifschitz] Quantenmechanik, Lehrbuch der Theoretischen Physik Band 3 (Verlag Harri Deutsch, Frankfurt 1988)

[Liboff] Richard L. Liboff, *Introductory Quantum Mechanics* (Addison-Wesley, Reading, Massachusetts, 1998)

[Marchildon] Louis Marchildon, *Quantum mechanics: From Basic Principles to Numerical Methods and Applications* (Springer, Berlin, 2002)

[Messiah] Albert Messiah, *Quantum mechanics* (Dover, Mineola N.Y., 1999)

[Musiol,Ranft,Reif,Seeliger] Gerhard Musiol, Johannes Ranft, Roland Reif, Dieter Seeliger, *Kern- und Elementarteilchenphysik* (VCH, Weinheim, 1988)

[Nachtmann] Otto Nachtmann, *Phänomene und Konzepte der Elementarteilchenphysik* (Vieweg, Braunschweig 1986)

[Reed] Michael Reed, Barry Simon, *Functional Analysis I* (Academic Press, San Diego 1980)

[Schwabl] Franz Schwabl, *Quantenmechanik* (Springer, Berlin, 2002)

[Zettili] Nouredine Zettili, *Quantum mechanics: Concepts and Applications* (John Wiley & Sons, New York, 2001)