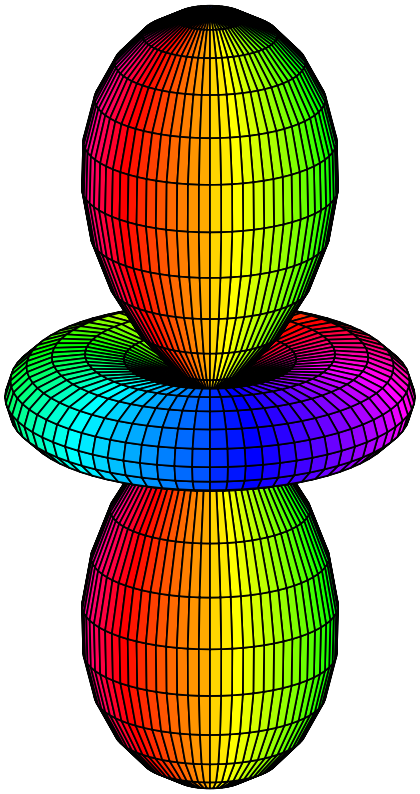
Script

QUANTUM THEORY I

3rd Edition July 2025

by Jakob Weiß, Helmut Hörner, and Tobias Schäfer
based on the lecture by Prof. Andreas Grüneis, Prof. Stefan Rotter
for the Bachelor's program **Technical Physics**
at the **Technische Universität Wien**



0 Preface

Preface to the zeroth edition This script is based on the contents of the course `Quantum Theory I`, which is offered to students of the Bachelor's program in Technical Physics at the Vienna University of Technology. The main topics of this script can be divided into the following areas: (i) failure of classical physics and beginnings of quantum theory, (ii) foundations and formal structure of non-relativistic quantum theory, (iii) applications of quantum theory to fundamental problems and questions. The last two topics are particularly extensive. The script often uses the Dirac notation, which enables an elegant and compact notation of mathematical expressions. We particularly focus on the quantum-theoretical description of the harmonic oscillator and the hydrogen atom, with a closer look at the theory of angular momentum and spin in the latter. The topic of approximation methods is only briefly touched upon, as a more detailed treatment will be given in the continuation course `Quantum Theory II` in the Master's program.

About the design of this script In order not to impede readability, different text boxes are used in this script. We distinguish between motivation boxes, in-depth boxes, and example boxes of the form:

Motivation

Will delve into the motivation behind various concepts of quantum theory.

In-Depth

Will provide context for various concepts of quantum theory or offer mathematical in-depth explanations.

Example

Allows for a more detailed investigation of specific examples.

The content of the motivation and in-depth boxes is intended as learning support and does not have to be considered as additional test material - the example boxes are mainly used to examine different problem situations in more detail. In order to improve readability in complex formulas and derivations, terms that cancel each other out are colored red (e.g. $b + a - a = b$, or $\frac{abc}{b} = ac$). If we use $+0$ or $\cdot 1$ within a formula, we color these terms blue (e.g. $b = a - a + b$ or $ab = ab\frac{c}{c}$).

In addition, important formulas, central statements of a section or fundamental results are explicitly highlighted - the reader is encouraged to pay special attention to formulas of the following form:

$$E = mc^2 \tag{0.1}$$

The script is adapted to the lecture `Quantum Theory I` at the Vienna University of Technology by Prof. A. Grüneis and Prof. S. Rotter. The purpose of this is to use the script in addition to or as a supplement to the blackboard and slides.

Literature In addition, the following books are recommended for further study:

- *Quantum Mechanics, Vol. 1* by C. Cohen-Tannoudji
- *Modern Quantum Mechanics* by J.J. Sakurai
- *Principles of Quantum Mechanics* by R. Shankar
- *Theoretische Physik 3* by M. Bartelmann

About the cover page The cover page shows the spherical harmonics function $Y_2^0(\vartheta, \varphi)$. We will get to know the spherical harmonics functions as the eigenfunctions in the spherically symmetric position space of angular momentum. In the special case of an electron in the hydrogen atom (which we will discuss in detail), we use $Y_2^0(\vartheta, \varphi)$ to localize the electron around the proton with a certain probability.

Preface to the first edition In this updated edition, numerous corrections have been made based on the feedback from the students. In addition, chapters 1, 7, and 8 („Failure of Classical Physics“, „Spin“, and „Perturbation Theory“) have been revised and expanded. Furthermore, a new chapter 9 („Concepts of Quantum Theory“) has been added, in which advanced concepts and technical applications of quantum theory are explained. The authors are of course still grateful for any feedback that helps to further improve the script!

Preface to the second edition In the second edition, not only were errors corrected throughout the script (partly based on feedback from the students, for which the authors are very grateful), but also chapters 2 and parts of chapter 3 were extensively revised and expanded, with the aim of presenting the material and mathematical derivations even more clearly.

Preface to the third edition The present English version of the third edition is a translation of the original German version. It was created in August 2025 by Helmut Hörner with AI assistance. In this third edition, a number of errors reported by attentive students via email or GitHub have been corrected. In particular, the flawed example calculation on page 103f regarding the action of the creation and annihilation operator has been fixed. Additionally, in Chapter *4.3.2 Coherent Glauber States*, the notation has been improved to make the derivation of the time evolution of Glauber states easier to follow. In Chapter *8 Perturbation Theory*, the notation was aligned with that used in the lecture, eliminating discrepancies between the derivation and application of perturbation theory. Furthermore, a technical error had caused half a paragraph to be missing from the English version of the second edition in subsection 9.5.4. This has also been corrected.

Collaboration on GitHub An online version of the script is freely available at `https://github.com/Quantentheorie-1/`. Similarly, the forum at `https://github.com/Quantentheorie-1/Skriptum/issues` on GitHub can be used to report errors or provide suggestions for improving the script. The authors are happy to incorporate the comments and suggestions received into the script and will provide corresponding updated versions on GitHub.

Contents

1 Failure of Classical Physics

At the beginning of the 20th century, physics seemed almost complete. The physical laws known at that time could explain nearly all observations, and there were only a few discrepancies between experiment and theory in some isolated areas. This is why Max Planck was even advised against studying theoretical physics by Philipp von Jolly in 1874, with the following justification:

> "[Theoretical physics is] a highly developed, almost fully matured science that has now, after the discovery of energy, been crowned in a sense, and would soon probably have assumed its final stable form. There might perhaps be a speck of dust or a bubble to examine and to categorize in one corner or another, but the system as a whole stood fairly secure, and theoretical physics was approaching that degree of completion which geometry, for example, had possessed for centuries."

It was assumed that these remaining contradictions would soon resolve with the existing theories of what was then called classical physics. Instead, in attempting to clarify these apparent trifles, a revolution in physics was set in motion: quantum theory was born.

1.1 Radiation of Black Bodies

Motivation: The Problem with Blackbody Radiation

Every real body emits electromagnetic (EM) radiation depending on its temperature T. The radiation spectrum depends both on the temperature of the body and on the reflectance of the surface. A so-called *Black Body* is an idealized real body: it is assumed that all incident EM radiation is completely absorbed. Therefore, in a Black Body, the spectrum of the emitted electromagnetic radiation depends only on the temperature T.

Although no real Black Body exists, it can be experimentally mimicked by a cavity with opaque, sufficiently absorbing inner walls. If this cavity has a small hole, radiation entering through the hole undergoes so many reflections inside the cavity at the absorbing inner walls that it can practically no longer escape from the hole. By heating the walls of the cavity to a temperature T, the hole behaves like a Black Body with a surface equivalent to the cavity. Thus, Blackbody radiation is also synonymously called *Cavity Radiation*.

The problem now is: when attempting to calculate the spectrum of cavity radiation using the laws of classical physics, the Rayleigh-Jeans radiation law (1.12) is derived. While it agrees well with measurements at low frequencies, it does not at high frequencies. It suggests that the energy density increases quadratically with frequency. A Black Body would, therefore, have to emit an infinite amount of energy! This problem is known as the "ultraviolet catastrophe". Another radiation law from classical physics is Wien's Radiation Law (1.20), which was justified using methods from statistical physics. While it can adequately represent the spectrum at high frequencies, it completely fails at low frequencies.

Max Planck tried to unify these radiation laws. However, he only succeeded when, in December 1900, in an „act of desperation", he hypothesized that light is emitted only in discrete energy packets, so-called "quanta" $E = h\nu = \hbar\omega$. This was the birth of quantum physics. The natural constant $h = 6.62607015 \cdot 10^{-34}$ Js is today known as the "Planck constant", while we refer to $\hbar = h/2\pi$ as the "Reduced Planck Constant".

The spectral energy density of a Black Body can be defined as follows:

$$\varepsilon(\omega)d\omega = \langle E(\omega)\rangle\, n(\omega)d\omega \tag{1.1}$$

Here, $\langle E(\omega)\rangle$ is the average energy of the vibrational modes of the radiation field in the Black Body, and $n(\omega)$ is the mode density as a function of the angular frequency ω.

1.1.1 Mode Density of the Black Body

In a closed cavity, only certain vibrational modes can form. The mode density $n(\omega)$ describes the absolute number of modes $N(\omega)$ per volume V in an infinitesimal frequency interval $d\omega$:

$$n(\omega) = \frac{\text{Number of modes}}{\text{Volume} \cdot \text{Frequency interval}} = \frac{1}{V}\frac{dN(\omega)}{d\omega} \tag{1.2}$$

In the following, the mode density will be derived: We consider a cubic volume $V = L^3$ with edge length L, where a standing electromagnetic field $\mathbf{E}(\mathbf{r})$ forms. The surface of the cube is assumed to be perfectly conducting; that is, the electromagnetic field has no component parallel to the surface. Solving Maxwell's equations inside the volume yields a possible field of the form:

$$\mathbf{E}(\mathbf{r}) = \begin{pmatrix} E_x(\mathbf{r}) \\ E_y(\mathbf{r}) \\ E_z(\mathbf{r}) \end{pmatrix} = \begin{pmatrix} E_{x0}\cos(k_x x)\sin(k_y y)\sin(k_z z) \\ E_{y0}\sin(k_x x)\cos(k_y y)\sin(k_z z) \\ E_{z0}\sin(k_x x)\sin(k_y y)\cos(k_z z) \end{pmatrix} \tag{1.3}$$

Consider the case where $x = L$. According to our boundary conditions, the components of the field $\mathbf{E}(x = L, y, z)$ parallel to the cube surface must vanish, which means $E_y = E_z = 0$ must hold. Consequently, the terms $E_{y0}\sin(k_x x)$ and $E_{z0}\sin(k_x x)$ must vanish, which only holds if $\sin(k_x x)$ vanishes with $k_x = n_x \pi x/L$ with $n_i \in \mathbb{N}^+$. This leads to a quantization of the corresponding wave number $k_x = n_x \pi/L$. The same applies analogously in the y and z directions, so we can generally write:

$$k_x = n_x\frac{\pi}{L}, \quad k_y = n_y\frac{\pi}{L}, \quad k_z = n_z\frac{\pi}{L} \quad \text{with} \quad n_x, n_y, n_z \in \mathbb{N}^+ \tag{1.4}$$

Thus, only those waves $\mathbf{E}(\mathbf{r})$ whose wave vector $\mathbf{k}$ is defined through the integer coefficients n_x, n_y, and n_z can be realized in the cubic box:

$$\mathbf{k} = \begin{pmatrix} k_x \\ k_y \\ k_z \end{pmatrix} \overset{(1.4)}{=} \frac{\pi}{L}\begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix} \quad \text{with} \quad n_x, n_y, n_z \in \mathbb{N}^+ \tag{1.5}$$

The magnitude of this wave vector can therefore also be expressed through the coefficients n_x, n_y, and n_z:

$$|\mathbf{k}| \overset{(1.5)}{=} \frac{\pi}{L}\sqrt{n_x^2 + n_y^2 + n_z^2}$$

Each possible $\mathbf{k}$-vector („state“) can thus be assigned a point in a three-dimensional space with the axes k_x, k_y, and k_z (the so-called $\mathbf{k}$-space)–two adjacent values of k_x always have the constant distance $\Delta k_x = \pi/L$ due to (1.4). The same holds for Δk_y and Δk_z:

$$\Delta k_x = \Delta k_y = \Delta k_z = \frac{\pi}{L} \tag{1.6}$$

We now want to know how many possible $\mathbf{k}$-vectors there are whose magnitude $|\mathbf{k}|$ is smaller than a given wave number k_0. To do this, we first assign each state $\mathbf{k}(n_x, n_y, n_z)$, which we have so far represented only as a point in k-space, now a cubic volume $V_\mathbf{k}$. Since all points have constant distances, each point has the same associated volume:

$$V_\mathbf{k} = \Delta k_x \Delta k_y \Delta k_z \overset{(1.6)}{=} \left(\frac{\pi}{L}\right)^3 \tag{1.7}$$

To determine how many states $N(k_0)$ there are that satisfy the condition $|\mathbf{k}| < k_0$, we check how often the volume $V_\mathbf{k}$ fits into the volume of a sphere with radius k_0.
Note that:

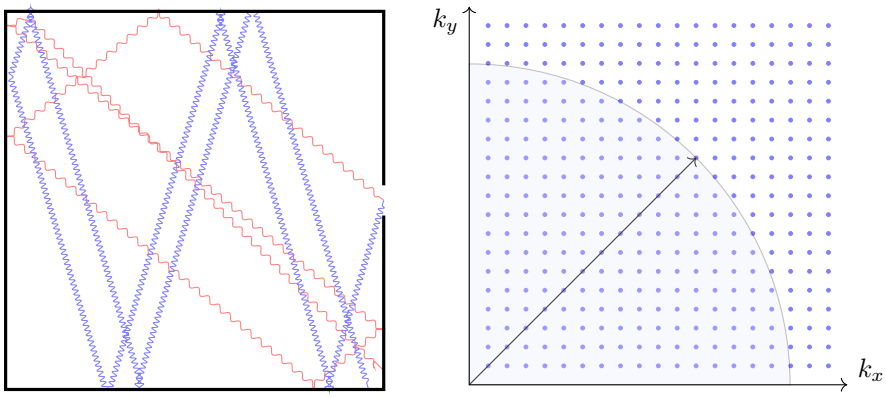


Fig. 1: (left) Radiation forms a standing wave field in the Black Body, in which only the existence of certain wave vectors is possible. (right) Possible states in wave vector space are represented by points. Therefore, only a finite number of wavelengths are allowed in a sphere with a finite radius k.

- k_x, k_y, and k_z can only take positive values. Thus, not the whole sphere volume $V_{\text{Sphere}} = \frac{4}{3}\pi k_0^3$, but the positive octant with $V_{\text{oct}} = \frac{1}{8}\frac{4}{3}\pi k_0^3$ is considered.
- due to the two possible polarization directions of the electromagnetic field $\mathbf{E}(\mathbf{r})$, *two* states can be assigned to all possible k_x, k_y, and k_z. This results in an additional factor of 2.

We thus find for the number of realizable states with the boundary condition $|\mathbf{k}| < k_0$:

$$N(k_0) = 2\frac{V_{\text{oct}}}{V_{\mathbf{k}}} = 2\frac{\frac{1}{8}\frac{4\pi}{3}k_0^3}{(\pi/L)^3} = \frac{\pi}{3}k_0^3\left(\frac{L}{\pi}\right)^3$$

Since we want to express the number of modes as a function of angular frequency $\omega = k_0 c$, not wave number k_0, we substitute $k_0 = \omega/c$ in the above expression, yielding:

$$N(\omega) = \frac{\pi}{3}\frac{\omega^3}{c^3}\frac{L^3}{\pi^3} = \frac{L^3}{3\pi^2}\frac{\omega^3}{c^3} = \frac{V}{3\pi^2}\frac{\omega^3}{c^3} \tag{1.8}$$

From (1.8), the mode density can now be derived using (1.2). Thus we get:

$$n(\omega) \overset{(1.2)}{=} \frac{1}{V}\frac{d}{d\omega}N(\omega) \overset{(1.8)}{=} \frac{\omega^2}{\pi^2 c^3} \tag{1.9}$$

1.1.2 Spectral Energy Density of the Black Body

Deeper Understanding: Classical Energy Density of the Black Body

The average frequency-dependent energy $\langle E\rangle$ according to the laws of classical physics can be calculated using occupancy probability. We use the *Boltzmann distribution* $P(E) = e^{-\beta E}$ with the Boltzmann factor $\beta^{-1} = k_B T$. If the energy can be varied continuously, the averaging is done by integration. We calculate $\langle E\rangle$ as the first moment of E using the following relation:

$$\langle E\rangle = \frac{\int_0^\infty dE\, E\cdot P(E)}{\int_0^\infty dE\, P(E)} = \frac{\int_0^\infty dE\, E\cdot e^{-\beta E}}{\int_0^\infty dE\, e^{-\beta E}} \tag{1.10}$$

The integral in the denominator of the fraction is the normalization and can be solved by

trivial integration:

$$N = \int_0^\infty dE\, e^{-\beta E} = -\frac{1}{\beta}\left[e^{-\beta E}\right]_{E=0}^\infty = -\frac{1}{\beta}(0-1) = \frac{1}{\beta}$$

For evaluating the integral in the numerator, one can either integrate by parts, or use Feynman's trick, where clever differentiation can simplify integrations:

$$\overline{E} = \int_0^\infty dE\, E \cdot e^{-\beta E} = -\int_0^\infty dE\, \frac{\partial}{\partial\beta} e^{-\beta E} = -\frac{\partial}{\partial\beta}\int_0^\infty dE\, e^{-\beta E} = -\frac{\partial}{\partial\beta}\frac{1}{\beta} = \frac{1}{\beta^2}$$

We obtain for the expected value of the energy $\langle E\rangle$ in the classical case:

$$\langle E\rangle = \frac{\overline{E}}{N} = \beta\frac{1}{\beta^2} = \frac{1}{\beta} = k_B T \tag{1.11}$$

Note that according to the equipartition theorem, an average energy of $k_B T/2$ is assigned to each degree of freedom that appears quadratically in the Hamiltonian $H_0 = \frac{\varepsilon_2}{2}\int d^3r(\mathbf{E}^2 + c^2\mathbf{B}^2)$. In using the Boltzmann distribution, it is already taken into account here that each mode with energy E contains both potential energy (quadratic in position) and kinetic energy (quadratic in momentum). The spectral energy density in the classical case is thus given as:

$$\varepsilon(\omega) \overset{(1.1)}{=} \langle E(\omega)\rangle\, n(\omega) \overset{(1.11)}{=} k_B T\, n(\omega) \overset{(1.9)}{=} \frac{k_B T \omega^2}{\pi^2 c^3} \tag{1.12}$$

This corresponds to the Rayleigh-Jeans radiation law. It asserts that the energy density ε increases quadratically with the frequency ω. A Black Body would, therefore, have to emit an infinite amount of energy! This is referred to as the ultraviolet catastrophe.

Max Planck made the momentous assumption that we are dealing with *discrete* energy values in electromagnetic radiation, not continuous ones:

$$E_n = \hbar\omega n \tag{1.13}$$

Therefore, in forming the expectation value $\langle E\rangle$ the integral must be replaced by a sum. $\langle E\rangle$ is thus given by:

$$\langle E\rangle = \frac{\sum_{n=0}^\infty E_n \cdot P(E_n)}{\sum_{n=0}^\infty P(E_n)} = \frac{\hbar\omega\sum_{n=0}^\infty n\cdot e^{-\beta\hbar\omega n}}{\sum_{n=0}^\infty e^{-\beta\hbar\omega n}} \tag{1.14}$$

The sum in the denominator of the fraction is again a normalization. It can be determined with the geometric series formula since the exponential function is less than one in all cases. We substitute $x = e^{-\beta\hbar\omega}$ in the following:

$$N_{\text{dis}} = \sum_{n=0}^\infty e^{-\beta\hbar\omega n} = \sum_{n=0}^\infty \left(e^{-\beta\hbar\omega}\right)^n \overset{\text{subst}}{=} \sum_{n=0}^\infty x^n = \frac{1}{1-x} \tag{1.15}$$

For the numerator in (1.14), a similar trick can be used as in the evaluation of the average energy $\overline{E}$ in the case of continuous energies:

$$\begin{aligned}\overline{E}_{\text{dis}} &= \sum_{n=0}^\infty n e^{-\beta\hbar\omega n} = \sum_{n=0}^\infty n\left(e^{-\beta\hbar\omega}\right)^n \overset{\text{subst}}{=} \sum_{n=0}^\infty n x^n = \sum_{n=0}^\infty x n x^{n-1} = \sum_{n=0}^\infty x\frac{d}{dx}x^n = \\ &= x\frac{d}{dx}\sum_{n=0}^\infty x^n \overset{(1.15)}{=} x\frac{d}{dx}\frac{1}{1-x} = x(1-x)^{-2} = \frac{x}{(1-x)^2}\end{aligned} \tag{1.16}$$

Thus, the averaging for discrete energy values results finally in:

$$\langle E\rangle = \hbar\omega\frac{\overline{E}_{\text{dis}}}{N_{\text{dis}}} = \hbar\omega\frac{x/(1-x)^2}{1/(1-x)} = \hbar\omega\frac{x(1-x)}{(1-x)^2} = \frac{\hbar\omega}{1/x-1} \overset{\text{subst}}{=} \frac{\hbar\omega}{e^{\beta\hbar\omega}-1} \tag{1.17}$$

Therefore, the spectral energy density $\varepsilon(\omega)$ ultimately follows:

$$\varepsilon(\omega) \overset{(1.1)}{=} \langle E(\omega) \rangle \, n(\omega) \overset{(1.17)}{=} \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} n(\omega) \overset{(1.9)}{=} \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} \frac{\omega^2}{\pi^2 c^3}$$

This result is none other than the famous *Planck's Radiation Law*:

$$\varepsilon(\omega) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{e^{\beta\hbar\omega} - 1} \tag{1.18}$$

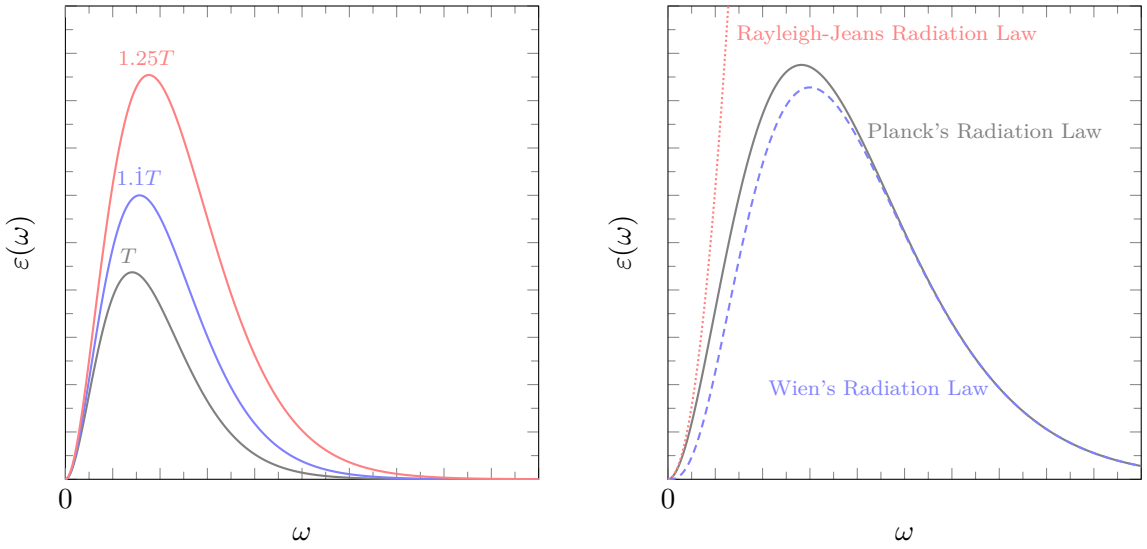


Fig. 2: (left) Planck's radiation law as a function of angular frequency ω at different temperatures. (right) Planck's radiation law with approximations at low and high energies.

Approximations for the limiting cases $\omega \to 0$ and $\omega \to \infty$ can now be investigated. Considering only low frequencies $\omega \to 0$ leads to the classical *Rayleigh-Jeans Radiation Law* (1.12) using the approximation $e^{\beta\hbar\omega} \approx 1 + \beta\hbar\omega$:

$$\lim_{\omega\to 0} \varepsilon(\omega) \overset{(1.18)}{=} \lim_{\omega\to 0} \left[\frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{e^{\beta\hbar\omega} - 1} \right] = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{1 + \beta\hbar\omega + \mathcal{O}(\omega^2) - 1} \approx \frac{\omega^2}{\beta\pi^2 c^3} = \frac{k_B T}{\pi^2 c^3} \omega^2 \tag{1.19}$$

Forming the limit for high frequencies yields $e^{\beta\hbar\omega} - 1 \approx e^{\beta\hbar\omega}$, resulting in the *Wien's Radiation Law*:

$$\lim_{\omega\to\infty} \varepsilon(\omega) \overset{(1.18)}{=} \lim_{\omega\to\infty} \left[\frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{e^{\beta\hbar\omega} - 1} \right] \approx \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{e^{\beta\hbar\omega}} = \frac{\hbar\omega^3}{\pi^2 c^3} e^{-\beta\hbar\omega} \tag{1.20}$$

The derivative with respect to ω from Wien's Radiation Law (1.20) also follows Wien's Displacement Law. It states that the frequency $\omega_{\max}$, at which a Black Body with temperature T emits its most intense radiation, is directly proportional to the temperature T:

$$0 = \frac{d}{d\omega} \left(\omega^3 e^{-\beta\hbar\omega} \right) = 3\omega^2 e^{-\beta\hbar\omega} - \hbar\omega^3 \beta e^{-\beta\hbar\omega} = \omega^2 \left(3 - \hbar\omega\beta \right) e^{-\beta\hbar\omega}$$

The value $\omega = \omega_{\max}$, at which the last part of the expression vanishes, corresponds to the intensity maximum. Both $\hbar$ and k_B are constant quantities, whereby Wien's Displacement Law can be simplified to the following equation:

$$\frac{\omega_{\max}}{T} = \text{const.} \tag{1.21}$$

We have thus managed to show that Wien's Displacement Law, named after Wilhelm Wien, directly follows from Planck's Radiation Law.

1.2 Photoelectric Effect

Motivation: The Photoelectric Effect

When illuminating a metal plate, which is negatively charged with respect to its surroundings, with light of sufficiently high frequency (e.g., UV light), electrons are released from the metal plate. According to the laws of classical physics, the energy of a light wave is given by its intensity. Therefore, it would be expected that the kinetic energy of the electrons released from the plate (and therefore, also their speed in the non-relativistic limit) would increase proportionally with the light intensity.

However, as physicist Philipp Lenard discovered in 1902, this is not the case. The speed of the outgoing electrons does not depend on the intensity, but on the *frequency* of the incoming light! When monochromatic light is shone on the metal plate, all outgoing electrons have the same speed. While the *number* of electrons increases when the intensity is increased, the *speed* of each electron remains unchanged.

In 1905, Albert Einstein succeeded in explaining this effect using the light quantum hypothesis. The incoming light is imagined as being composed of individual „light quanta“(photons). For monochromatic light, each photon has the same energy $E_\gamma = \hbar\omega$ and can only transfer this energy to the knocked out electron. Thus, all electrons have the same speed. An increased intensity of the monochromatic light beam results in more photons hitting the metal plate per unit time, hence more electrons are emitted, but each electron still has the same speed.

In 1921, Albert Einstein was awarded the Nobel Prize for this light quantum hypothesis, although against the opposition of numerous colleagues. Ironically, the antisemitic physicist Philipp Lenard, whose experiments initially prompted the light quantum hypothesis, was among those who protested against awarding the Nobel Prize to Einstein.

The energy density ω_{EM} of electromagnetic radiation is given in the classical electromagnetic theory by the electric field strength $\mathbf{E}$, the electric flux density $\mathbf{D}$, as well as the magnetic field strength $\mathbf{H}$ and the magnetic flux density $\mathbf{B}$. With the electric and the magnetic constants ε_0 and μ_0, we can express the flux densities in terms of their respective field strengths. Thus, we find a relationship for ω_{EM}, which depends on the square of the field strengths, i.e., the *intensities*:

$$\begin{aligned}
\omega_{\mathrm{EM}} &= \frac{1}{2}(\mathbf{E}\cdot\mathbf{D} + \mathbf{H}\cdot\mathbf{B}) = \left| \mathbf{D} = \varepsilon_0\mathbf{E},\ \mathbf{H} = \frac{1}{\mu_0}\mathbf{B} \right. \\
&= \frac{1}{2}\left(\varepsilon_0\mathbf{E}^2 + \frac{\varepsilon_0}{\varepsilon_0\mu_0}\mathbf{B}^2\right) = \\
&= \frac{\varepsilon_0}{2}\left(\mathbf{E}^2 + \frac{1}{\varepsilon_0\mu_0}\mathbf{B}^2\right) = \left| \frac{1}{\varepsilon_0\mu_0} = c^2 \right. \\
&= \frac{\varepsilon_0}{2}(\mathbf{E}^2 + c^2\mathbf{B}^2)
\end{aligned} \tag{1.22}$$

The electromagnetic radiation is attributed a particle character according to Einstein's quantum mechanical interpretation, and the „elementary charges“of the field are referred to as *photons*. According to the heuristic assumption of Planck in deriving his radiation law, these carry the following energy E_γ:

$$E_\gamma = \hbar\omega \tag{1.23}$$

The energy of light radiation (or electromagnetic radiation in general) can take only integer multiples of the energy of a photon. However, the energy of a photon is determined exclusively by the frequency of the photon and is completely independent of the light beam's intensity. The *number* of photons thus determines the amplitude of the light and is therefore responsible for the intensity of the radiation.
But what momentum does a single photon have? According to the special theory of relativity, for a moving particle with a rest mass m_0, it holds that:

$$p^2c^2 = E^2 - m_0^2c^4 \tag{1.24}$$

A photon has a vanishing rest mass $m_0 = m_\gamma = 0$; therefore, the relativistic energy relationship (1.24) simplifies to $E^2 = p_\gamma^2 c^2$. Consequently, the momentum of a photon can be calculated by taking the square root, resulting in:

$$p_\gamma = \frac{E_\gamma}{c} \overset{(1.23)}{=} \frac{\hbar\omega}{c} = \hbar|\mathbf{k}| = \frac{h}{\lambda} \tag{1.25}$$

For the vectorial momentum $\mathbf{p}_\gamma$ of a single photon, it follows:

$$\mathbf{p}_\gamma = \hbar\mathbf{k} \tag{1.26}$$

When a photon collides with a bound electron, it can release the electron from the potential if the energy E_γ is sufficient.

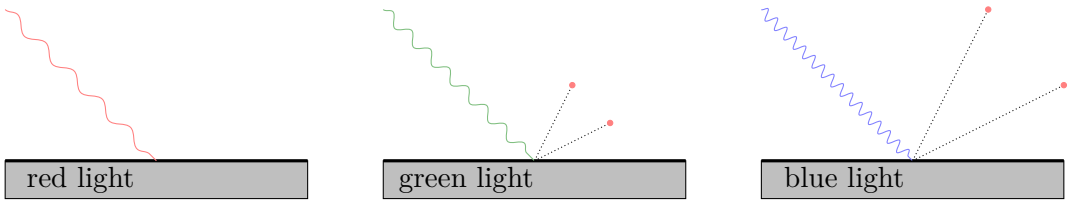


Fig. 3: Photons hit a surface and depending on the frequency, electrons are ejected from it. The E_{kin} of the electrons is directly dependent on the light frequency.

The photon energy $E_\gamma = \hbar\omega$ must exceed the work function W (which is the binding energy of the electron), and the remaining energy is converted into kinetic energy E_{kin} of the now free electron:

$$E_{\text{kin}} = E_\gamma - W = \hbar\omega - W \tag{1.27}$$

If the intensity of the light and thus the number of photons is increased, more electrons are indeed emitted, but the kinetic energy (and thus the speed) of the electrons depends solely on the frequency of the light! If $E_\gamma = \hbar\omega < W$, no electrons are emitted from the metal plate, even if exposure duration and intensity are increased.

1.3 Compton Effect

Motivation: The Compton Effect

The American physicist Arthur Holly Compton discovered the following phenomenon in 1922: When a material is irradiated with X-rays, the scattered radiation experiences a frequency shift to smaller frequencies (i.e., larger wavelengths). This shift is greater the larger the scattering angle and is almost independent of the material.

Compton could explain this effect with the photon model as an elastic collision between a photon with energy $E_\gamma = \hbar\omega$ and momentum $\mathbf{k} = \hbar\mathbf{k}$ and a weakly bound electron of the scattering material. For his discovery, today called the Compton Effect, he was awarded the Nobel Prize in 1927.

The *Compton Effect* is the elastic scattering of photons by (almost) free electrons, which leads to the redshift ($\Delta\lambda = \lambda' - \lambda > 0$) of the scattered light. The photon is characterized by initial and final wavelengths (λ and λ'), as well as by an initial and final momentum ($\mathbf{p}$ and $\mathbf{p}'$). The electron with mass m_e is assumed to be at rest before the interaction ($E_\mathrm{e} = m_\mathrm{e}c^2$ and $p_\mathrm{e} = 0$). The scattering geometry is shown in Figure 4.

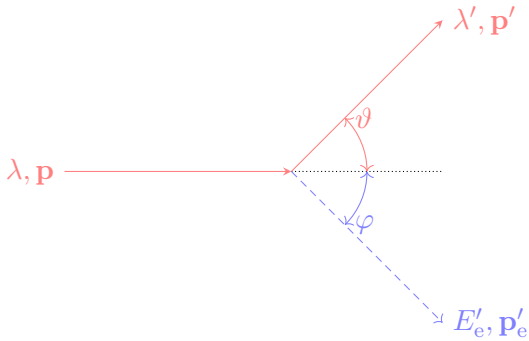


Fig. 4: The Compton Effect as the scattering of a photon by an electron.

The description follows the special theory of relativity. Therefore, the total energy of the electron before the scattering process E_e and after the scattering process E'_e is given by:

$$E_\mathrm{e} = m_\mathrm{e}c^2 \quad \text{and} \quad E'_\mathrm{e} = \sqrt{p'^2_\mathrm{e}c^2 + m^2_\mathrm{e}c^4} \tag{1.28}$$

During the scattering process, both energy and momentum are conserved. Energy conservation can be expressed as:

$$E_\gamma + E_\mathrm{e} = E'_\gamma + E'_\mathrm{e} \tag{1.29}$$

With (1.23), we can express the photon energies in terms of their respective frequencies ω and ω', while for the electrons, the quantities from (1.28) are used. Simple rearrangement of the energy conservation (1.29) leads us to:

$$\hbar\left(\omega - \omega'\right) + m_\mathrm{e}c^2 = \sqrt{p'^2_\mathrm{e}c^2 + m^2_\mathrm{e}c^4} \tag{1.30}$$

To eliminate the square root expression, both sides are squared. Using the relation $\omega = c|\mathbf{k}| = ck$, (1.30) can now be specified as follows:

$$\begin{aligned} [c\hbar\left(k - k'\right) + m_\mathrm{e}c^2]^2 &= p'^2_\mathrm{e}c^2 + m^2_\mathrm{e}c^4 \implies \\ \hbar^2(k - k')^2 + m^2_\mathrm{e}c^2 + 2\hbar m_\mathrm{e}c(k - k') &= p'^2_\mathrm{e} + m^2_\mathrm{e}c^2 \end{aligned} \tag{1.31}$$

Momentum is also conserved, which we can present in a vectorial relation:

$$\mathbf{p}_\gamma + \mathbf{p}_\mathrm{e} = \mathbf{p}'_\gamma + \mathbf{p}'_\mathrm{e} \tag{1.32}$$

The electron is at rest before scattering, so $\mathbf{p}_\mathrm{e} = 0$ holds. Substituting this into the momentum conservation equation (1.32) simplifies to $\mathbf{p}_\gamma = \mathbf{p}'_\gamma + \mathbf{p}'_\mathrm{e}$, where the photon's momentum follows from (1.26). Solving for the momentum of the scattered electron $\mathbf{p}'_\mathrm{e}$, we get the relationship $\mathbf{p}'_\mathrm{e} = \mathbf{p}_\gamma - \mathbf{p}'_\gamma = \hbar\,(\mathbf{k} - \mathbf{k}')$. Thus, for the magnitude of momentum $p'^2_\mathrm{e} = |\mathbf{p}'_\mathrm{e}|^2$:

$$p'^2_\mathrm{e} = \mathbf{p}'_\mathrm{e} \cdot \mathbf{p}'_\mathrm{e} = \hbar^2(\mathbf{k}\cdot\mathbf{k} + \mathbf{k}'\cdot\mathbf{k}' - 2\mathbf{k}\cdot\mathbf{k}') = \hbar^2\left(k^2 + k'^2 - 2kk'\cos(\vartheta)\right) \tag{1.33}$$

The angle ϑ is the scattering angle of the photon according to Figure 4. Expression (1.33) for the scalar p'^2_e can now be substituted into the energy conservation equation (1.31):

$$\begin{aligned}
\hbar^2(k-k')^2 + m_\mathrm{e}^2c^2 + 2\hbar m_\mathrm{e}c(k-k') &= \hbar^2(k^2 + k'^2 - 2kk'\cos(\vartheta)) + m_\mathrm{e}^2c^2 \implies \\
\hbar(k^2 + k'^2 - 2kk') + 2m_\mathrm{e}c(k-k') &= \hbar(k^2 + k'^2 - 2kk'\cos(\vartheta)) \implies \\
-2\hbar kk' + 2m_\mathrm{e}c(k-k') &= -2\,\hbar kk'\cos(\vartheta) \implies \\
m_\mathrm{e}c(k-k') &= \hbar kk' - \hbar kk'\cos(\vartheta)) \implies \\
m_\mathrm{e}c(k-k') &= \hbar kk'(1-\cos(\vartheta))
\end{aligned}$$

A trigonometric relationship for the cosine term is $1-\cos(\vartheta) = 2\sin^2(\vartheta/2)$; additionally applying the relation $k = \omega/c = 2\pi\nu/c = 2\pi/\lambda$, the wavelength change $\Delta\lambda = \lambda' - \lambda$ can explicitly be determined:

$$1 - \cos(\vartheta) = 2\sin^2\left(\frac{\vartheta}{2}\right) = \frac{m_\mathrm{e}c}{\hbar}\frac{k-k'}{kk'} = \frac{m_\mathrm{e}c}{\hbar}\left(\frac{1}{k'} - \frac{1}{k}\right) = \frac{m_\mathrm{e}c}{\hbar 2\pi}(\lambda' - \lambda) = \frac{m_\mathrm{e}c}{h}\Delta\lambda$$

Overall, when photons scatter off electrons, their wavelengths change according to the Compton effect, depending on the scattering angle ϑ:

$$\Delta\lambda = \frac{2h}{m_\mathrm{e}c}\sin^2\left(\frac{\vartheta}{2}\right) = 2\lambda_\mathrm{c}\sin^2\left(\frac{\vartheta}{2}\right)$$

We call $\lambda_\mathrm{c} = 2.4262 \times 10^{-11}\,\mathrm{m}$ the *Compton wavelength* of the electron; in other words, the wavelength of a photon with an energy equivalent to the rest energy of an electron. Overall, we find the wavelength shift of a scattered photon in the context of the Compton Effect as:

$$\Delta\lambda = 2\lambda_\mathrm{c}\sin^2\left(\frac{\vartheta}{2}\right) \tag{1.34}$$

It must hold that $\Delta\lambda > 0$: If $\Delta\lambda < 0$ were the case, the wavelength would become shorter after scattering, corresponding to an increase in energy and thus violating energy conservation. The energy loss of the photon after scattering is transferred to the electron's kinetic energy.

According to (1.34), the frequency shift and thus also the energy transfer to the electron is greatest when the photon is reflected by the electron, resulting in backscattering ($\vartheta = \pi$), and smallest when the photon does not change angles, experiencing forward scattering ($\vartheta = 0$).

1.4 De-Broglie Hypothesis for Matter Waves

Motivation: Particle and Wave Nature

After it was shown that light has both wave and particle characteristics, the French physicist Louis-Victor de Broglie postulated a – somewhat complementary – hypothesis in 1924 in his dissertation: He postulated that every (moving) matter particle also exhibits wave properties, and that the *wave-particle duality* is a universal phenomenon.

This was first confirmed in 1927 in experiments on the diffraction of electrons on thin metal foils by Clinton Davisson and Lester Germer. Distinct interference patterns were observable in these experiments, which can only be explained by the wave nature of electrons. De Broglie received the Nobel Prize in 1929 for his theory of matter waves.

According to the special theory of relativity, each moving particle with rest mass m is assigned energy $E = \sqrt{p^2c^2 + m^2c^4}$. De Broglie postulates that this energy – just like for photons – occurs only in „energy packets“:

$$E = \hbar\omega \tag{1.35}$$

This allows us to write the angular frequency ω as:

$$\hbar\omega = \sqrt{p^2c^2 + m^2c^4} \implies \omega = \frac{1}{\hbar}\sqrt{p^2c^2 + m^2c^4} \tag{1.36}$$

De Broglie assumes that a moving, massive particle with velocity v corresponds to a wave packet with group velocity v_G, which is defined generally as:

$$v \equiv v_\mathrm{G} = \frac{\partial\omega}{\partial k} \tag{1.37}$$

Substituting the expression for ω from (1.36) allows for the following calculation:

$$v = \frac{1}{\hbar}\frac{\partial}{\partial k}\sqrt{p^2c^2 + m^2c^4} = \frac{1}{\hbar}\frac{\partial p}{\partial k}\frac{\partial}{\partial p}\sqrt{p^2c^2 + m^2c^4} = \frac{1}{\hbar}\frac{1}{2}\frac{\partial p}{\partial k}\frac{1}{\sqrt{p^2c^2 + m^2c^4}}2pc^2 \tag{1.38}$$

In the non-relativistic regime, where $m^2c^4 \gg p^2c^2$, it simplifies to $\sqrt{p^2c^2 + m^2c^4} \approx mc^2$. Hence, expression (1.38) simplifies as follows:

$$v \approx \frac{1}{\hbar}\frac{\partial p}{\partial k}\frac{pc^2}{mc^2} = \frac{1}{\hbar}\frac{\partial p}{\partial k}\frac{mv}{m} \implies \hbar v = \frac{\partial p}{\partial k}v \implies \partial p = \hbar\,\partial k \tag{1.39}$$

If both sides of (1.39) are integrated, a relation for the momentum p is obtained (in the non-relativistic approximation):

$$\int dp = \hbar\int dk + p_0 \implies p = \hbar k + p_0$$

p_0 is the integration constant in this context. Since $k = 0$ implies $p = 0$, it can be deduced that $p_0 = 0$ – we finally obtain the following relationship between the momentum p and the wave number k of a moving, massive particle:

$$\mathbf{p} = \hbar\mathbf{k} \tag{1.40}$$

Thus, the de-Broglie wavelength λ can be assigned to a particle with momentum $p = mv$ (with $p \equiv |\mathbf{p}|$ and $v \equiv |\mathbf{v}|$):

$$\lambda = \frac{2\pi}{k} = \frac{2\pi\hbar}{p} = \frac{2\pi\hbar}{mv} \tag{1.41}$$

Utilizing the classic Newtonian relationships $E = mv^2/2$ and $p = mv$, it is also easy to establish a relationship between angular frequency ω and wave number k using the de-Broglie equations:

$$E = \frac{mv^2}{2} = \frac{m^2v^2}{2m} = \frac{\mathbf{p}^2}{2m} \overset{(1.40)}{=} \frac{\hbar^2k^2}{2m} \overset{!}{=} \hbar\omega$$

Dividing both sides of the above equation by $\hbar$ produces the *dispersion relation* $\omega(\mathbf{k})$ for free matter waves:

$$\omega(\mathbf{k}) = \frac{\hbar k^2}{2m} \tag{1.42}$$

From (1.42), the phase velocity v_{Ph} and the group velocity v_{G} (for the non-relativistic limit $v \ll c$) can be derived:

$$v_{\mathrm{Ph}} = \frac{\omega(\mathbf{k})}{k} \overset{(1.42)}{=} \frac{\hbar k}{2m} \quad \text{and} \quad v_{\mathrm{G}} = \frac{\partial}{\partial k}\omega(\mathbf{k}) \overset{(1.42)}{=} \frac{\hbar k}{m} \tag{1.43}$$

Thus, a free matter wave exhibits a quadratic dispersion relation ($E \propto k^2$), as opposed to the linear dispersion relation of photons ($E_\gamma \propto k$).

1.5 Discrete Spectral Lines

Motivation: Spectral Lines of the Hydrogen Atom

When a hydrogen atom is excited, it emits light not in a continuous spectrum, but in a line spectrum, where only particular frequencies appear. Conversely, during absorption, discrete absorption lines are also observed at the same positions in the spectrum. This phenomenon cannot be explained by classical physics. Nonetheless, in 1888, the Swedish physicist Johannes Rydberg could empirically establish the following formula, today known as the Rydberg Formula, which describes the hydrogen line spectrum very accurately:

$$\frac{1}{\lambda} = R_{\mathrm{H}}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right)$$

The spectroscopic quantity $R_{\mathrm{H}} = 1.096\,775\,83 \times 10^7\,\mathrm{m}^{-1}$ is referred to as the *Rydberg Constant* for hydrogen, while the parameters n_1 and n_2 are whole numbers greater than zero, with $n_2 > n_1$. Certain series of spectral lines, each distinguished by its n_1, could be found experimentally: $n_1 = 1$ (Lyman series), $n_1 = 2$ (Balmer series), $n_1 = 3$ (Paschen series), etc.

Only with the help of quantum physics could the *Rydberg Formula* be justified and derived. Today, it is understood that the parameters n_1 and n_2 represent possible values of the principal quantum number n, describing the quantized energy levels of the electron in the hydrogen atom's shell. The spectral lines arise from the discrete "allowed" transitions between different energy levels. In quantum physics, the Rydberg Formula can thus be expressed as:

$$\frac{1}{\lambda} = \frac{\Delta E}{hc} = \frac{\mathrm{Ry}}{hc}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right) \tag{1.44}$$

$\mathrm{Ry} = 13.605\,693\,\mathrm{eV}$ is the Rydberg constant in energy units. The precise derivation of the energy levels of the hydrogen atom (and thus the Rydberg Formula) is provided in a later chapter of this text.

Figure 5 shows both the emission spectrum of helium He and the absorption spectrum of H. The emission spectrum occurs when energy is supplied to a helium gas, causing excitation–successive de-excitation leads to the emission of photons at certain frequencies according to (1.44), which

manifests as discrete spectral lines. This spectral "fingerprint" enables the identification of elements within a gas. If white light is shone through a gas, excitation of the gas atoms occurs at particular wavelengths; photons of that energy are absorbed, hence do not appear in the initial beam after passing through the gas.

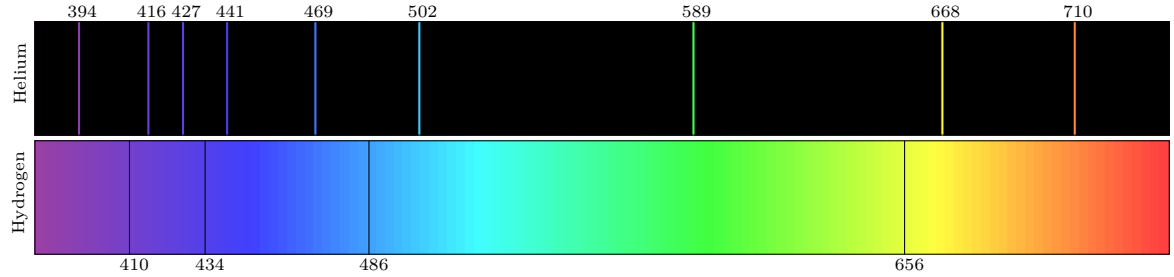


Fig. 5: (top) The emission spectrum of helium; the spectral lines are indicated as wavelengths in nm. (bottom) The absorption spectrum of hydrogen.

2 Schrödinger Equation

Significance of the Schrödinger Equation

If particles and matter also have wave properties according to the hypothesis of DE BROGLIE, the question arises, which wave equation describes these matter waves. The answer was given by ERWIN SCHRÖDINGER when he formulated in 1926 the today named after him *Schrödinger equation* (2.8), which represents the fundamental equation of non-relativistic quantum mechanics. The Schrödinger equation proved to be so successful, among other things in the description of the hydrogen atom, that ERWIN SCHRÖDINGER was awarded the Nobel Prize for his work as early as 1933.

In stark contrast to classical mechanics, where the state of a system at a given time can be described by the positions $\mathbf{r}_i(t)$ and momenta $\mathbf{p}_i(t)$ of all particles, the state of a quantum mechanical system is represented by a so-called *wave function* $\psi(\mathbf{r}, t)$. Here, the Schrödinger equation is the differential equation with which one can determine the temporal evolution of this wave function (and thus of the quantum mechanical system).

2.1 Motivation of the Schrödinger Equation

The Schrödinger equation cannot be directly derived from classical physics. However, starting from the idea of the matter wave and the quantization of energy in the form of "energy packets" $E = \hbar\omega$, it can at least be "motivated" as we will demonstrate in the following.

We consider a moving mass particle (for simplicity initially only in one dimension), based on DE BROGLIE'S idea of the matter wave in free space. In other words, we postulate that the moving mass particle can be described by a wave equation. The general solution of a one-dimensional wave equation is a superposition of plane waves of the form:

$$\psi(x,t) = C\, e^{i(kx-\omega t)} \tag{2.1}$$

If we take the second spatial derivative of that wave function (2.1), we obtain:

$$\frac{\partial^2}{\partial x^2}\psi(x,t) = -k^2\, C\, e^{i(kx-\omega t)} = -k^2\, \psi(x,t) \tag{2.2}$$

According to DE BROGLIE, for the momentum of matter waves $\mathbf{p} = \hbar\mathbf{k}$ applies following (1.40) and thus $p^2 = \hbar^2 k^2$ (with $|\mathbf{p}|^2 \equiv p^2$ and $|\mathbf{k}|^2 \equiv k^2$). From this follows $k^2 = p^2/\hbar^2$, which we can substitute into (2.2):

$$\frac{\partial^2}{\partial x^2}\psi(x,t) = -\frac{p^2}{\hbar^2}\,\psi(x,t)$$

The kinetic energy of a moving particle can be represented (in the non-relativistic case) as $E_{\text{kin}} = mv^2/2 = m^2v^2/(2m) = p^2/(2m)$. This can be rearranged to $p^2 = 2mE_{\text{kin}}$. Substituting into the above equation, we get the following differential equation:

$$\begin{aligned} \frac{\partial^2}{\partial x^2}\psi(x,t) &= -\frac{2m}{\hbar^2}E_{\text{kin}}(x,t)\psi(x,t) \implies \\ E_{\text{kin}}(x,t)\;\psi(x,t) &= -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x,t) \end{aligned} \tag{2.3}$$

For the total energy, we have $E = E_{\text{kin}} + E_{\text{pot}}$. We represent the potential energy E_{pot} initially via a time-independent energy potential $V(x)$, from which follows: $E = E_{\text{kin}}(x,t) + V(x)$, or

$E_{\text{kin}}(x,t) = E - V(x)$. Substituting again into (2.3), we obtain a differential equation for the total energy E:

$$\begin{aligned}
[E - V(x)]\,\psi(x,t) &= -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x,t) \implies \\
E\psi(x,t) &= -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x,t) + V(x)\psi(x,t) = \\
&= \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\psi(x,t)
\end{aligned} \tag{2.4}$$

With this, we have nearly the (one-dimensional) Schrödinger equation! But only nearly, because there is a (second) derivative with respect to position, but not yet a time derivative. Also, we need an expression for the total energy E. Here, de Broglie helps us again. Let us consider the time derivative of the wave function (2.1):

$$\frac{\partial}{\partial t}\psi(x,t) = -i\omega\, C e^{i(kx-\omega t)} = -i\omega\,\psi(x,t) \tag{2.5}$$

According to (1.35), de Broglie postulates for matter waves $E = \hbar\omega$, or $\omega = E/\hbar$. Thus, we can rewrite (2.5) as:

$$\frac{\partial}{\partial t}\psi(x,t) = -\frac{i}{\hbar}E\psi(x,t) \implies E\psi(x,t) = i\hbar\frac{\partial}{\partial t}\psi(x,t) \tag{2.6}$$

Substituting this insight into (2.4), we finally obtain the complete (one-dimensional) Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\psi(x,t) = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\psi(x,t) \tag{2.7}$$

If you go from one dimension to $\mathbb{R}^3$, you need to replace the second spatial derivative with the Laplace operator Δ. In Cartesian coordinates, Δ is simply defined as:

$$\Delta = \boldsymbol{\nabla}^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

Combining all the above points, you get a general expression for the Schrödinger equation of a particle:

$$\left(-\frac{\hbar^2}{2m}\Delta + V(\mathbf{r})\right)\psi(\mathbf{r},t) = i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t) \tag{2.8}$$

The kinetic and potential energy terms on the left in (2.8) lead to an expression corresponding to the classical Hamiltonian $H(\mathbf{r},\mathbf{p})$. This similarity will be discussed in more detail in Appendix 10.1.

For a free particle, that is, for vanishing potential ($V(\mathbf{r}) = 0$), a simplified expression applies in three-dimensional space:

$$-\frac{\hbar^2}{2m}\Delta\psi(\mathbf{r},t) = i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t) \tag{2.9}$$

For a system consisting of N interacting particles in a potential $V_1(\mathbf{r}_i)$, the interaction force $V_2(\mathbf{r}_{\text{rel}})$ between the individual particles with respect to each other ($\mathbf{r}_{\text{rel}} = \mathbf{r}_i - \mathbf{r}_j$) must also be included:

$$\left[\sum_{i=1}^{N}\left(\frac{-\hbar^2}{2m_i}\Delta_i + V_1(\mathbf{r}_i)\right) + \frac{1}{2}\sum_{i\neq j}^{N} V_2(\mathbf{r}_i - \mathbf{r}_j)\right]\psi(\mathbf{r}_1,\ldots,\mathbf{r}_N,t) = i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r}_1,\ldots,\mathbf{r}_N,t) \tag{2.10}$$

If we consider only a system of N particles that are not interacting with each other, the Schrödinger equation (2.10) simplifies again to:

$$\sum_{i=1}^{N}\left(\frac{-\hbar^2}{2m_i}\Delta_i + V_1(\mathbf{r}_i)\right)\psi(\mathbf{r}_1,\ldots,\mathbf{r}_N,t) = i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r}_1,\ldots,\mathbf{r}_N,t) \tag{2.11}$$

2.1.1 Correspondence Principle

Motivation: Representation of Classical Measurable Quantities in Quantum Mechanics

In classical mechanics, measurable system properties A, called *observables*, are described as a function of the phase space, i.e., as a function of all (generalized) coordinates q_i and momenta p_i: $A(t) = A(q_1,\ldots,q_N,p_1,\ldots,p_N,t)$.

However, as we have seen, a quantum mechanical system is not described by the coordinates and momenta of all particles, but by a wave function. Thus, instead of a function with q_i and p_i as function arguments for each classical observable $A(t)$, we need a mathematical formalism that allows us to calculate the desired physical observables for a given wave function $\psi(\mathbf{r},t)$. This is achieved with the help of operators on Hilbert spaces, where certain classical observables are replaced by their corresponding quantum mechanical operators.

In fact, it turns out that for each classical observable $A(t)$, there is an equivalent operator $\hat{A}$ that can act on the wave function $\psi(\mathbf{r},t)$: $A(q_1,\ldots,q_N,p_1,\ldots,p_N,t) \leftrightarrow \hat{A}\psi(\mathbf{r},t)$. This is called the *correspondence principle*. The properties of such operators will be discussed in more detail in later chapters.

Although we shall rigorously define operators later, we want to briefly introduce some important operators at this point, which will simplify the representation of future equations.

For the energy E, the already derived correspondence identity in (2.6) applies:

$$E \leftrightarrow \hat{E} = i\hbar\frac{\partial}{\partial t} \tag{2.12}$$

If we let $\hat{E}$ act on a wave function, we obtain the corresponding energy. Let's next see what the operator $-i\hbar\frac{\partial}{\partial x}$ (which appears in second order in the Schrödinger equation) yields when acting directly on $\psi(x,t)$:

$$-i\hbar\frac{\partial}{\partial x}\psi(x,t) = -i\hbar\frac{\partial}{\partial x}\left(C\,e^{i(kx-\omega t)}\right) = \hbar k\left(C\,e^{i(kx-\omega t)}\right) = \hbar k\,\psi(x,t) \tag{2.13}$$

According to de Broglie, $p = \hbar k$, thus $-i\hbar\frac{\partial}{\partial x}$ is associated with the momentum operator:

$$p \leftrightarrow \hat{p} = -i\hbar\frac{\partial}{\partial x} \tag{2.14}$$

For the momentum operator in three-dimensional space $\mathbb{R}^3$, we analogously find:

$$\mathbf{p} \leftrightarrow \hat{\mathbf{p}} = -i\hbar\boldsymbol{\nabla} \tag{2.15}$$

In view of the three-dimensional Schrödinger equation (2.8), we find for the quadratic momentum operator $\hat{\mathbf{p}}^2$ the following relation:

$$\mathbf{p}^2 \leftrightarrow \hat{\mathbf{p}}^2 = -\hbar^2\Delta \tag{2.16}$$

Let's look once more at the Schrödinger equation (2.8). As noted in (2.12), the right, time-dependent side gives the total energy E. But then, of course, the left side of (2.8) must also provide the total energy! Thus, the following expression is also an energy operator, which we henceforth refer to as the *Hamiltonian operator* $\hat{H}$:

$$\hat{H} = -\frac{\hbar^2}{2m}\Delta + V(\mathbf{r}) \tag{2.17}$$

For a particle with mass m in an arbitrary potential $V(\mathbf{r})$, $\hat{H}$ can be written with the help of (2.15) as:

$$E \leftrightarrow \hat{H} = -\frac{\hbar^2}{2m}\Delta + V(\mathbf{r}) \overset{(2.16)}{=} \frac{\hat{\mathbf{p}}^2}{2m} + V(\hat{\mathbf{r}}) \tag{2.18}$$

In this way, the Schrödinger equation from (2.8) can be compactly written:

$$\hat{H}\psi(\mathbf{r},t) = i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t) \tag{2.19}$$

2.1.2 Probability Density

The probability density $\rho(\mathbf{r},t)$, which is described by the expression $\rho(\mathbf{r},t)d\mathbf{r}$ with which probability the particle is located at time t in an infinitesimally small volume element $d\mathbf{r}$, can be calculated from the wave function $\psi(\mathbf{r},t)$ as follows:

$$\rho(\mathbf{r},t) = \psi^*(\mathbf{r},t)\psi(\mathbf{r},t) = |\psi(\mathbf{r},t)|^2 \tag{2.20}$$

One can see that from experimentally determined values of $\rho(\mathbf{r},t) \equiv \rho$, the underlying wave function $\psi(\mathbf{r},t)$ can never be completely reconstructed, because the magnitude square eliminates all phase information: A wave function $\tilde{\psi} = \psi e^{i\varphi}$ with real phase φ yields exactly the same probability density as ψ. By calculating the complex conjugate of (2.20), one obtains the relation:

$$\rho^* = (\psi(\mathbf{r},t)^*\psi(\mathbf{r},t))^* = \psi(\mathbf{r},t)\psi(\mathbf{r},t)^* = \psi(\mathbf{r},t)^*\psi(\mathbf{r},t) = \rho$$

This must obviously be the case, because the probability density $\rho(\mathbf{r},t)$ is a real number.
In the one-dimensional case, the probability $w(t)$ that a particle is located at time t within a specific range between $x = a$ and $x = b$ is given by:

$$w(t) = \int_a^b dx\, \rho(x,t) = \int_a^b dx\, |\psi(x,t)|^2 \tag{2.21}$$

Since the particle considered (in the one-dimensional case) must with absolute certainty at any time be located somewhere between $x = -\infty$ and $x = +\infty$, the integration of $\rho(x,t)$ from $x = -\infty$ to $x = +\infty$ must yield exactly 1 in the sense of a probability interpretation:

$$\int_{-\infty}^{\infty} dx\, \rho(x,t) = \int_{-\infty}^{\infty} dx\, |\psi(x,t)|^2 = 1 \tag{2.22}$$

In the three-dimensional case, similarly the integration over the entire space V must again yield exactly 1:

$$\int_V dV\, \rho(\mathbf{r},t) = \int_V dV\, |\psi(\mathbf{r},t)|^2 = 1 \tag{2.23}$$

It should be noted that this normalization *does not* automatically result from the Schrödinger equation. Rather, the wave function $\psi(\mathbf{r},t)$ must be provided with an appropriate constant in a separate calculation step to satisfy the normalization conditions (2.22) or (2.23).

For a normalization of the wave function with a suitable constant to be possible, all valid wave functions must belong to the so-called *square-integrable functions*, i.e., satisfy the following condition:

$$\int_V dV\, |\psi(\mathbf{r},t)|^2 < \infty \tag{2.24}$$

The fact that $w(t)$ can be regarded as a probability is motivated by two assumptions: $w(t)$ is real and positive definite ($w(t) \geq 0$). The first condition follows from $\rho^* = \rho$:

$$w^* = \int_a^b dx\, \rho^* = \int_a^b dx\, \rho = w$$

Since $w^* = w$, w must be real. Together with the postulate of the positive definite metric, $w(t)$ is consistent with the interpretation as a probability.

Example: Calculation of a Probability Density

We assume a wave function ψ, which is composed of the superposition of two separate wave functions ψ_1 and ψ_2. It holds:

$$\psi = a\,\psi_1 + b\,\psi_2$$

The probability density can now be calculated via the square of the modulus:

$$\begin{aligned}
|\psi|^2 = \psi^*\psi &= (a^*\psi_1^* + b^*\psi_2^*)(a\,\psi_1 + b\,\psi_2) = \\
&= |a|^2|\psi_1|^2 + |b|^2|\psi_2|^2 + a^*\psi_1^* b\psi_2 + b^*\psi_2^* a\psi_1 = \quad | \; z = a^*\psi_1^* b\psi_2,\ z^* = b^*\psi_2^* a\psi_1 \\
&= |a|^2|\psi_1|^2 + |b|^2|\psi_2|^2 + z + z^* = \quad | \; z + z^* = 2\,\mathrm{Re}(z) \\
&= |a|^2|\psi_1|^2 + |b|^2|\psi_2|^2 + 2\,\mathrm{Re}(a^*\psi_1^* b\psi_2)
\end{aligned}$$

Besides the expected terms describing the probability of the presence of the respective wave functions ψ_1 and ψ_2, a further expression called the *interference term* is obtained. This term is responsible for special phenomena in quantum mechanics, for example, explaining the interference pattern in the double-slit experiment. The form of the interference term is to be understood using two complex numbers $\alpha = a\psi_1$ and $\beta = b\psi_2$:

$$\alpha^*\beta + \alpha\beta^* = (\alpha\beta^*)^* + \alpha\beta^* = \mathrm{Re}(\alpha\beta^*) - i\,\mathrm{Im}(\alpha\beta^*) + \mathrm{Re}(\alpha\beta^*) + i\,\mathrm{Im}(\alpha\beta^*) = 2\,\mathrm{Re}(\alpha\beta^*)$$

Another computation possibility for the complex numbers is:

$$\begin{aligned}
\alpha^*\beta + \alpha\beta^* &= [\mathrm{Re}(\alpha) - i\,\mathrm{Im}(\alpha)][\mathrm{Re}(\beta) + i\,\mathrm{Im}(\beta)] + [\mathrm{Re}(\alpha) + i\,\mathrm{Im}(\alpha)][\mathrm{Re}(\beta) - i\,\mathrm{Im}(\beta)] = \\
&= 2\,\mathrm{Re}(\alpha)\,\mathrm{Re}(\beta) + 2\,\mathrm{Im}(\alpha)\,\mathrm{Im}(\beta) = 2\,[\mathrm{Re}(\alpha\beta) + \mathrm{Im}(\alpha\beta)]
\end{aligned}$$

2.1.3 Probability Current Density

Motivation: Probability as a Conserved Quantity

As we have already determined, the probability of finding an observed particle somewhere in the entire space must always be one. Strictly speaking, this is a constraint of Schrödingerian quantum mechanics. In high-energy physics, particles can indeed be created or annihilated – an aspect considered in quantum field theory (QFT). The probability of presence can be higher at time t_1 in one space region than in another, and at a later time t_2, it may be the reverse. But overall, the total probability across the entire space always remains one, thus the total probability is a conserved quantity!

The entire space thus resembles a container in some way, in which the probability density can "flow" from one place to another, but additional probability is never "created or destroyed". Therefore, there must be a continuity equation, and a probability current density must be definable.

For the position probability density, we have previously determined the following expression: $\rho(\mathbf{r},t) = |\psi(\mathbf{r},t)|^2$, where ρ is a positive definite quantity. Integrating over the entire space, ρ always yields 1, meaning that the total probability remains conserved. Simultaneously, $\rho(\mathbf{r},t)$ can change over time in each space region. Therefore, we expect there must be a *probability current density* $\mathbf{j}$ that, together with ρ, satisfies a *continuity equation* (similar to, e.g., the charge carrier density and charge current density in electrodynamics fulfilling a continuity equation):

$$\frac{\partial \rho}{\partial t} = -\boldsymbol{\nabla} \cdot \mathbf{j} \tag{2.25}$$

It becomes evident: The faster the probability density ρ changes over time, the greater must be the spatial change of the probability current density $\mathbf{j}$. Now we attempt to find an expression for the probability current density $\mathbf{j}$ such that the above equation (2.25) is satisfied. To do this, we need to form the time derivative of $\rho = \psi^*\psi$:

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial t}(\psi^*\psi) = \frac{\partial \psi^*}{\partial t}\psi + \psi^*\frac{\partial \psi}{\partial t} \tag{2.26}$$

To convert the time derivative of ψ into a space derivative, we use the Schrödinger equation (2.9) for a free particle:

$$-\frac{\hbar^2}{2m}\Delta\psi = i\hbar\frac{\partial \psi}{\partial t} \implies \frac{\partial \psi}{\partial t} = -\frac{\hbar}{2im}\Delta\psi \tag{2.27}$$

We obtain an expression for the time derivative of ψ^* by taking the complex conjugate of both sides of the Schrödinger equation (2.9):

$$\left(-\frac{\hbar^2}{2m}\Delta\psi\right)^* = \left(i\hbar\frac{\partial \psi}{\partial t}\right)^* \implies -\frac{\hbar^2}{2m}\Delta\psi^* = -i\hbar\frac{\partial \psi^*}{\partial t} \implies \frac{\partial \psi^*}{\partial t} = \frac{\hbar}{2im}\Delta\psi^* \tag{2.28}$$

Substituting (2.28) and (2.27) into (2.26), we get:

$$\frac{\partial \rho}{\partial t} = \frac{\hbar}{2im}\left[(\Delta\psi^*)\psi - \psi^*\Delta\psi\right] \tag{2.29}$$

We now have an equation that on the left side includes the time derivative of ρ and on the right side includes space derivatives of ψ and ψ^*. However, we cannot yet bring the right side of (2.29) into direct agreement with the continuity equation (2.25) because the latter is determined by the divergence of a (probability) current. Therefore, a further transformation is required, for which we must first derive the following identities using the product rule:

$$\begin{aligned} \boldsymbol{\nabla}\cdot[(\boldsymbol{\nabla}\psi^*)\psi] &= (\Delta\psi^*)\psi + (\boldsymbol{\nabla}\psi^*)(\boldsymbol{\nabla}\psi) \implies (\Delta\psi^*)\psi = \boldsymbol{\nabla}\cdot[(\boldsymbol{\nabla}\psi^*)\psi] - (\boldsymbol{\nabla}\psi^*)(\boldsymbol{\nabla}\psi) \\ \boldsymbol{\nabla}\cdot(\psi^*\boldsymbol{\nabla}\psi) &= (\boldsymbol{\nabla}\psi^*)(\boldsymbol{\nabla}\psi) + \psi^*\Delta\psi \implies \psi^*\Delta\psi = \boldsymbol{\nabla}\cdot(\psi^*\boldsymbol{\nabla}\psi) - (\boldsymbol{\nabla}\psi^*)(\boldsymbol{\nabla}\psi) \end{aligned} \tag{2.30}$$

By substituting both relationships from (2.30) into (2.29), we finally get:

$$\begin{aligned}
\frac{\partial \rho}{\partial t} &= \frac{\hbar}{2im}\left[\boldsymbol{\nabla}\cdot[(\boldsymbol{\nabla}\psi^*)\psi] - (\boldsymbol{\nabla}\psi^*)(\boldsymbol{\nabla}\psi) - \boldsymbol{\nabla}\cdot(\psi^*\boldsymbol{\nabla}\psi) + (\boldsymbol{\nabla}\psi^*)(\boldsymbol{\nabla}\psi)\right] = \\
&= \frac{\hbar}{2im}\boldsymbol{\nabla}\cdot[(\boldsymbol{\nabla}\psi^*)\psi - \psi^*\boldsymbol{\nabla}\psi] = \\
&= -\boldsymbol{\nabla}\cdot\Big[\psi^*\frac{\hbar}{2im}\boldsymbol{\nabla}\psi + \psi\Big(-\frac{\hbar}{2im}\Big)\boldsymbol{\nabla}\psi^*\Big] = \quad \big|\ \boldsymbol{\alpha} = \psi^*\frac{\hbar}{2im}\boldsymbol{\nabla}\psi \\
&= -\boldsymbol{\nabla}\cdot(\boldsymbol{\alpha} + \boldsymbol{\alpha}^*) = -\boldsymbol{\nabla}\cdot[2\,\mathrm{Re}\,(\boldsymbol{\alpha})] = \\
&= -\boldsymbol{\nabla}\cdot\underbrace{\mathrm{Re}\left(\psi^*\frac{\hbar}{im}\boldsymbol{\nabla}\psi\right)}_{\mathbf{j}} \overset{(2.25)}{=} -\boldsymbol{\nabla}\cdot\mathbf{j}
\end{aligned} \tag{2.31}$$

Comparing this result with the continuity equation (2.25), we can immediately read off the sought expression for the probability current density $\mathbf{j}(\mathbf{r},t)$:

$$\mathbf{j}(\mathbf{r},t) = \mathrm{Re}\left(\psi^*\frac{\hbar}{im}\boldsymbol{\nabla}\psi\right) = \frac{\hbar}{2im}\left[\psi^*\boldsymbol{\nabla}\psi - (\boldsymbol{\nabla}\psi^*)\psi\right] \tag{2.32}$$

In this derivation, we used the Schrödinger equation for a free particle. Expression (2.32) and the continuity equation (2.25) are valid (without proof) for a particle in a (real-valued) potential as well. Moreover, it follows from (2.25) that the probability density remains conserved. To demonstrate this, we apply Gauss' theorem (with surface normal $\mathbf{n}$):

$$-\int_V dV\,\boldsymbol{\nabla}\cdot\mathbf{j} = -\oint_{\partial V} dS\,\mathbf{n}\cdot\mathbf{j} = 0$$

If $\psi(\mathbf{r},t)$ is a square-integrable function, it can be normalized, so that the probability of presence at the edge of the system vanishes. From this follows the conservation of norm with real potentials.

2.1.4 Time-independent Schrödinger Equation

In this chapter, we show that the Schrödinger equation can be solved by separating the wave function into a space-dependent and a time-dependent part, and how the solution that depends only on space relates to the so-called *time-independent Schrödinger equation*.

The Schrödinger equation (2.8) and the resulting wave function $\psi(\mathbf{r},t)$ are, evidently, time and space-dependent:

$$\left(-\frac{\hbar^2}{2m}\Delta + V(\mathbf{r})\right)\psi(\mathbf{r},t) = i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t)$$

As long as the potential $V(\mathbf{r})$ depends only on the location, we can choose an ansatz where the wave function $\psi(\mathbf{r},t)$ is factorizable into a space-dependent part and a time-dependent part:

$$\psi(\mathbf{r},t) = \tilde{\psi}(\mathbf{r})\chi(t) \tag{2.33}$$

Substituting the ansatz (2.33) for these special solutions into the Schrödinger equation (2.8), we succeed in separating space and time variables:

$$\begin{aligned}
-\frac{\hbar^2}{2m}\Delta\left(\tilde{\psi}(\mathbf{r})\chi(t)\right) + V(\mathbf{r})\tilde{\psi}(\mathbf{r})\chi(t) &= i\hbar\frac{\partial}{\partial t}\left(\tilde{\psi}(\mathbf{r})\chi(t)\right) \\
-\frac{\hbar^2}{2m}\chi(t)\Delta\tilde{\psi}(\mathbf{r}) + V(\mathbf{r})\tilde{\psi}(\mathbf{r})\chi(t) &= i\hbar\tilde{\psi}(\mathbf{r})\frac{d\chi(t)}{dt} \quad \big|\ \div\left[\tilde{\psi}(\mathbf{r})\chi(t)\right] \\
\frac{\hbar^2}{2m}\frac{1}{\tilde{\psi}(\mathbf{r})}\Delta\tilde{\psi}(\mathbf{r}) + V(\mathbf{r}) &= i\hbar\frac{1}{\chi(t)}\frac{d\chi(t)}{dt}
\end{aligned} \tag{2.34}$$

The left side of (2.34) depends only on $\mathbf{r}$, the right side depends only on t. To satisfy the equation for all $\mathbf{r}$ and t, both sides must correspond to a constant. We wisely call this constant E, and first look at the right side of equation (2.34), which we also multiply by $\chi(t)$:

$$i\hbar\frac{d}{dt}\chi(t) = E\chi(t) \tag{2.35}$$

A comparison with (2.12) shows: Since $\hat{E} = i\hbar\frac{d}{dt}$ is the energy operator, the chosen constant E is obviously the total energy. A solution for $\chi(t)$ can be easily found:

$$i\hbar\frac{d\chi}{dt} = E\chi \mid \cdot\frac{1}{i\hbar\chi} \implies \frac{1}{\chi}\frac{d\chi}{dt} = \frac{E}{i\hbar} \mid \cdot dt \implies \frac{1}{\chi}d\chi = \frac{E}{i\hbar}dt \implies \int\frac{1}{\chi}d\chi = \frac{E}{i\hbar}\int dt \implies$$
$$\ln(\chi) = \frac{E}{i\hbar}t + c = -i\frac{Et}{\hbar} + c \implies \chi(t) = e^{-iEt/\hbar+c} = e^{-iEt/\hbar}e^{c} = Ne^{-iEt/\hbar}$$

We simply choose the constant $N = 1$, because we can, due to the ansatz $\psi(\mathbf{r},t) = \tilde{\psi}(\mathbf{r})\chi(t)$, shift all normalization into $\tilde{\psi}(\mathbf{r})$. It holds $E = \hbar\omega$ or $\omega = E/\hbar$, enabling us to write the solution for $\chi(t)$ as:

$$\chi_n(t) = e^{-iE_nt/\hbar} = e^{-i\omega_nt} \tag{2.36}$$

Due to the mathematical form of $\chi_n(t)$, the time-dependent part of the solution corresponds to a *phase factor*. The index n suggests, as we will see, that there may be several (often infinitely many) solutions of the Schrödinger equation with different energies E_n. Let's now look at the left – space-dependent – side of equation (2.34) and equate it again to the constant energy E. After multiplying with $\tilde{\psi}(\mathbf{r})$:

$$-\frac{\hbar^2}{2m}\Delta\tilde{\psi}(\mathbf{r}) + V(\mathbf{r})\tilde{\psi}(\mathbf{r}) = E\tilde{\psi}(\mathbf{r}) \tag{2.37}$$

Equation (2.37) is the time-independent Schrödinger equation. We can write it compactly using the Hamiltonian operator (2.18):

$$\hat{H}\tilde{\psi}_n(\mathbf{r}) = E_n\tilde{\psi}_n(\mathbf{r}) \tag{2.38}$$

Again, we use the index n, indicating that there may be several solutions with different energies. The time-independent Schrödinger equation is generally a linear, elliptic, partial differential equation, finding its solution will be the task of the following chapters.

Stationary Solutions We have shown that, given a time-independent potential $V(\mathbf{r})$, special solutions of the Schrödinger equation can be factorized into a space-dependent part $\tilde{\psi}_n(\mathbf{r})$ and a time-dependent part $\chi_n(t)$ so that:

$$\psi_n(\mathbf{r},t) = \tilde{\psi}_n(\mathbf{r})\chi_n(t) \overset{(2.36)}{=} \tilde{\psi}_n(\mathbf{r})e^{-i\omega_nt} \tag{2.39}$$

The index n again characterizes different solutions with different energies E_n. Observing these special solutions closely, the associated probability of presence is time-independent (stationary):

$$\rho = \psi^*(\mathbf{r},t)\psi(\mathbf{r},t) = \tilde{\psi}^*(\mathbf{r})e^{i\omega t}\,\tilde{\psi}(\mathbf{r})e^{-i\omega t} = |\tilde{\psi}(\mathbf{r})|^2 \tag{2.40}$$

One speaks of *stationary solutions* of the Schrödinger equation, which are determined by solving the time-independent Schrödinger equation (2.38) and are each associated with a well-defined energy E_n.

Superposition of States When is it not sufficient to solve only the time-independent Schrödinger equation (2.38)? Since the Schrödinger equation is a linear differential equation (meaning, ψ appears only as a first power), the superposition principle applies: arbitrary linear combinations of solutions are again solutions. Thus, we may be interested in solutions representing a superposition of several stationary solutions, such as:

$$\Psi(\mathbf{r},t) = N\left(\psi_1(\mathbf{r},t) + \psi_2(\mathbf{r},t)\right) = N\left(\tilde{\psi}_1(\mathbf{r})e^{-i\omega_1 t} + \tilde{\psi}_2(\mathbf{r})e^{-i\omega_2 t}\right) \tag{2.41}$$

The constant N is a normalization constant to be chosen such that the normalization condition (2.23) is satisfied. Calculating the probability density ρ for $\Psi(\mathbf{r},t)$, it is seen that it is *not* time-independent:

$$\begin{aligned}
\rho &= \Psi^*(\mathbf{r},t)\Psi(\mathbf{r},t) = \\
&= |N|^2\left[\tilde{\psi}_1^*(\mathbf{r})e^{i\omega_1 t} + \tilde{\psi}_2^*(\mathbf{r})e^{i\omega_2 t}\right]\left[\tilde{\psi}_1(\mathbf{r})e^{-i\omega_1 t} + \tilde{\psi}_2(\mathbf{r})e^{-i\omega_2 t}\right] = \\
&= |N|^2\Big[|\tilde{\psi}_1(\mathbf{r})|^2 + |\tilde{\psi}_2(\mathbf{r})|^2 + \tilde{\psi}_1^*(\mathbf{r})\tilde{\psi}_2(\mathbf{r})e^{i(\omega_1-\omega_2)t} + \tilde{\psi}_1(\mathbf{r})\tilde{\psi}_2^*(\mathbf{r})e^{i(\omega_2-\omega_1)t}\Big] = \quad | \ \alpha = \tilde{\psi}_1^*\tilde{\psi}_2 e^{i\Delta\omega t} \\
&= |N|^2\left[|\tilde{\psi}_1(\mathbf{r})|^2 + |\tilde{\psi}_2(\mathbf{r})|^2 + \alpha + \alpha^*\right] = \quad | \ \alpha + \alpha^* = 2\,\mathrm{Re}(\alpha) \\
&= |N|^2\left[|\tilde{\psi}_1(\mathbf{r})|^2 + |\tilde{\psi}_2(\mathbf{r})|^2 + 2\,\mathrm{Re}\left(\tilde{\psi}_1^*(\mathbf{r})\tilde{\psi}_2(\mathbf{r})e^{i(\omega_1-\omega_2)t}\right)\right]
\end{aligned} \tag{2.42}$$

The wave function $\Psi(\mathbf{r},t)$, which cannot be assigned to a definite energy E_n, is not a solution to the time-independent Schrödinger equation, and its time development results from the superposition of several stationary solutions (with time-dependent phase factors). Alternatively, one can attempt a direct solution of the time-dependent Schrödinger equation.

2.1.5 Properties of the Schrödinger Equation

An exact solution of the Schrödinger equation is only possible for very few systems. Through a thorough mathematical classification of the Schrödinger equation, helpful insights into possible solutions can nevertheless be derived. Properties of the Schrödinger equation are:

- **Sturm-Liouville Problem**: The one-dimensional, time-independent Schrödinger equation can be understood as a Sturm-Liouville eigenvalue problem. This means that we can find real, ordered eigenvalues λ_n ($\lambda_1 < \lambda_2 < \dots$), whose corresponding eigenfunctions will have $n-1$ zeros. These "nodes" lead us to the *node theorem*, allowing an ordering of the eigenvalues. The eigenfunctions are orthogonal to each other and form a complete basis.
- **Partial Differential Equation**: The time-independent Schrödinger equation is a partial differential equation because it includes both partial derivatives with respect to position $\mathbf{r}$ and time t. However, the time derivative is a simple derivative, while there is a second derivative concerning the position: thus, it is a first-order differential equation in time and a second-order one in space. The lack of symmetry in space and time derivatives is problematic for a relativistic application of the Schrödinger equation because, in relativity, space and time coordinates are treated equitably. This issue is solved by the Dirac equation, which, however, is not covered in this lecture.
- **Parabolic Differential Equation**: The time-dependent Schrödinger equation belongs to the class of parabolic differential equations because the time component is influenced by a first time derivative. Due to the dependence on position and time, solving the Schrödinger equation is both a boundary value and an initial value problem. Another parabolic differential equation is the heat equation.
- **Linear Differential Equation**: The Schrödinger equation is a linear differential equation, meaning ψ appears only as a first power. This, in turn, means that even a superposition of solutions must again be a solution of the Schrödinger equation. To show

this, assume that ψ and ϕ represent valid wave functions. Then the same must hold for $\Psi = \alpha\psi + \beta\phi$:

$$\left(\hat{H} - i\hbar\frac{\partial}{\partial t}\right)\Psi = \left(\hat{H} - i\hbar\frac{\partial}{\partial t}\right)(\alpha\psi + \beta\phi) = \alpha\left(\hat{H} - i\hbar\frac{\partial}{\partial t}\right)\psi + \beta\left(\hat{H} - i\hbar\frac{\partial}{\partial t}\right)\phi = 0 \quad \square$$

- **Homogeneous Differential Equation**: The homogeneity of the Schrödinger equation is a prerequisite for the wave function to be normalizable at all times. If this assumption is abandoned, the conservation of the norm of the Schrödinger equation is violated (likewise, this is the case when the potential contains imaginary terms).
- **Born's Probability Interpretation**: The wave functions $\psi(\mathbf{r}, t)$ as solutions of the Schrödinger equation escape a direct physical interpretation. Yet $\rho(\mathbf{r}, t) = |\psi(\mathbf{r}, t)|^2$ from (2.20) can be understood as a probability density.

In-Depth: No Norm Conservation with Complex Potential

Typically (and throughout this script), the potential term in the Schrödinger equation is real-valued. However, in certain contexts, it can be practical to assume a complex-valued potential $V(\mathbf{r})$:

$$\left(-\frac{\hbar^2}{2m}\Delta + V(\mathbf{r})\right)\psi(\mathbf{r}, t) = i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r}, t) \quad \text{with} \quad \hat{H} = -\frac{\hbar^2}{2m}\Delta + V(\mathbf{r}) \tag{2.43}$$

Since $V(\mathbf{r})$ now has imaginary parts, $V(\mathbf{r}) \neq V^*(\mathbf{r})$ and the Hamiltonian is thus no longer Hermitian overall ($\hat{H} \neq \hat{H}^*$). We now want to clarify the effect on the continuity equation. We begin with the same approach (2.26) as with the free particle:

$$\frac{\partial\rho}{\partial t} = \frac{\partial}{\partial t}(\psi^*\psi) = \frac{\partial\psi^*}{\partial t}\psi + \psi^*\frac{\partial\psi}{\partial t}$$

To establish a relation between $\frac{\partial\psi}{\partial t}$ and $\Delta\psi$, we divide both sides of the (complex-conjugated) Schrödinger equation (2.43) by $i\hbar$:

$$\frac{\partial\psi}{\partial t} = -\frac{\hbar}{2im}\Delta\psi + \frac{V}{i\hbar}\psi \quad \text{and} \quad \frac{\partial\psi^*}{\partial t} = \frac{\hbar}{2im}\Delta\psi^* - \frac{V^*}{i\hbar}\psi^* \tag{2.44}$$

We now substitute both relations from (2.44) into the approach (2.26):

$$\begin{aligned}\frac{\partial\rho}{\partial t} &= \left[\frac{\hbar}{2im}\Delta\psi^* - \frac{V^*}{i\hbar}\psi^*\right]\psi + \psi^*\left[-\frac{\hbar}{2im}\Delta\psi + \frac{V}{i\hbar}\psi\right] = \\ &= \underbrace{\frac{\hbar}{2im}\left[(\Delta\psi^*)\psi - \psi^*\Delta\psi\right]}_{-\nabla\cdot\mathbf{j}} + \frac{1}{i\hbar}\psi^*\psi\,(V - V^*)\end{aligned} \tag{2.45}$$

The first term on the right side of equation (2.45) is identical to expression (2.29), which in further progression was identified in (2.31) as $-\boldsymbol{\nabla}\cdot\mathbf{j}$. For the expression $(V - V^*)$, the following calculation rule for complex numbers applies: $V - V^* = 2i\,\mathrm{Im}(V)$. Thus, we obtain the following modified continuity equation:

$$\frac{\partial\rho}{\partial t} = -\boldsymbol{\nabla}\cdot\mathbf{j} + \frac{2}{\hbar}\mathrm{Im}(V) \tag{2.46}$$

The continuity equation now includes an additional term due to the complex-valued potential, which, depending on the sign of $\mathrm{Im}(V)$, corresponds to a particle source or sink.

Consequently, the total probability ρ (i.e., the integrated probability density over the entire space) is generally *not* conserved. This allows relatively simple simulation of physical situations where additional particles can emerge or disappear.

In-Depth: No Norm Conservation with Inhomogeneous Schrödinger Equation

The physically correct Schrödinger equation is a homogeneous differential equation. To simplify the simulation of specific physical configurations, it might be helpful to add an inhomogeneity term. Therefore, we assume a Schrödinger equation for a free particle with a (complex) inhomogeneity $s(\mathbf{r}, t)$:

$$-\frac{\hbar^2}{2m}\Delta\psi(\mathbf{r},t) + s(\mathbf{r},t) = i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t) \quad \text{with} \quad \hat{H} = -\frac{\hbar^2}{2m}\Delta \tag{2.47}$$

At least here $\hat{H} = \hat{H}^*$ (the Hamiltonian is Hermitian). This is, however, not the case for a complex-valued inhomogeneity: $s(\mathbf{r}, t) \neq s^*(\mathbf{r}, t)$. Again, we aim to derive a continuity equation. Proceeding similarly as in (2.26):

$$\frac{\partial\rho}{\partial t} = \frac{\partial}{\partial t}(\psi^*\psi) = \frac{\partial\psi^*}{\partial t}\psi + \psi^*\frac{\partial\psi}{\partial t}$$

As in the preceding example, we divide the Schrödinger equation (2.47) by $i\hbar$ to create a simple connection between the temporal and spatial derivatives. Following, we take the complex conjugate of both sides of the equation:

$$\frac{\partial\psi}{\partial t} = -\frac{\hbar}{2im}\Delta\psi + \frac{1}{i\hbar}s \quad \text{and} \quad \frac{\partial\psi^*}{\partial t} = \frac{\hbar}{2im}\Delta\psi^* - \frac{1}{i\hbar}s^* \tag{2.48}$$

With this, we can proceed with the tried-and-true approach by substituting (2.48) into (2.26):

$$\begin{aligned}\frac{\partial\rho}{\partial t} &= \left[\frac{\hbar}{2im}\Delta\psi^* - \frac{1}{i\hbar}s^*\right]\psi + \psi^*\left[-\frac{\hbar}{2im}\Delta\psi + \frac{1}{i\hbar}s\right] = \\ &= \frac{\hbar}{2im}\left[(\Delta\psi^*)\psi - \psi^*\Delta\psi\right] - \frac{1}{i\hbar}[s^*\psi - \psi^*s] = \quad | \ \alpha = s^*\psi \\ &= \underbrace{\frac{\hbar}{2im}\left[(\Delta\psi^*)\psi - \psi^*\Delta\psi\right]}_{-\nabla\cdot\mathbf{j}} - \frac{1}{i\hbar}(\alpha - \alpha^*)\end{aligned} \tag{2.49}$$

The first term on the right of equation (2.49), as we have learned from (2.29) and (2.31), is identical to $-\boldsymbol{\nabla}\cdot\mathbf{j}$. For the expression $(\alpha - \alpha^*)$, the following calculation rule for complex numbers applies: $\alpha - \alpha^* = 2i\,\mathrm{Im}(\alpha) \equiv 2i\,\mathrm{Im}(s^*\psi)$. Consequently, we obtain the following modified continuity equation:

$$\frac{\partial\rho}{\partial t} = -\boldsymbol{\nabla}\cdot\mathbf{j} - \frac{2}{\hbar}\mathrm{Im}(s^*\psi) \tag{2.50}$$

Once again, the continuity equation includes an additional term which, depending on the sign of $\mathrm{Im}(s^*\psi)$, acts absorbingly or enhancingly, breaking the conservation property of the continuity equation. In this manner, one can fairly easily calculate damping effects.

2.2 Wave Packets

Motivation: Describing Classical Processes with Quantum Physics

We have already shown that the one-dimensional Schrödinger equation (2.7) is solved by plane waves of the form $\psi(x,t) = Ae^{i(kx-\omega t)}$. However, when we want to represent something as elementary as the motion of a single particle in space with it, we encounter multiple problems:

- Classical particles are precisely localizable. We expect the quantum theory to represent the position of a particle, at least with the corresponding quantum mechanical probability.
- In contrast to this expectation, plane waves are entirely delocalized: The probability density of a plane wave $\rho = |\psi(x,t)|^2 = |Ae^{i(kx-\omega t)}|^2$ is constant at every point in space $\rho = |A|^2$.
- Thus, $\psi(x,t)$ cannot be normalized according to the normalization condition (2.22): The integral of a constant over an infinite space region goes against infinity!

We will show in this chapter that the solution to these problems lies in combining an infinite number of plane waves in superposition, so that the overall solution ultimately yields a spatially localized wave packet, identifying it with a particle that (even if with corresponding "quantum fuzziness") follows the laws of classical physics in specific limit cases.

We seek a suitable superposition of solutions of the Schrödinger equation ψ_i. Thus, one might try the following ansatz (A here being the overall normalization constant):

$$\phi(x,t) = A\sum_i a_i\psi_i(x,t) = A\sum_i a_i e^{i(k_i x-\omega_i t)} \tag{2.51}$$

The sum in (2.51) represents a discrete superposition of plane waves. In this discrete superposition, the various coefficients a_i are complex numbers defining the weightings of the individual solutions in the superposition. However, such a discrete superposition never leads to a solution localized in space. We can achieve this with a continuous superposition of the following form:

$$\phi(x,t) = A\int_{-\infty}^{\infty} dk\, a(k)\psi_k(x,t) = A\int_{-\infty}^{\infty} dk\, a(k)e^{i(kx-\omega_k t)} \tag{2.52}$$

Instead of the discrete weighting coefficients a_i in (2.51), a weighting function $a(k)$ must be used in the continuous superposition (2.52). It turns out that the following weighting function is well suited for the formation of localized wave packets:

$$a(k) = \frac{1}{\sqrt{2\pi}}e^{-(k-k_0)^2 d^2} \tag{2.53}$$

This corresponds to a Gaussian function shifted by k_0; $a(k)$ clearly has the property of tending towards zero in infinity, possibly ensuring normalizability over the entire space. By substituting (2.53) into (2.52), we obtain:

$$\phi(x,t) = \frac{A}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk\, e^{-(k-k_0)^2 d^2} e^{i(kx-\omega_k t)} \tag{2.54}$$

For the time $t = 0$, the superposition (2.54) could also be written in the following form:

$$\phi(x,0) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk\, \phi_k(k)e^{ikx} \quad \text{with} \quad \phi_k(k) = Ae^{-(k-k_0)^2 d^2} \tag{2.55}$$

But this now corresponds exactly to the (inverse) Fourier transform of the function $\phi_k(k)$ from the k-space into the real space (with the convention of the pre-factor $1/\sqrt{2\pi}$ in both the forward and the inverse transformation)! This means we can also understand our approach this way: First, we "build" a wave packet in real space using the Gaussian function $\phi_k(k)$ and inverse Fourier transform at the time $t = 0$. In this context, the factor k_0 in the Gaussian function corresponds to the initial momentum $p_0 = \hbar k_0$ of the wave packet. We then ask how this wave packet will evolve over time. Finally, we realize that the term e^{ikx} in the Fourier integral corresponds to a plane wave, and we already know the (trivial) time evolution $e^{-i\omega_k t}$ for plane waves. Therefore, we can simply add $e^{-i\omega_k t}$ to the integrand to describe the time evolution, leading us back to exactly the same solution as before in (2.54):

$$\phi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk\, \phi_k(k) e^{ikx} e^{-i\omega_k t} \quad \text{with} \quad \phi_k(k) = A e^{-(k-k_0)^2 d^2} \tag{2.56}$$

We must still express the dependence between the angular frequency ω_k and the wave number k: It holds $E_k = \hbar\omega_k$ and therefore $\omega_k = E_k/\hbar$; based on $E = p^2/2m$ and $p = \hbar k$, we can thus express the angular frequency ω_k as a function of the wave number k as follows:

$$\omega_k = \frac{E_k}{\hbar} = \frac{p^2}{2m\hbar} = \frac{\hbar^2 k^2}{2m\hbar} = \frac{\hbar k^2}{2m} \tag{2.57}$$

After substituting this relationship into (2.54), we finally derive the following expression for the wave packet $\phi(x,t)$:

$$\phi(x,t) = \frac{A}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk\, e^{-(k-k_0)^2 d^2} e^{i\left(kx - \frac{\hbar k^2}{2m} t\right)} \tag{2.58}$$

By substituting (2.57) into (2.56), we obtain the equivalent representation in the "Fourier view":

$$\phi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk\, \phi_k(k) e^{ikx} e^{-i\frac{\hbar k^2}{2m} t} \quad \text{with} \quad \phi_k(k) = A e^{-(k-k_0)^2 d^2} \tag{2.59}$$

If we substitute (2.59) into the Schrödinger equation (2.9) for a free particle, we can confirm that the wave packet $\phi(x,t)$ is a valid solution of the Schrödinger equation for all times t:

$$\begin{aligned}
-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\phi(x,t) &= i\hbar\frac{d}{dt}\phi(x,t) \quad \overset{(2.59)}{\Longrightarrow} \\
-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk\, \phi_k(k) e^{ikx} e^{-i\frac{\hbar k^2}{2m}t}\right] &= i\hbar\frac{d}{dt}\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk\, \phi_k(k) e^{ikx} e^{-i\frac{\hbar k^2}{2m}t}\right] \quad \Longrightarrow \\
-\frac{\hbar^2}{2m}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk\, \phi_k(k)\left(\frac{d^2}{dx^2}e^{ikx}\right) e^{-i\frac{\hbar k^2}{2m}t} &= i\hbar\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk\, \phi_k(k) e^{ikx}\left(\frac{d}{dt}e^{-i\frac{\hbar k^2}{2m}t}\right) \quad \Longrightarrow \\
\frac{\hbar^2 k^2}{2m}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk\, \phi_k(k) e^{ikx} e^{-i\frac{\hbar k^2}{2m}t} &= \frac{\hbar^2 k^2}{2m}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk\, \phi_k(k) e^{ikx} e^{-i\frac{\hbar k^2}{2m}t} \quad \overset{(2.59)}{\Longrightarrow} \\
\frac{\hbar^2 k^2}{2m}\phi(x,t) &= \frac{\hbar^2 k^2}{2m}\phi(x,t) \quad \square
\end{aligned}$$

We have thus found a solution for the time evolution of the wave packet without explicitly solving the Schrödinger equation as a differential equation.

2.2.1 Calculation of the Wave Packet in Position Space

In what follows, the integral (2.58) shall be executed. To do this, we first rewrite (2.58) as follows:

$$\phi(x,t) \overset{(2.58)}{=} \frac{A}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk\, e^{\zeta(k,x,t)} \quad \text{with} \quad \zeta(k,x,t) = -(k-k_0)^2 d^2 + i\left(kx - \frac{\hbar k^2}{2m}t\right) \tag{2.60}$$

Next, we convert the exponent $\zeta(k,x,t)$ to complete a square in k:

$$\begin{aligned}
\zeta(k,x,t) &= i\left(kx - \frac{\hbar k^2}{2m}t\right) - \left(k^2 - 2kk_0 + k_0^2\right)d^2 = \\
&= ixk - \frac{i\hbar t}{2m}k^2 - d^2k^2 + 2k_0d^2k - k_0^2d^2 = \\
&= -\underbrace{\left(\frac{i\hbar t}{2m} + d^2\right)}_{a}k^2 + \underbrace{\left(ix + 2k_0d^2\right)}_{2b}k - \underbrace{k_0^2d^2}_{c} = \qquad (2.61)\\
&= -ak^2 + 2bk - c \;|\div(-a) \implies \\
\frac{\zeta(k,x,t)}{-a} &= \underbrace{k^2 - 2k\frac{b}{a} + \left(\frac{b}{a}\right)^2}_{(k-b/a)^2} - \left(\frac{b}{a}\right)^2 + \frac{c}{a} = \left(k - \frac{b}{a}\right)^2 - \frac{b^2}{a^2} + \frac{c}{a} \;|\cdot(-a) \implies \\
\zeta(k,x,t) &= -a\left(k - \frac{b}{a}\right)^2 + \frac{b^2}{a} - c \qquad (2.62)
\end{aligned}$$

In placeholders a, b, and c, all terms constant from the perspective of the integral have been combined. The integration variable k now appears only in the expression $-a(k-b/a)^2$. This is advantageous as all terms not dependent on k can be extracted from the integral, and the remaining term of integration over a (shifted) Gaussian curve, for which we already know the result $\sqrt{\pi/a}$:

$$\begin{aligned}
\phi(x,t) &\overset{(2.60)}{=} \frac{A}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} dk\, e^{\zeta(k,x,t)} \overset{(2.62)}{=} \frac{A}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} dk\, e^{-a\left(k-\frac{b}{a}\right)^2 + \frac{b^2}{a} - c} = \\
&= \frac{A}{\sqrt{2\pi}} e^{\frac{b^2}{a}-c} \underbrace{\int_{-\infty}^{+\infty} dk\, e^{-a\left(k-\frac{b}{a}\right)^2}}_{\sqrt{\pi/a}} = \frac{A}{\sqrt{2a}} e^{\frac{b^2}{a}-c} \qquad (2.63)
\end{aligned}$$

The offset b/a does not change the result $\sqrt{\pi/a}$ of the integral in (2.63), which becomes clear when b/a is real-valued (it doesn't matter where the Gaussian function has its maximum, as long as we integrate from $-\infty$ to $+\infty$). Verification shows this is also the case when b/a is complex-valued, as in our case. Substituting the original values for a, b, and c from (2.61) provides the analytical representation of the wave packet in real space:

$$\phi(x,t) = \frac{A}{\sqrt{2\left(\frac{i\hbar t}{2m} + d^2\right)}} \exp\left[\frac{\left(\frac{ix}{2} + k_0d^2\right)^2}{\frac{i\hbar t}{2m} + d^2} - k_0^2d^2\right] \equiv N\, e^{\varphi(x,t)} \qquad (2.64)$$

The wave function $\phi(x,t)$ itself is not directly measurable. However, the probability of finding a particle (based on $\rho = |\phi(x,t)|^2$) can be measured. Thus, we assess the square of the modulus $|\phi(x,t)|^2$.

$$|\phi(x,t)|^2 \overset{(2.64)}{=} |N|^2\,|e^{\varphi}|^2 = |N|^2\, e^{\mathrm{Re}(\varphi)\,+i\,\mathrm{Im}(\varphi)} e^{\mathrm{Re}(\varphi)\,-i\,\mathrm{Im}(\varphi)} = |N|^2\, e^{2\,\mathrm{Re}(\varphi)} \qquad (2.65)$$

First, we handle the factor $|N|^2$ in (2.65):

$$\begin{aligned}
|N|^2 = N^*N &\overset{(2.64)}{=} \frac{A^*A}{\sqrt{2\left(d^2 - \frac{i\hbar t}{2m}\right)2\left(d^2 + \frac{i\hbar t}{2m}\right)}} = \frac{|A|^2}{\sqrt{4\left[d^4 + \frac{\hbar^2t^2}{4m^2}\right]}} = \frac{|A|^2}{\sqrt{4d^4\left[1 + \frac{\hbar^2t^2}{4m^2d^4}\right]}} \implies \\
|N|^2 &= \frac{|A|^2}{2d^2\sqrt{1 + \left(\frac{\hbar t}{2md^2}\right)^2}} \equiv \frac{|A|^2}{2d^2\sqrt{1 + \Delta^2(t)}} \quad \text{with} \quad \Delta(t) \overset{\text{def}}{=} \frac{\hbar t}{2md^2} \qquad (2.66)
\end{aligned}$$

Next, we need to compute the exponent $2\,\mathrm{Re}(\varphi)$ in (2.65):

$$\begin{aligned}
2\,\mathrm{Re}(\varphi) &\overset{(2.64)}{=} 2\,\mathrm{Re}\left[\frac{(\frac{ix}{2}+k_0d^2)^2}{d^2+\frac{i\hbar t}{2m}} - k_0^2d^2\right] = \\
&= 2\,\mathrm{Re}\left[\frac{(-\frac{x^2}{4}+k_0^2d^4+ixk_0d^2)}{(d^2+\frac{i\hbar t}{2m})}\frac{(d^2-\frac{i\hbar t}{2m})}{(d^2-\frac{i\hbar t}{2m})} - k_0^2d^2\right] = \\
&= 2\,\mathrm{Re}\left[\frac{-\frac{x^2d^2}{4}+k_0^2d^6+ixk_0d^4+i\frac{x^2\hbar t}{8m}-i\frac{k_0^2d^4\hbar t}{2m}+\frac{xk_0d^2\hbar t}{2m}}{d^4+(\frac{\hbar t}{2m})^2} - k_0^2d^2\right] = \\
&= 2\left[\frac{-\frac{x^2d^2}{4}+k_0^2d^6+\frac{xk_0d^2\hbar t}{2m}}{d^4\left[1+(\frac{\hbar t}{2md^2})^2\right]}\frac{\frac{1}{d^2}}{\frac{1}{d^2}} - k_0^2d^2\right] = \\
&= 2\left[\frac{-\frac{x^2}{4}+k_0^2d^4+xk_0d^2\frac{\hbar t}{2md^2}}{d^2\left[1+(\frac{\hbar t}{2md^2})^2\right]} - k_0^2d^2\right] = \Big|\,\Delta(t)\overset{\text{def}}{=}\frac{\hbar t}{2md^2} \\
&= 2\,\frac{-\frac{x^2}{4}+k_0^2d^4+xk_0d^2\Delta(t)-k_0^2d^4\left(1+\Delta^2(t)\right)}{d^2\left(1+\Delta^2(t)\right)} = \\
&= -2\frac{1}{4}\,\frac{x^2-4xk_0d^2\Delta(t)+4k_0^2d^4\Delta^2(t)}{d^2\left(1+\Delta^2(t)\right)} = \\
&= -\frac{x^2-4xd^2k_0\Delta(t)+4d^4\left[k_0\Delta(t)\right]^2}{2d^2\left(1+\Delta^2(t)\right)}
\end{aligned} \tag{2.67}$$

For the quantum particle, the momentum relationship $p_0 = \hbar k_0$ holds (where p_0 represents the initial momentum). With $p_0 = mv_0$, this turns into $mv_0 = \hbar k_0$, which simplifies to $k_0 = mv_0/\hbar$ (using v_0 as the initial velocity). The auxiliary quantity $\Delta(t)$ can consequently be rewritten as:

$$k_0\Delta(t) = \frac{mv_0}{\hbar}\frac{\hbar t}{2md^2} = \frac{v_0t}{2d^2}$$

Substituting this back in the numerator of the recently derived expression (2.67) gives:

$$2\,\mathrm{Re}(\varphi) = -\frac{x^2-4xd^2\frac{v_0t}{2d^2}+4d^4\frac{v_0^2t^2}{4d^4}}{2d^2\left(1+\Delta^2(t)\right)} = -\frac{x^2-2xv_0t+v_0^2t^2}{2d^2(1+\Delta^2(t))} = -\frac{(x-v_0t)^2}{2d^2(1+\Delta^2(t))} \tag{2.68}$$

Now we can substitute (2.66) and (2.68) into (2.65), yielding an (as yet unnormalized) expression for $|\phi(x,t)|^2$:

$$|\phi(x,t)|^2 = \frac{|A|^2}{2d^2\sqrt{1+\Delta^2(t)}}e^{-\frac{(x-v_0t)^2}{2d^2(1+\Delta^2(t))}} \tag{2.69}$$

Now let's compare this result with the normalized Gaussian distribution:

$$N(x;\mu_0,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-(x-\mu_0)^2}{2\sigma^2}} \tag{2.70}$$

For the normalized Gaussian distribution $N(x)$, it holds that $\int_{-\infty}^{\infty} dx\, N(x) = 1$. We expect the same for our result $|\phi(x,t)|^2$, which also takes the form of a Gaussian function. By comparing the exponents in (2.70) and (2.69), we can immediately read off the variance σ^2 and the standard deviation σ for $|\phi(x,t)|^2$:

$$\sigma^2 = d^2\left[1+\Delta^2(t)\right] \iff \sigma = d\sqrt{1+\Delta^2(t)} \tag{2.71}$$

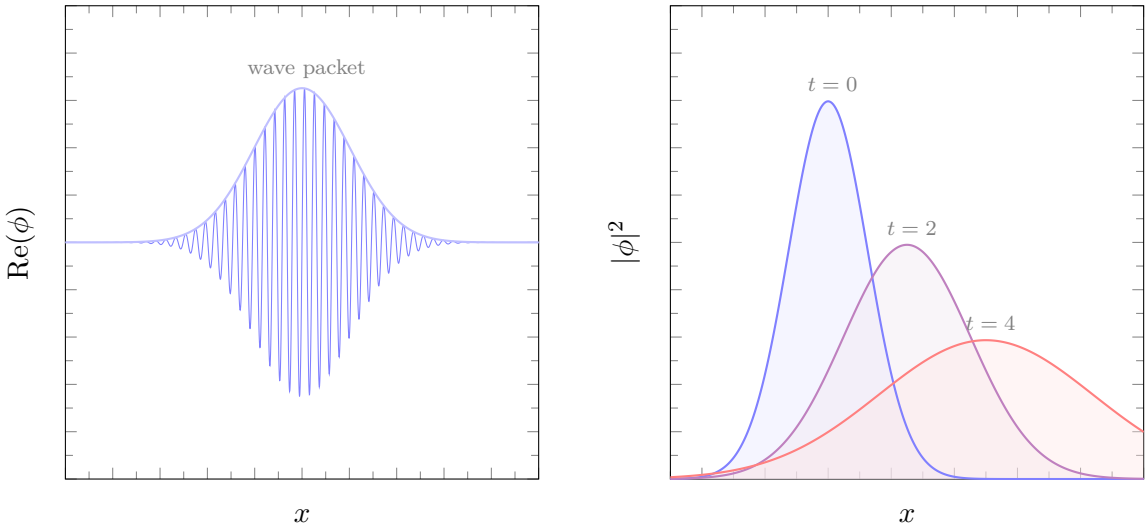


Fig. 6: (left) Representation of the real part of a wave packet as a superposition of plane waves. (right) Schematic representation of the temporal evolution of a wave packet: As time progresses, there is a shift in the maximum, but also a spreading of the probability of presence in space. The norm of the probability density remains conserved.

Substituting this into (2.69), we obtain:

$$|\phi(x,t)|^2 = \frac{|A|^2}{2d\sigma} e^{-\frac{(x-v_0 t)^2}{2\sigma^2}} \tag{2.72}$$

By comparing the prefactors of (2.72) and (2.70), the missing normalization $|A|^2$ can finally be determined:

$$\frac{|A|^2}{2d\sigma} = \frac{1}{\sqrt{2\pi}\sigma} \implies |A|^2 = \frac{2d\sigma}{\sqrt{2\pi}\sigma} = \sqrt{\frac{2}{\pi}} d \tag{2.73}$$

Substituting this into (2.69), we finally obtain the following normalized expression for the probability density of the wave packet:

$$|\phi(x,t)|^2 = \frac{1}{\sqrt{2\pi d^2 (1+\Delta^2(t))}} e^{-\frac{(x-v_0 t)^2}{2d^2(1+\Delta^2(t))}} \quad \text{with} \quad \Delta(t) = \frac{\hbar t}{2md^2} \tag{2.74}$$

Or, even more compactly:

$$|\phi(x,t)|^2 = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-v_0 t)^2}{2\sigma^2}} \quad \text{with} \quad \sigma = d\sqrt{1+\Delta^2(t)} \quad \text{and} \quad \Delta(t) = \frac{\hbar t}{2md^2} \tag{2.75}$$

From (2.74) and (2.75), we can immediately read two facts: Firstly: The maximum of the wave packet moves with $x_0(t) = v_0 t$ to larger x values, which corresponds exactly to the classical equation of motion $x(t) = v_0 t$ for a moving free particle. Secondly: As the wave packet (for positive v_0) moves towards larger x values, the variance $\sigma^2 = d^2(1+\Delta^2(t))$ increases as time t progresses, which means that the wave packet **"spreads out"** in the position representation.

2.2.2 Heisenberg's Uncertainty Principle

Motivation: Thought Experiments on the Uncertainty Principle

In this section, we will demonstrate another fundamental property of quantum physics: The position and momentum of a particle are so-called **"complementary observables"**, which means it is not possible to determine both observables with arbitrary precision. This is not due to technical inadequacies of the measuring instruments used. Rather, it is a fundamental property of quantum systems.

One way to understand this phenomenon visually is to illustrate the process of measuring a location: The most obvious way to determine the location of a particle is to observe the particle. However, this necessarily means interacting with the particle. Because, even if you just "look at" a particle, you have to illuminate it – the light interacts with the particle and changes its momentum. In classical physics, you can make the interaction arbitrarily small by reducing the light intensity as much as you want. In quantum physics, however, you are forced to shoot at least one photon with momentum $|\mathbf{p}| = \hbar k = h/\lambda$ at the presumed location of the particle and then see if the photon was deflected. Therefore, the interaction during the location measurement cannot be reduced arbitrarily! An experimental physicist might come up with the clever idea of using a photon with a larger wavelength to reduce the momentum transfer to the particle. This idea has only one flaw: It also reduces the spatial resolution of the measurement because a photon with a larger wavelength is less well localized.

But what if we shoot a stream of particles in the z-direction at a plate with a very small hole? Suppose all particles have a very precisely determined momentum before they hit the plate. Behind the plate, you can only find those particles that happened to hit the hole exactly. That would mean that the location (at least in the x and y directions, i.e., transverse to the direction of flight) is precisely determined and the momentum was already known before the position measurement! But it turns out: even this attempt to circumvent the laws of quantum physics fails. Because the location of all particles that made it through the hole is so precisely known in the transverse direction, the momentum of these particles in the transverse direction is now particularly uncertain. Each particle flies away in a different transverse direction after the hole. And this is precisely what we would expect from the diffraction of a wave at a small opening! This thought experiment demonstrates that at a fundamental level, a quantum particle does not have a well-defined momentum once its location is known very precisely.

In the previous section, we determined the standard deviation σ of $|\phi(x,t)|^2$ in (2.71). It is the standard deviation of the probability of presence in *position space*. To emphasize this, we will call it σ_x in this section:

$$\sigma_x \overset{(2.71)}{=} d\sqrt{1+\Delta^2(t)} \tag{2.76}$$

We now want to correlate σ_x with the standard deviation σ_k of $|\phi_k(k)|^2$ in *k-space*. We already know the function $\phi_k(k)$ from the **"Fourier perspective"** (2.59). It is important to note that $\phi_k(k)$ does not depend on time because the time evolution in the Fourier integral affects only the position space. First, we determine $|\phi_k(k)|^2$:

$$|\phi_k(k)|^2 \overset{(2.59)}{=} \left|Ae^{-(k-k_0)^2d^2}\right|^2 = |A|^2 e^{-2(k-k_0)^2d^2} \tag{2.77}$$

By comparing the exponent in (2.77) with the exponent of the normalized Gaussian distribution

(2.70), we find:

$$2d^2 = \frac{1}{2\sigma_k^2} \implies \sigma_k^2 = \frac{1}{4d^2} \implies \sigma_k = \frac{1}{2d} \tag{2.78}$$

σ_k does not depend on time, just as $\phi(k)$ doesn't. By forming the product of both standard deviations, we can estimate its minimum size:

$$\sigma_x \sigma_k = \frac{1}{2d} d \underbrace{\sqrt{1 + \Delta^2(t)}}_{\geq 1} \geq \frac{1}{2} \tag{2.79}$$

When transitioning from the standard deviation σ_k in k-space to the standard deviation σ_p in momentum space, an additional factor of $\hbar$ appears in the weight function $a(k)$ from (2.53) due to $p = \hbar k$. This results in the *Heisenberg Uncertainty Principle* named after Werner Heisenberg:

$$\sigma_x \sigma_p \geq \frac{\hbar}{2} \tag{2.80}$$

This means: The smaller σ_x is (i.e., the more precisely the position is determined), the larger σ_p must become to satisfy the uncertainty principle (2.80). Therefore, we cannot determine position and momentum with arbitrary precision. A more accurate measurement of one observable results in the complementary quantity being determined less precisely.

2.2.3 Expectations of the Wave Packet

It has already been shown that the maximum of the Gaussian wave packet in position representation changes over time with $x_0 + v_0 t$. From this, one can conclude (because of the symmetry of the Gaussian function) that this will also apply to the expected value $\langle \hat{x} \rangle$. This will now be explicitly shown in the formalism of quantum mechanics, by integrating x over the entire one-dimensional space with the probability density $\rho(x, t)$:

$$\langle \hat{x} \rangle = \int_{-\infty}^{+\infty} dx\, \rho(x,t) x = \int_{-\infty}^{+\infty} dx\, |\phi(x,t)|^2 x \tag{2.81}$$

The following two remarks apply:

1. The notation $\langle \hat{x} \rangle$ (with "operator hat") suggests that even the position measurement corresponds to an operator $\hat{x}$. In position representation, however, the operator with $\hat{x} \to x$ is quite trivial.

2. If there were only a countable, discrete set of positions x_i, at which the particle could be found with probability $P_i(t)$, one would intuitively calculate the expectation $\langle \hat{x} \rangle$ as a weighted probability as follows:

$$\langle \hat{x} \rangle = \sum_i P_i(t) x_i \tag{2.82}$$

   The integral (2.81) is simply the generalization of the sum (2.82) for an uncountable, continuous set of possible positions.

Continuing the calculation started in (2.81):

$$\begin{aligned}
\langle \hat{x} \rangle &= \int_{-\infty}^{+\infty} dx\, |\phi(x,t)|^2 (x - v_0 t + v_0 t) = \\
&= \int_{-\infty}^{+\infty} dx\, |\phi(x,t)|^2 (x - v_0 t) + v_0 t \underbrace{\int_{-\infty}^{+\infty} dx\, |\phi(x,t)|^2}_{=1 \text{ (normalized)}} \overset{(2.75)}{=} \\
&= \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{+\infty} dx\, e^{-\frac{(x-v_0 t)^2}{2\sigma}} (x - v_0 t) + v_0 t = \quad | \, \xi = x - v_0 t \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \underbrace{\int_{-\infty}^{+\infty} d\xi\, \underbrace{e^{-\frac{\xi^2}{2\sigma}}}_{\text{even}} \cdot \underbrace{\xi}_{\text{odd}}}_{=0} + v_0 t = v_0 t
\end{aligned} \tag{2.83}$$

In the last step, we used the fact that an integral with symmetric limits over the product of an even and an odd function vanishes. As already suspected, the spatial expectation $\langle \hat{x} \rangle = v_0 t$, making it clear that the center of the wave packet follows the classical equation of motion and moves over time in the direction determined by the sign of v_0.

To calculate the momentum expectation value, we use the previously motivated (one-dimensional) momentum operator $\hat{p}$ from (2.14):

$$\hat{p} \to -i\hbar \frac{d}{dx} \equiv \frac{\hbar}{i} \frac{d}{dx}$$

We showed in (2.13) that this momentum operator acts on a plane wave as follows:

$$\frac{\hbar}{i} \frac{d}{dx} e^{i(kx - \omega t)} = p\, e^{i(kx - \omega t)}$$

If we multiply the complex conjugate plane wave on the left side, we obtain as a result the momentum of the individual plane wave:

$$e^{-i(kx-\omega t)} \frac{\hbar}{i} \frac{d}{dx} e^{i(kx-\omega t)} = e^{-i(kx-\omega t)}\, p\, e^{i(kx-\omega t)} = p$$

Now our wave packet $\phi(x,t)$ according to (2.52) is simply at any time a continuous superposition of plane waves. Therefore, we can express the expectation value $\langle \hat{p} \rangle$ again as a weighted average as follows:

$$\langle \hat{p} \rangle = \int_{-\infty}^{\infty} dx\, \phi^*(x,t) \frac{\hbar}{i} \frac{d}{dx} \phi(x,t) \tag{2.84}$$

Comparing this with our approach (2.81) for calculating $\langle \hat{x} \rangle$, we see that we could have also written (2.81) as follows:

$$\langle \hat{x} \rangle = \int_{-\infty}^{\infty} dx\, \phi^*(x,t)\, x\, \phi(x,t) \tag{2.85}$$

The integrals (2.84) and (2.85) have the same structure, and it turns out that one can calculate the expectation value $\langle \hat{A} \rangle$ of any observable $\hat{A}$ for a wave function $\phi(x,t)$ as follows:

$$\langle \hat{A} \rangle = \int_{-\infty}^{\infty} dx\, \phi^*(x,t)\, \hat{A}^{\{x\}}\, \phi(x,t) \tag{2.86}$$

The notation $\hat{A}^{\{x\}}$ is meant to suggest that in the integral, the operator $\hat{A}$ must be written in the position basis. This means that, for example, $\frac{\hbar}{i}\frac{d}{dx}$ can only be used as a momentum operator if the wave function is a function of x, and we integrate over x. If the wave function, for

instance, were a function of p, and we were integrating over p, the momentum operator would look different. We will learn more about this in the next chapter.

Therefore, we use the approach (2.84) and insert the Fourier representation (2.59) of $\phi(x,t)$:

$$\begin{aligned}
\langle\hat{p}\rangle &= \int_{-\infty}^{\infty} dx\, \phi^*(x,t)\left(-i\hbar\frac{d}{dx}\right)\phi(x,t) \overset{(2.59)}{=} \\
&= \int_{-\infty}^{\infty} dx\, \left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk'\, \phi_k(k') e^{ik'x} e^{-i\frac{\hbar k'^2}{2m}t}\right]^* \frac{\hbar}{i}\frac{d}{dx}\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk\, \phi_k(k) e^{ikx} e^{-i\frac{\hbar k^2}{2m}t}\right] = \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\, \left[\int_{-\infty}^{\infty} dk'\, \phi_k^*(k') e^{-ik'x} e^{i\frac{\hbar k'^2}{2m}t}\right]\left[\int_{-\infty}^{\infty} dk\, \phi_k(k)\frac{\hbar}{i}\frac{d}{dx} e^{ikx} e^{-i\frac{\hbar k^2}{2m}t}\right] = \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\, \left[\int_{-\infty}^{\infty} dk'\, \phi_k^*(k') e^{-ik'x} e^{i\frac{\hbar k'^2}{2m}t}\right]\left[\int_{-\infty}^{\infty} dk\, \phi_k(k)\frac{\hbar}{i} ike^{ikx} e^{-i\frac{\hbar k^2}{2m}t}\right] = \\
&= \frac{\hbar}{2\pi}\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dk' \int_{-\infty}^{\infty} dk\, \phi_k^*(k')\phi_k(k)\, e^{-ik'x} e^{ikx}\, e^{i\frac{\hbar k'^2}{2m}t} e^{i\frac{\hbar k^2}{2m}t}\, k = \\
&= \frac{\hbar}{2\pi}\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dk' \int_{-\infty}^{\infty} dk\, \phi_k^*(k')\phi_k(k)\, e^{i(k-k')x}\, e^{i\frac{\hbar}{2m}(k'^2-k^2)t}\, k = \\
&= \frac{\hbar}{2\pi}\int_{-\infty}^{\infty} dk' \int_{-\infty}^{\infty} dk\, \phi_k^*(k')\phi_k(k) \underbrace{\left[\int_{-\infty}^{\infty} dx\, e^{i(k-k')x}\right]}_{2\pi\delta(k-k')} e^{i\frac{\hbar}{2m}(k'^2-k^2)t}\, k = \\
&= \frac{\hbar}{2\pi}\int_{-\infty}^{\infty} dk' \int_{-\infty}^{\infty} dk\, \phi_k^*(k')\phi_k(k)\, 2\pi\delta(k-k') e^{i\frac{\hbar}{2m}(k'^2-k^2)t}\, k = \\
&= \hbar\int_{-\infty}^{\infty} dk\, \phi_k^*(k)\phi_k(k)\, e^{i\frac{\hbar}{2m}(k^2-k^2)t}\, k = \hbar\int_{-\infty}^{\infty} dk\, |\phi_k(k)|^2\, k
\end{aligned} \tag{2.87}$$

A comparison of (2.87) with (2.81) shows that the integrals look entirely equivalent. We therefore attempt a solution strategy analogous to (2.83):

$$\begin{aligned}
\langle\hat{p}\rangle &= \hbar\int_{-\infty}^{+\infty} dk\, |\phi_k(k)|^2 (k - k_0 + k_0) = \\
&= \hbar\int_{-\infty}^{+\infty} dk\, |\phi_k(k)|^2 (k - k_0) + \hbar k_0 \underbrace{\int_{-\infty}^{+\infty} dk\, |\phi_k(k)|^2}_{=1 \text{ (normalized)}} \overset{(2.59)}{=} \\
&= \hbar\int_{-\infty}^{+\infty} dk\, \left|A\, e^{-(k-k_0)^2 d^2}\right|^2 (k - k_0) + \hbar k_0 \\
&= \hbar|A|^2 \int_{-\infty}^{+\infty} dk\, e^{-2(k-k_0)^2 d^2} (k - k_0) + \hbar k_0 = \quad | \, \xi = k - k_0 \\
&= \hbar|A|^2 \underbrace{\int_{-\infty}^{+\infty} d\xi\, \underbrace{e^{-2\xi^2 d^2}}_{\text{even}} \cdot \underbrace{\xi}_{\text{odd}}}_{=0} + \hbar k_0 = \hbar k_0 = p_0 = m v_0
\end{aligned} \tag{2.88}$$

In the last line, the same argument as in (2.83) is applied again: The integration with symmetric limits over the product of an even and an odd function vanishes. The normalization in the second line is a consequence of the *Parseval's equation*:

$$\int_{-\infty}^{+\infty} dx\, |\phi(x,t)|^2 = \int_{-\infty}^{+\infty} dk\, |\phi_k(k,t)|^2 \tag{2.89}$$

The normalization of a state $\phi(x,t)$ remains even after a Fourier transformation to $\phi_k(k,t)$ (and vice versa). Since in (2.74) the position wave function was constructed so that the total integrated probability is equal to one, the probability in momentum space must also lead to the

same result.

In summary, one obtains the following results for the expected values of position and momentum:

$$\langle \hat{x} \rangle = v_0 t \quad \text{and} \quad \langle \hat{p} \rangle = m v_0 = m \frac{d \langle \hat{x} \rangle}{dt}$$

These relations satisfy the equations of motion $p = mv = m\dot{x}$ of classical mechanics. In the lecture `Quantentheorie II` (Quantum Theory II), we will see that this property is related to the Ehrenfest theorem, named after Paul Ehrenfest, which states that for special systems, the expected values of quantum mechanical observables correspond to the classical quantities of the equations of motion. This is the case when we have potentials $V(x)$ that include at most polynomial terms up to the second order in x (i.e., $V(x) \propto x^0, x^1, x^2$). For a potential of the form $V(x) \propto x^3$, the Ehrenfest theorem would no longer be satisfied. In such a case, there would be discrepancies between quantum mechanical and classical measured quantities. In the case of our wave packet, we solved the Schrödinger equation for a free particle ($V(x) = 0$), making the Ehrenfest theorem valid.

2.3 Particle in a Potential Well

Motivation: A Simple Model for Bound Particles Reveals Astonishing Quantum Properties

The term *potential well* refers to a spatial region where a potential $V(\mathbf{r})$ or, in the one-dimensional case, $V(x)$ has a local minimum. Earth's gravitational field is an example of such a potential well in classical physics. The potential of the strong nuclear force, which causes protons to remain together in the atomic nucleus, is an example of a potential well in quantum physics.

In this section, we consider two simple one-dimensional, model potential wells: the infinitely deep potential well and the finitely deep potential well. Both are suitable for demonstrating essential differences between classical physics and quantum physics.

The "infinitely deep potential well" is formed by a potential that is only zero within a limited spatial region and infinite otherwise. One can imagine a quantum particle being perfectly enclosed in this potential well. However, if a quantum particle is confined to a certain spatial region, its position is to some extent determined: It is certainly inside the potential well and not outside. As we previously explained, this means its momentum is no longer precisely determinable! That is, in every measurement, we would find that the particle would have a certain momentum in the negative or positive x-direction. For a quantum particle, it is utterly impossible to simply "lie completely motionless" on the bottom of a potential well! It always has a minimum kinetic energy, and thus (unlike a classical particle) a so-called *ground state energy* that is non-zero. Additionally, we will see that (unlike in classical physics) only discrete (quantized) energy states are allowed.

Similar effects can be observed (although with different numerical values) in the finitely deep potential well, where the potential is again zero within a limited spatial region but otherwise takes on a finite positive value. In such a finitely deep potential well, another effect can be observed that does not exist in classical physics: A particle trapped in the well can penetrate into the walls of the potential well!

2.3.1 Convexity Relations

Based on the stationary Schrödinger equation and the form of the potential, certain statements about the characteristic of the wave function $\psi(x)$ can be made. In the following, the time-independent, one-dimensional Schrödinger equation for stationary states is considered further:

$$\frac{d^2}{dx^2}\psi(x) = \frac{2m}{\hbar^2}\left(V(x) - E\right)\psi(x)$$

The left side of the differential equation is related to the kinetic energy (since the position representation of the quadratic momentum operator is used here and $\hat{p}^2 \propto \hat{E}_{\text{kin}}$ applies) and can be used within the framework of an analysis of the wave function to describe the curvature behavior of $\psi(x)$.

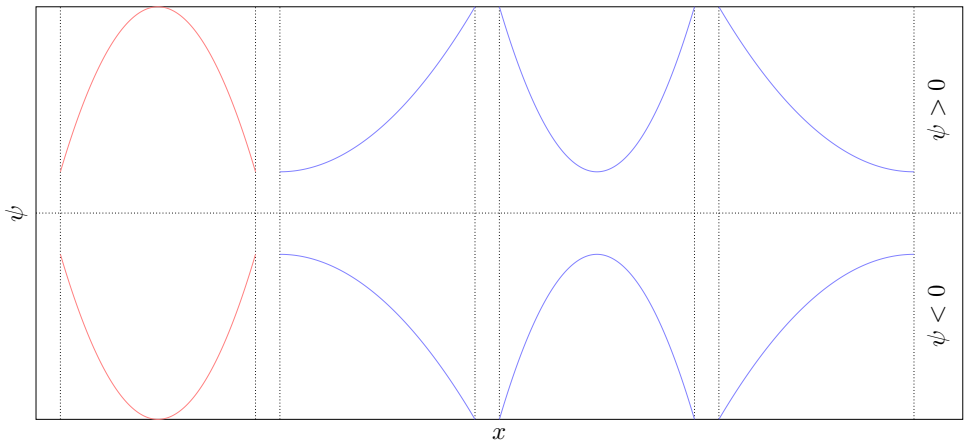


Fig. 7: Curvature behavior of wave functions at different energies. While the red function corresponds to $E > V(x)$, the blue functions concern the case $E < V(x)$.

It is distinguished whether the wave function curves towards or away from the x-axis. This is determined at fixed E by the energy of the potential and the sign of the wave function (assumed here to be a real function); in other words, the cases in which the potential is greater or less than the energy of the particle are distinguished. Specifically, the following cases can be distinguished:

a.) $V(x) < E$: If the energy of the particle exceeds the respective potential, this is referred to as a classically allowed region, and the wave function curves towards the axis. Whether the wave function is concave or convex is determined by the sign of $\psi(x)$:

1.) $\psi(x) > 0$: One obtains a concave wave function, the curvature occurs in the direction of the axis.

2.) $\psi(x) < 0$: One obtains a convex wave function. If one examines the transition between negative and positive wave function, a point of inflection is found at this point.

b.) $V(x) > E$: The energy of the potential exceeds that of the particle, and this is referred to as a classically forbidden region. However, as will be shown in the following chapters, the probability of stay is not zero even in such regions. The convexity behavior is again determined by the sign of the wave function:

3.) $\psi(x) > 0$: One obtains a convex wave function.

4.) $\psi(x) < 0$: One obtains a concave wave function.

2.3.2 Symmetry Relations

In addition to the energy of the system and the sign of the wave function, the actual form of the potential $V(\mathbf{r})$ can also be examined. Particularly interesting is the case in which the potential is completely symmetric, as this property directly carries over to the wave functions as solutions of the Schrödinger equation. That is, if we solve the Schrödinger equation with a symmetric potential $V(\mathbf{r})$, we will obtain wave functions $\psi(x)$ that are either symmetric (even) or antisymmetric (odd):

- $\psi(x) = \psi(-x)$: The wave function is symmetric or even.
- $\psi(x) = -\psi(-x)$: The wave function is antisymmetric or odd.

The reason why the symmetry properties of the potential $V(\mathbf{r})$ influence the wave functions in this manner will be explained more rigorously in later chapters and in the lecture `Quantentheorie II` (Quantum Theory II).

2.3.3 Infinitely Deep Potential Well

Instead of considering a free particle, an interaction with a one-dimensional potential $V(x)$ is introduced for the first time. This has the simple form:

$$V(x) = \begin{cases} 0, & 0 \leq x \leq L \\ \infty, & \text{otherwise} \end{cases} \tag{2.90}$$

The potential well has a width L and is bounded by infinitely high potential walls. Inside the well, $V(x) = 0$, and the Schrödinger equation simplifies to (2.9). However, the limitation due to the potential form now provides boundary conditions. The wave function $\psi(x)$ cannot penetrate into the infinitely high walls. This physical condition, therefore, suggests Dirichlet boundary conditions:

$$\psi(0) = \psi(L) = 0 \tag{2.91}$$

To determine the time-independent eigenfunctions, we can choose as a solution approach to the Schrödinger equation a superposition of a left-traveling and a right-traveling *plane wave*:

$$\psi(x) = Ae^{+ikx} + Be^{-ikx} \tag{2.92}$$

The wavenumber k can be associated with energy using the de Broglie formalism: $k^2 = 2mE/\hbar^2$. Substituting into (2.9), one obtains:

$$\frac{-\hbar^2}{2m}\frac{\partial^2\psi(x)}{\partial x^2} = \frac{-\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\left(Ae^{+ikx} + Be^{ikx}\right) = \frac{-\hbar^2}{2m}(-k^2)\left(Ae^{+ikx} + Be^{ikx}\right) \overset{(2.92)}{=} E\psi(x)$$

The wave function $\psi(x)$ vanishes at the boundaries of the potential, allowing the constants A, B, and k to be determined. It holds:

$$\psi(0) = A + B = 0 \implies A = -B$$

At $x = L$, the wave function also vanishes; we immediately use the result $A = -B$:

$$\psi(L) = A\left(e^{+ikL} - e^{-ikL}\right) = 2iA\sin(kL) = C\sin(kL) = 0 \implies k_n = \frac{n\pi}{L}$$

The sine function vanishes only when the argument kL takes whole multiples of π, where $n \in \mathbb{N}^+$ (even $n = 0$ is not a valid solution, as in this case $\psi(x)$ would not be normalizable). We denote n as a *quantum number* - this means that n determines the microscopic state of a quantum system but can only take discrete values. The constant C must now be determined by normalizing $\psi(x)$:

The particle is with absolute certainty inside the potential well, as it cannot penetrate into its walls. The probability of stay is therefore:

$$\frac{1}{C^2} \stackrel{!}{=} \int_0^L dx\, \sin^2(k_n x) = \frac{1}{2}\int_0^L dx\, [1 - \cos(2k_n x)] = \frac{L}{2}$$

The wave function that describes the particle in the potential well is thus given by the following real function:

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi}{L}x\right) \tag{2.93}$$

From (2.93) and the graphically represented probability of stay in Figure 8, the convexity behavior of the wave function can also be recognized: For the infinitely high potential well, only a classically allowed region is possible, so $\psi(x)$ always curves towards the axis.
The quantized values of k_n are a consequence of the boundary conditions and accordingly translate to the energies of the stationary states n:

$$E_n = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2 \pi^2}{2mL^2} n^2 \tag{2.94}$$

The energy E_n thus increases quadratically with n, and energy levels can only exist at permitted n values.

The set of these eigenstates $\psi_n(x)$ forms a complete orthonormal system. The orthogonality of the wave functions will be demonstrated in the following:

$$\begin{aligned}
\psi_n(x)\psi_m(x) &= \frac{2}{L}\int_0^L dx \sin\left(\frac{\pi n x}{L}\right)\sin\left(\frac{\pi m x}{L}\right) = \\
&= \frac{1}{L}\int_0^L dx\left[\cos\left(\frac{\pi(n-m)x}{L}\right) - \cos\left(\frac{\pi(n+m)x}{L}\right)\right] = \\
&= \frac{1}{L}\left[\frac{L\sin(\pi(n-m)x/L)}{\pi(n-m)} - \frac{L\sin(\pi(n+m)x/L)}{\pi(n+m)}\right]_0^L = \\
&= \frac{\sin(\pi(n-m))}{\pi(n-m)} - \frac{\sin(\pi(n+m))}{\pi(n+m)}
\end{aligned}$$

This relation vanishes for all $n \neq m$, as each sine function involves a natural multiple of π - only if $n = m$ the first term must be further considered. Since in this case, zero is divided by zero, L'Hôpital's rule can be applied (let it be $x = n - m$):

$$\lim_{x\to 0} \frac{\frac{d}{dx}\sin(\pi x)}{\frac{d}{dx}\pi x} = \lim_{x\to 0} \frac{\pi\cos(\pi x)}{\pi} = 1$$

The product only does not vanish if the indices are equal and thus represent the same wave function, allowing us to speak of orthogonality. Due to the fact that $\psi_n(x)$ is also normalized, it even holds orthonormality:

$$\psi_n(x)\psi_m(x) = \delta_{nm} \tag{2.95}$$

The number of zero crossings $\tilde{n}_0$ of $\psi_n(x)$ can be used to establish an energetic hierarchy of the system (*nodel theorem*). The larger $\tilde{n}_0$ becomes, the higher the frequency of the eigenstate and thus its respective energy. The *principal* or *energy quantum number* n can be determined over:

$$n = \tilde{n}_0 + 1 \tag{2.96}$$

Counting the zero crossings (or *nodes*) of the wave function, we can thus determine the quantum number n of a state in an infinitely deep potential well.

2.3.4 Finitely Deep Potential Well

Again, a particle is to be trapped in a one-dimensional potential well, but this time it has a finite depth:

$$V(x) = \begin{cases} 0, & -L \leq x \leq L \\ V_0 > 0, & \text{otherwise} \end{cases} \tag{2.97}$$

Two cases can generally be distinguished here: The energy of the particle E exceeds or falls below the energy of the potential V_0. The case where $E > V_0$ will not be addressed in the following, as it leads to no bound state but to a non-normalizable scattering state. When $E < V_0$, the particle cannot leave the potential well, and there exist normalizable eigenstates at discrete energies.
Unlike in (2.91), the wave function of the particle can now penetrate the finite potential barrier - to find a solution, the entire space can be divided into three regions: $\psi_{\mathrm{I}}(x)$ in $(-\infty, -L)$, $\psi_{\mathrm{II}}(x)$ in $[-L, L]$, and $\psi_{\mathrm{III}}(x)$ in (L, ∞). The wave function $\psi_{\mathrm{II}}(x)$ corresponds to that of a free particle, and the wave number k yields the known de Broglie relation:

$$k = \frac{\sqrt{2mE}}{\hbar} \tag{2.98}$$

For the wave functions $\psi_{\mathrm{I}}(x)$ and $\psi_{\mathrm{III}}(x)$, the finite potential needs to be considered in both cases in the same way. The Schrödinger equation corresponds in this case to:

$$\frac{d^2}{dx^2}\psi_{\mathrm{I,III}}(x) = \frac{2m}{\hbar^2}\left(V_0 - E\right)\psi_{\mathrm{I,III}}(x)$$

The energy terms on the right side of the equation can be replaced by a constant, which is certainly a positive quantity due to $E < V_0$ - for solving the differential equation, a real exponential function is chosen. For the constant κ, it holds:

$$\kappa = \frac{\sqrt{2m(V_0 - E)}}{\hbar} \tag{2.99}$$

A general approach for a solution can then be given:

$$\psi(x) = \begin{cases} A_1 e^{-\kappa x} + A_2 e^{+\kappa x}, & x < -L \\ B_1 e^{-ikx} + B_2 e^{+ikx}, & -L \leq x \leq L \\ C_1 e^{-\kappa x} + C_2 e^{+\kappa x}, & L < x \end{cases} \tag{2.100}$$

For the wave functions $\psi_{\mathrm{I}}(x)$ and $\psi_{\mathrm{III}}(x)$ in the potential, it must hold that $A_1 = C_2 = 0$, as the wave function must not grow exponentially (otherwise it would no longer be normalizable). At the junctions $x = -L$ and $x = L$, however, it must hold:

$$\begin{aligned} \psi_{\mathrm{I}}(-L) &= \psi_{\mathrm{II}}(-L) \quad \text{and} \quad \psi_{\mathrm{II}}(L) = \psi_{\mathrm{III}}(L) \\ \psi'_{\mathrm{I}}(-L) &= \psi'_{\mathrm{II}}(-L) \quad \text{and} \quad \psi'_{\mathrm{II}}(L) = \psi'_{\mathrm{III}}(L) \end{aligned} \tag{2.101}$$

The transition should occur without discontinuities and be continuously differentiable.
A further simplification of the problem can be made by incorporating the symmetry of the potential. The potential in (2.97) is symmetric around $x = 0$, which results in the entire Hamiltonian operator $\hat{H}$ exhibiting these symmetry properties. In later chapters, it will be shown that the symmetry of the Hamiltonian operator impacts the eigenstates: only symmetric or antisymmetric eigenstates can exist, making it clear that $A_2 = C_1 = A$ must hold. For the symmetric or even solution, it generally holds $f(x) = f(-x)$, allowing the wave function $\psi_+(x)$ to be written as:

$$\psi_+(x) = \begin{cases} Ae^{+\kappa x}, & x < -L \\ B\cos(kx), & -L \leq x \leq L \\ Ae^{-\kappa x}, & L < x \end{cases}$$

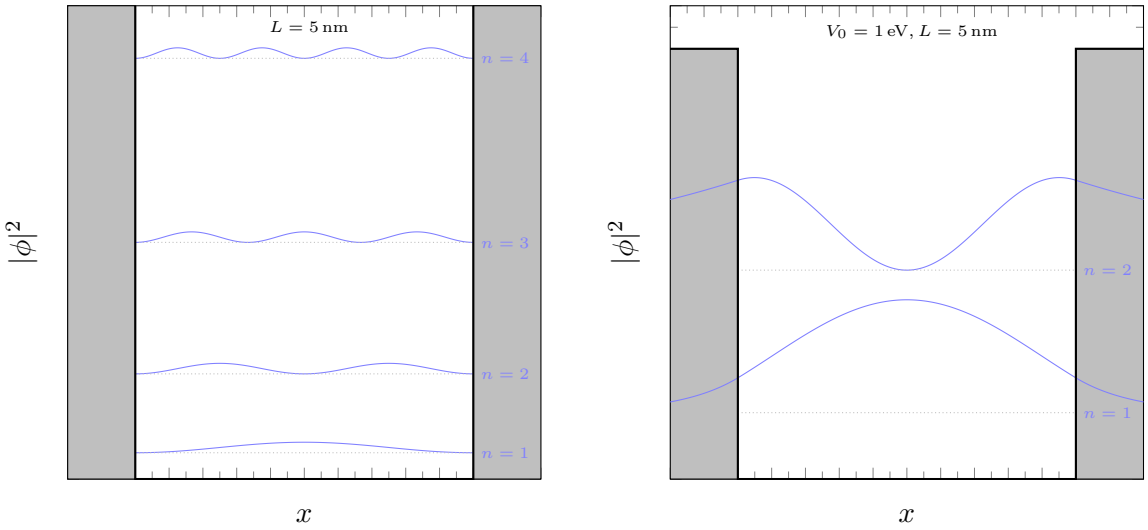


Fig. 8: (left) Infinitely deep potential well with the probabilities of stay of stationary states with the four lowest eigenenergies. (right) The finite potential well has, for a given potential depth V_0 and width L, only a finite number of bound states, whose probabilities of stay are depicted here.

For the antisymmetric or odd solution, it generally holds $f(x) = -f(-x)$, resulting in the waveform $\psi_-(x)$:

$$\psi_-(x) = \begin{cases} -Ae^{+\kappa x}, & x < -L \\ B\sin(kx), & -L \leq x \leq L \\ Ae^{-\kappa x}, & L < x \end{cases}$$

The boundary conditions allow us to express A through B; B can later be determined through normalization conditions:

$$\psi_{\mathrm{II},+}(L) = \psi_{\mathrm{III},+}(L) \iff B\cos(kL) = Ae^{-\kappa L} \quad \text{thus} \quad A = B\cos(kL)e^{+\kappa L}$$

Additionally, the differential boundary condition allows us to establish a connection between k and κ:

$$\frac{\psi'_{\mathrm{II},+}(L)}{\psi_{\mathrm{II},+}(L)} = \frac{\psi'_{\mathrm{III},+}(L)}{\psi_{\mathrm{III},+}(L)} \iff \frac{-k\sin(kL)}{\cos(kL)} = \frac{-\kappa Ae^{-\kappa L}}{Ae^{-\kappa L}} \quad \text{so} \quad k\tan(kL) = \kappa$$

Through the given parameters E and V_0, constraints for k and κ can be defined - only certain (but arbitrarily many) pairs of k and κ are possible through V_0:

$$k^2 + \kappa^2 = \frac{2m}{\hbar^2}V_0 = \eta^2$$

Thus, one has two conditions available for k and κ to calculate concrete values. For even eigenfunctions, it thus holds:

$$k\tan(kL) = \kappa \quad \text{and} \quad k^2 + \kappa^2 = \eta^2 \tag{2.102}$$

Depending on the depth V_0 and width L of the potential, only finitely many pairs of k and κ can be found. For odd eigenfunctions, one obtains a similar relationship:

$$-k\cot(kL) = \kappa \quad \text{and} \quad k^2 + \kappa^2 = \eta^2 \tag{2.103}$$

Both (2.102) and (2.103) provide solutions for k and κ only after numerical (or graphical) evaluation. Unlike with the square dependence on n in the infinite potential well from (2.94), the

eigenenergies follow no clear progression of a quantum number n. The wider the potential well becomes, the smaller the spacing between energy levels, and the number of possible eigenstates increases. This is also the case when V_0 increases. A calculated example of the probability density of a particle in two possible states is given in Figure 8. The graphical ordering of the wave function is justified by the node count (or in the case of the probability density by the zero points) according to (2.96).

2.4 Particles at the Potential Barrier

Motivation: Scattering Phenomena and Tunnel Effect

While we dealt in the previous sections with potential wells where a particle is bound in a specific spatial region, we now examine the effects that occur when particles coming from infinity encounter simple *potential barriers* or potential steps.

In the following sections, not only will important mathematical tools for dealing with such cases (such as the transfer and scattering matrix) be presented, but the astonishing phenomenon of the tunnel effect will also be described. In this phenomenon, a particle – despite not having sufficient energy according to the laws of classical physics – can penetrate a potential barrier (in other words, "tunnel through").

2.4.1 Particles at the Potential Step

Let us consider, as the simplest case, a one-dimensional potential $V(x)$ with the energy height V_0, which has the following step form (see Fig. 9, left):

$$V(x) = \begin{cases} 0, & x < 0 \ \text{(Region I)} \\ V_0, & x \geq 0 \ \text{(Region II)} \end{cases} \tag{2.104}$$

It is assumed that the kinetic energy satisfies $E > V_0$ (this is referred to as *superbarrier scattering*). The entire available space can be divided into two subregions: In Region I $(-\infty, 0)$ the wave function $\psi_{\mathrm{I}}(x)$ holds, and in Region II $[0, +\infty)$ the wave function $\psi_{\mathrm{II}}(x)$. Assuming an incoming wave from the left, we expect it to be partially reflected, meaning $\psi_{\mathrm{I}}(x)$ is described by both an incoming and an outgoing wave. On the other hand, Region II with the wave function $\psi_{\mathrm{II}}(x)$ will only be characterized by an outgoing wave. We therefore choose the Ansatz:

$$\psi(x) = \begin{cases} \psi_{\mathrm{I}}(x) & x < 0 \\ \psi_{\mathrm{II}}(x) & x \geq 0 \end{cases} \quad \text{with} \quad \begin{aligned} \psi_{\mathrm{I}}(x) &= Ae^{+ikx} + Be^{-ikx} \\ \psi_{\mathrm{II}}(x) &= Ce^{+ik'x} \end{aligned} \tag{2.105}$$

Anticipating the next step, we calculate the first spatial derivative of $\psi_{\mathrm{I}}(x)$ and $\psi_{\mathrm{II}}(x)$:

$$\begin{aligned} \psi'_{\mathrm{I}}(x) &= ikAe^{+ikx} - ikBe^{-ikx} \\ \psi'_{\mathrm{II}}(x) &= ik'Ce^{+ik'x} \end{aligned} \tag{2.106}$$

The wave number k is equivalent to the expression already found in (2.98), while for k', from the Schrödinger equation in analogy to (2.99), the following relation is obtained:

$$k' = \sqrt{\frac{2m(E - V_0)}{\hbar^2}} \tag{2.107}$$

The boundary conditions (continuity of the wave function and its derivative) are as follows:

$$\psi_{\mathrm{I}}(0) = \psi_{\mathrm{II}}(0) \quad \text{and} \quad \psi'_{\mathrm{I}}(0) = \psi'_{\mathrm{II}}(0) \tag{2.108}$$

From the first of these conditions in (2.108), a direct connection between the amplitudes of the wave functions is derived:

$$\psi_{\mathrm{I}}(0) = \psi_{\mathrm{II}}(0) \overset{(2.105)}{\Longrightarrow} A + B = C \tag{2.109}$$

The differential boundary condition from (2.108) provides another relation:

$$\psi'_{\mathrm{I}}(0) = \psi'_{\mathrm{II}}(0) \overset{(2.106)}{\Longrightarrow} ikA - ikB = ik'C \quad \Rightarrow \quad A - B = \frac{k'}{k}C \tag{2.110}$$

The amplitudes B and C are to be expressed in terms of the amplitude A of the incoming wave. Inserting (2.109) into (2.110) yields:

$$A - B = \frac{k'}{k}(A + B) \quad \Rightarrow \quad kA - kB = k'A + k'B \quad \Rightarrow \quad B = \frac{k - k'}{k + k'}A \tag{2.111}$$

By substituting (2.111) into the first boundary condition (2.109), an expression for the amplitude C of the outgoing wave is obtained:

$$C = A + \frac{k - k'}{k + k'}A = \frac{k + k' + k - k'}{k + k'}A \quad \Rightarrow \quad C = \frac{2k}{k + k'}A \tag{2.112}$$

Since a plane wave over the entire space cannot be normalized, A remains undetermined. The solution to the scattering problem can thus be represented by inserting (2.111) and (2.112) into the Ansatz (2.105) as follows:

$$\psi(x) = \begin{cases} \psi_{\mathrm{I}}(x) & x < 0 \\ \psi_{\mathrm{II}}(x) & x \geq 0 \end{cases} \quad \text{with} \quad \begin{aligned} \psi_{\mathrm{I}}(x) &= Ae^{+ikx} + A\frac{k - k'}{k + k'}e^{-ikx} \\ \psi_{\mathrm{II}}(x) &= A\frac{2k}{k + k'}e^{+ik'x} \end{aligned} \tag{2.113}$$

How can we understand the wave function $\psi(x)$ from (2.113)? In the region $x < 0$, we deal with the superposition of a wave moving to the right (incoming) and a wave moving to the left (reflected).

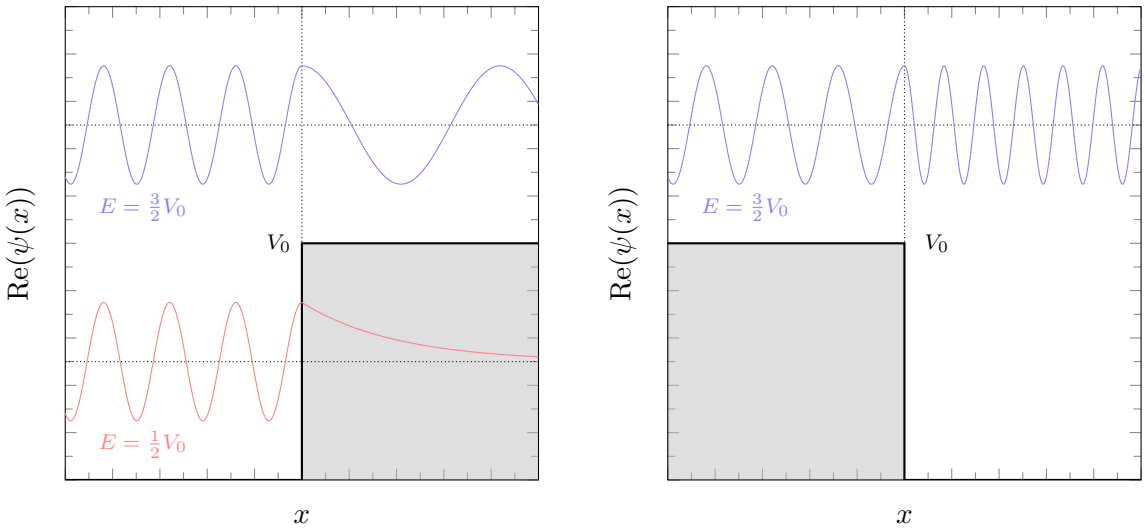


Fig. 9: (left) A particle wave incoming from the left interacts with a positive potential step, and depending on the energy E, either undergoes a frequency reduction or exponential decay. (right) Viewing a negative potential step, the frequency increases.

Note, however, that the direction of motion of the plane waves becomes clear only in the context of the time-dependent Schrödinger equation, which according to (2.36) provides a phase factor

$e^{-i\omega t}$ (thus $e^{i(kx-\omega t)}$ is a wave moving to the *right*). The amplitude of the reflected, left-moving part depends on the ratio of the energy E to the step height V_0. In the region of the potential ($x \geq 0$), there is a change in the frequency of the non-reflected component of the plane wave. Depending on the sign of the potential V_0, there is either an increase or decrease. With the wave vector k' from (2.107), one obtains:

$$\nu' = \frac{c}{\lambda'} = \frac{c}{2\pi}k' = \frac{c}{2\pi\hbar}\sqrt{2m(E-V_0)} \tag{2.114}$$

2.4.2 Transfer and Scattering Matrix

Let us consider a one-dimensional potential barrier $V(x)$ with finite width L and height V_0 (see Fig. 10, left):

$$V(x) = \begin{cases} 0, & x < 0 \quad \text{(Region I)} \\ V_0, & 0 \leq x \leq L \quad \text{(Region II)} \\ 0, & x > 0 \quad \text{(Region III)} \end{cases} \tag{2.115}$$

In this example, we assume that $E < V_0$, which causes an exponential decay of the wave function in the barrier (this is also known as *sub-barrier scattering*). The space is again divided into three distinct regions: Region I $(-\infty, 0)$ with $\psi_\text{I}(x)$, Region II $[0, L]$ with $\psi_\text{II}(x)$, and Region III (L, ∞) with $\psi_\text{III}(x)$. As in the previous chapter, the constants k and κ can be determined from the Schrödinger equation, which also match in this case with (2.98) and (2.99):

$$k = \frac{\sqrt{2mE}}{\hbar} \quad \text{and} \quad \kappa = \frac{\sqrt{2m(V_0-E)}}{\hbar}$$

The following solution approach to the Schrödinger equation for the potential barrier is chosen, assuming this time that there can be incoming, and therefore also reflected (outgoing) waves both from the left (in Region I) and from the right (in Region III):

$$\psi(x) = \begin{cases} \psi_\text{I}(x), & x < 0 \\ \psi_\text{II}(x), & 0 \leq x \leq L \\ \psi_\text{III}(x), & x > L \end{cases} \quad \text{with} \quad \begin{aligned} \psi_\text{I}(x) &= Ae^{+ikx} + Be^{-ikx} \\ \psi_\text{II}(x) &= Ce^{-\kappa x} + De^{+\kappa x} \\ \psi_\text{III}(x) &= Fe^{+ikx} + Ge^{-ikx} \end{aligned} \tag{2.116}$$

Again, we calculate the first spatial derivative of ψ_I, ψ_II, and ψ_III for later use:

$$\begin{aligned} \psi'_\text{I}(x) &= ikAe^{+ikx} - ikBe^{-ikx} \\ \psi'_\text{II}(x) &= -\kappa Ce^{-\kappa x} + \kappa De^{+\kappa x} \\ \psi'_\text{III}(x) &= ikFe^{+ikx} - ikGe^{-ikx} \end{aligned} \tag{2.117}$$

The following boundary conditions apply:

$$\begin{aligned} \psi_\text{I}(0) &= \psi_\text{II}(0) \quad \text{and} \quad \psi_\text{II}(L) = \psi_\text{III}(L) \\ \psi'_\text{I}(0) &= \psi'_\text{II}(0) \quad \text{and} \quad \psi'_\text{II}(L) = \psi'_\text{III}(L) \end{aligned} \tag{2.118}$$

The boundary conditions for $x = 0$ lead to:

$$\begin{aligned} &\text{(I)}| \quad \psi_\text{I}(0) = \psi_\text{II}(0) \overset{(2.116)}{\Longrightarrow} A + B = C + D \\ &\text{(II)}| \quad \psi'_\text{I}(0) = \psi'_\text{II}(0) \overset{(2.117)}{\Longrightarrow} ik(A-B) = \kappa(D-C) \implies A - B = \frac{\kappa}{ik}(D-C) \end{aligned}$$

The addition of both equations, or the subtraction of the second equation from the first, respectively gives:

$$\text{(I)+(II)}|\quad 2A = C + D + \frac{\kappa}{ik}(D - C) = \frac{ikC + ikD + \kappa D - \kappa C}{ik} = \frac{ik - \kappa}{ik}C + \frac{ik + \kappa}{ik}D \implies$$

$$A = \frac{1}{2ik}\left[(ik - \kappa)C + (ik + \kappa)D\right] \tag{2.119}$$

$$\text{(I)–(II)}|\quad 2B = C + D - \frac{\kappa}{ik}(D - C) = \frac{ikC + ikD - \kappa D + \kappa C}{ik} = \frac{ik + \kappa}{ik}C + \frac{ik - \kappa}{ik}D \implies$$

$$B = \frac{1}{2ik}[(ik + \kappa)C + (ik - \kappa)D] \tag{2.120}$$

This can be written more compactly as a matrix equation:

$$\begin{pmatrix} A \\ B \end{pmatrix} = \frac{1}{2ik}\begin{pmatrix} ik - \kappa & ik + \kappa \\ ik + \kappa & ik - \kappa \end{pmatrix}\begin{pmatrix} C \\ D \end{pmatrix} = \underbrace{\frac{1}{2k}\begin{pmatrix} k + i\kappa & k - i\kappa \\ k - i\kappa & k + i\kappa \end{pmatrix}}_{P}\begin{pmatrix} C \\ D \end{pmatrix} = P\begin{pmatrix} C \\ D \end{pmatrix} \tag{2.121}$$

Here, the matrix P is a *transfer matrix*, which immediately relates the amplitudes in Region I to the amplitudes in Region II. For the boundary condition at $x = L$, the relations are a bit more complicated due to the non-vanishing exponential function. To simplify the next step, the first equation is first multiplied by κ:

$$\text{(III)}|\quad \kappa\psi_{\mathrm{II}}(L) = \kappa\psi_{\mathrm{III}}(L) \overset{(2.116)}{\implies} \kappa Ce^{-\kappa L} + \kappa De^{+\kappa L} = \kappa Fe^{+ikL} + \kappa Ge^{-ikL}$$

$$\text{(IV)}|\quad \kappa\psi'_{\mathrm{II}}(L) = \kappa\psi'_{\mathrm{III}}(L) \overset{(2.117)}{\implies} -\kappa Ce^{-\kappa L} + \kappa De^{+\kappa L} = ikFe^{+ikL} - ikGe^{-ikL}$$

The addition of both equations, or the subtraction of the second equation from the first, gives:

$$\text{(III)–(IV)}|\quad 2C\kappa e^{-\kappa L} = (\kappa - ik)Fe^{+ikL} + (\kappa + ik)Ge^{-ikL}$$

$$\text{(III)+(IV)}|\quad 2D\kappa e^{+\kappa L} = (\kappa + ik)Fe^{+ikL} + (\kappa - ik)Ge^{-ikL}$$

By simple rearrangement, one obtains the following expressions for C and D:

$$C = \frac{\kappa - ik}{2\kappa}Fe^{+ikL+\kappa L} + \frac{\kappa + ik}{2\kappa}Ge^{-ikL+\kappa L}$$

$$D = \frac{\kappa + ik}{2\kappa}Fe^{+ikL-\kappa L} + \frac{\kappa - ik}{2\kappa}Ge^{-ikL-\kappa L}$$

Here too, the equation system can be more clearly presented in matrix form:

$$\begin{pmatrix} C \\ D \end{pmatrix} = \underbrace{\frac{1}{2\kappa}\begin{pmatrix} (\kappa - ik)e^{+L(ik+\kappa)} & (\kappa + ik)e^{-L(ik-\kappa)} \\ (\kappa + ik)e^{+L(ik-\kappa)} & (\kappa - ik)e^{-L(ik+\kappa)} \end{pmatrix}}_{Q}\begin{pmatrix} F \\ G \end{pmatrix} = Q\begin{pmatrix} F \\ G \end{pmatrix} \tag{2.122}$$

The matrix Q is also a *transfer matrix*. It relates the amplitudes in Region II to the amplitudes in Region III. Transfer matrices have the practical property that they can be easily multiplied to obtain a superior transfer matrix. Thus, if you multiply the transfer matrices P and Q, you obtain a new transfer matrix M, which describes the overall relationship between the amplitudes in Region I (left of the barrier) and the amplitudes in Region III (right of the barrier):

$$\begin{pmatrix} A \\ B \end{pmatrix} = \underbrace{PQ}_{M}\begin{pmatrix} F \\ G \end{pmatrix} = M\begin{pmatrix} F \\ G \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}\begin{pmatrix} F \\ G \end{pmatrix} \tag{2.123}$$

A *transfer matrix* therefore always relates the amplitudes *on the right* to the amplitudes *on the left* of a scatterer. If you already know the amplitudes F and G of the outgoing and incoming

wave on the *right* side of the barrier, then the transfer matrix M can be used to calculate the amplitudes A and B of the incoming and outgoing wave on the *left* side of the barrier. With the matrix M^{-1}, the calculation is possible in reverse direction. However, this is obviously not very helpful for most problem settings. If you want to use the transfer matrix M, for example, you must already know the amplitude of the outgoing wave on the right side. However, this will be somehow derived from the reflection of the incoming wave from the right and the transmission of the incoming wave from the left. It would therefore be better if there were a matrix that could calculate the unknown amplitudes of the outgoing waves on the left and right (B and F) from the known amplitudes of the incoming waves on the left and right (A and G). Such a matrix is called a *scattering matrix*. We want to denote it with S:

$$\begin{pmatrix} B \\ F \end{pmatrix} = S \begin{pmatrix} A \\ G \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} A \\ G \end{pmatrix} \tag{2.124}$$

We will now attempt to express the components S_{ij} of the *scattering matrix* S in terms of the components M_{ij} of the *transfer matrix* M. For this, we first write the matrix equation (2.123) of the transfer matrix M as a system of equations:

$$A = M_{11}F + M_{12}G \tag{2.125}$$
$$B = M_{21}F + M_{22}G \tag{2.126}$$

In contrast, the matrix equation (2.124) of the scattering matrix S corresponds to the following system of equations:

$$B = S_{11}A + S_{12}G \tag{2.127}$$
$$F = S_{21}A + S_{22}G \tag{2.128}$$

Our goal should be to express the amplitudes B and F as a function of the amplitudes A and G, but only using the matrix entries M_{ij}. For F, this is simple: We just need to rearrange equation (2.125) accordingly:

$$F = \frac{1}{M_{11}}A - \frac{M_{12}}{M_{11}}G \tag{2.129}$$

By comparing coefficients between (2.129) and (2.128), we can immediately deduce:

$$S_{21} = \frac{1}{M_{11}} \quad \text{and} \quad S_{22} = -\frac{M_{12}}{M_{11}} \tag{2.130}$$

Finally, inserting the relation (2.129) into equation (2.126), we obtain:

$$B = M_{21}\left(\frac{1}{M_{11}}A - \frac{M_{12}}{M_{11}}G\right) + M_{22}G = \frac{M_{21}}{M_{11}}A - \frac{M_{21}M_{12}}{M_{11}}G + M_{22}G \implies$$
$$B = \frac{M_{21}}{M_{11}}A + \frac{M_{22}M_{11} - M_{21}M_{12}}{M_{11}}G \tag{2.131}$$

A further coefficient comparison between (2.131) and (2.127) finally gives the following expressions for the matrix components S_{11} and S_{12}:

$$S_{11} = \frac{M_{21}}{M_{11}} \quad \text{and} \quad S_{12} = \frac{M_{22}M_{11} - M_{21}M_{12}}{M_{11}} \tag{2.132}$$

For the scattering matrix S, expressed in the components of the transfer matrix M, the following holds according to (2.130) and (2.132):

$$S = \frac{1}{M_{11}} \begin{pmatrix} M_{21} & M_{22}M_{11} - M_{21}M_{12} \\ 1 & -M_{12} \end{pmatrix} \tag{2.133}$$

2.4.3 Transmission and Reflection Probability

We will now consider how large the reflection and transmission probability is for an incoming particle at a potential barrier according to (2.115), depicted in Fig. 10 (left). To do this, let's assume that a particle approaches the potential barrier from the left. There is a certain probability of transmission or reflection. Because we assume that the particle is only incoming from the left (but no particle from the right), we can set the amplitude G of the wave incoming from the right in the Ansatz (2.116) to $G = 0$. Doing this, the equations (2.127) and (2.128) simplify as follows:

$$B = S_{11}A \tag{2.134}$$

$$F = S_{21}A \tag{2.135}$$

The simplified scattering equation (2.135) allows us to define and calculate the complex-valued *transmission coefficient* for transmission from left to right $t_{\mathrm{L}\to\mathrm{R}}$ as the ratio between the amplitude F of the *right outgoing* wave and the amplitude A of the *left incoming* wave. The square of the magnitude of this quantity leads to the *transmission probability* $T_{\mathrm{L}\to\mathrm{R}}$:

$$t_{\mathrm{L}\to\mathrm{R}} = \frac{F}{A} \overset{(2.135)}{=} S_{21} \quad \text{and} \quad T_{\mathrm{L}\to\mathrm{R}} = |t_{\mathrm{L}\to\mathrm{R}}|^2 = |S_{21}|^2 \overset{(2.130)}{=} \left|\frac{1}{M_{11}}\right|^2 \tag{2.136}$$

The complex-valued reflection coefficient r_{L} for particles incoming from the left is instead defined as the ratio between the amplitude B of the *left outgoing* wave and the amplitude A of the *left incoming* wave. Using equation (2.134), an expression can also be found for this. In a manner equivalent to before, the *reflection probability* R_{L} can also be calculated:

$$r_{\mathrm{L}} = \frac{B}{A} \overset{(2.134)}{=} S_{11} \quad \text{and} \quad R_{\mathrm{L}} = |r_{\mathrm{L}}|^2 = |S_{11}|^2 \overset{(2.132)}{=} \left|\frac{M_{21}}{M_{11}}\right|^2 \tag{2.137}$$

For the transmission coefficient $t_{\mathrm{L}\to\mathrm{R}} \in \mathbb{C}$, it holds that $\psi_{\mathrm{III}}(x) = t_{\mathrm{L}\to\mathrm{R}}\, A\, e^{+ikx}$ (in the case of an exclusively left incoming wave). The transmission probability $T_{\mathrm{L}\to\mathrm{R}} \in \mathbb{R}$ must lie within the interval $T_{\mathrm{L}\to\mathrm{R}} \in [0, 1]$.

Now let's assume that a particle exclusively hits the potential barrier from the right. This particle will either be transmitted from right to left or reflected back to the right. Since we assume that this time no particle is incoming from the left, we can now set the amplitude $A = 0$ in the Ansatz (2.116). This simplifies the equations (2.127) and (2.128) as follows:

$$B = S_{12}G \tag{2.138}$$

$$F = S_{22}G \tag{2.139}$$

The *transmission coefficient* from right to left $t_{\mathrm{R}\to\mathrm{L}}$ is the ratio between the amplitude B of the *left outgoing* wave and the amplitude G of the *right incoming* wave. The associated *transmission probability* $T_{\mathrm{R}\to\mathrm{L}}$ is the square of the magnitude of $t_{\mathrm{R}\to\mathrm{L}}$. Both can be calculated using the relation (2.138):

$$t_{\mathrm{R}\to\mathrm{L}} = \frac{B}{G} \overset{(2.138)}{=} S_{12} \quad \text{and} \quad T_{\mathrm{R}\to\mathrm{L}} = |t_{\mathrm{R}\to\mathrm{L}}|^2 = |S_{12}|^2 \tag{2.140}$$

The *reflection coefficient* from the right r_{R} is finally the ratio between the amplitude F of the *right outgoing* wave and the amplitude G of the *right incoming* wave. We can calculate it (and the associated reflection probability R_{R}) using equation (2.139).

$$r_{\mathrm{R}} = \frac{F}{G} \overset{(2.139)}{=} S_{22} \quad \text{and} \quad R_{\mathrm{R}} = |r_{\mathrm{R}}|^2 = |S_{22}|^2 \tag{2.141}$$

Thus, we can also write the scattering matrix S as follows:

$$S = \begin{pmatrix} r_{\mathrm{L}} & t_{\mathrm{R}\to\mathrm{L}} \\ t_{\mathrm{L}\to\mathrm{R}} & r_{\mathrm{R}} \end{pmatrix} \tag{2.142}$$

2.4.4 Probability Current Density in Scattering

Since $V(x)$ is purely real, there is no absorption or amplification within the potential. Therefore, the probability current density $j(x)$ must also be preserved during transmission through the potential barrier. For $j(x)$, it holds according to (2.32) in the one-dimensional, time-independent case:

$$j(x) = \mathrm{Re}\left(\psi^* \frac{\hbar}{im}\frac{d}{dx}\psi\right) \tag{2.143}$$

Let's assume a matter wave incoming from the left, which scatters at the barrier. It will be partially reflected back to the left and will partially run out to the right. There should be no wave incoming from the right. We can therefore set the amplitude $G = 0$ in our Ansatz (2.116). Additionally, we can express the amplitudes B and F in terms of the amplitude A of the incoming wave using the relations (2.134) and (2.135). We then obtain for $\psi_{\mathrm{I}}(x)$ and $\psi_{\mathrm{III}}(x)$:

$$\begin{aligned} \psi_{\mathrm{I}}(x) &= Ae^{+ikx} + Be^{-ikx} \overset{(2.134)}{\Longrightarrow} \quad \psi_{\mathrm{I}}(x) = Ae^{+ikx} + S_{11}Ae^{-ikx} \\ \psi_{\mathrm{III}}(x) &= Fe^{+ikx} \qquad\qquad\quad \overset{(2.135)}{\Longrightarrow} \quad \psi_{\mathrm{III}}(x) = S_{21}Ae^{+ikx} \end{aligned} \tag{2.144}$$

Without loss of generality, we now additionally assume that the incoming wave has an amplitude $A = 1$. This results in:

$$\psi_{\mathrm{I}}(x) = e^{+ikx} + S_{11}e^{-ikx} \quad \text{and} \quad \psi_{\mathrm{III}}(x) = S_{21}e^{+ikx} \tag{2.145}$$

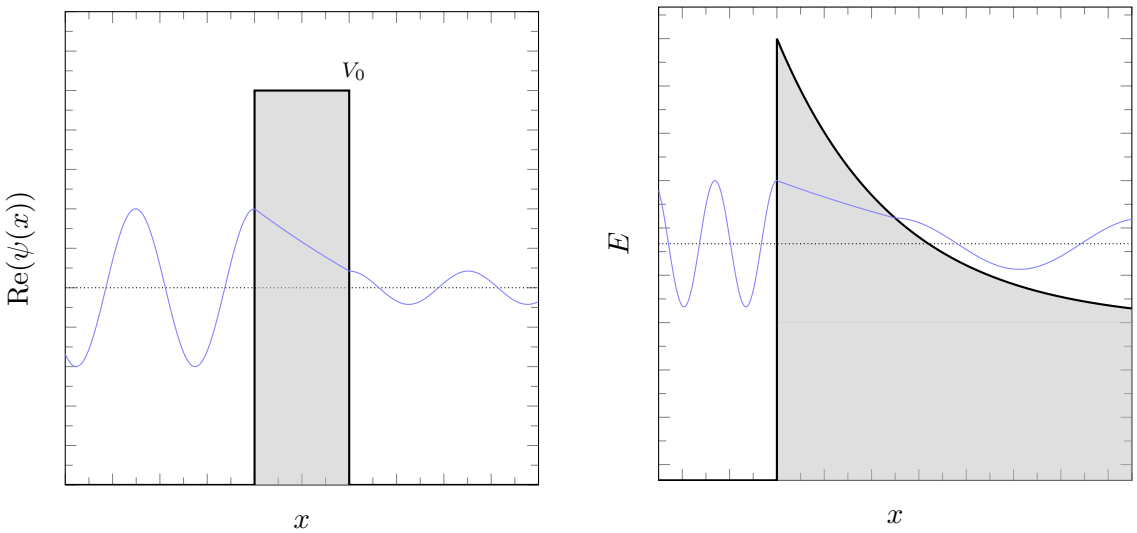


Fig. 10: (left) Transmission through a rectangular potential barrier. While the amplitude significantly decreases after transmission, there is no change in the particle's energy in this case. (right) Transmission of an α-particle, bound by the strong nuclear force, through the Coulomb barrier. Besides the decrease in amplitude, there is a reduction in the particle's kinetic energy.

The probability current density on the left side of the barrier yields:

$$\begin{aligned}
j_{\mathrm{I}}(x) &\overset{(2.143)}{=} \mathrm{Re}\left[\psi_{\mathrm{I}}^{*}\frac{\hbar}{im}\frac{d}{dx}\psi_{\mathrm{I}}\right] \overset{(2.145)}{=} \\
&= \mathrm{Re}\left[\left(e^{+ikx}+S_{11}e^{-ikx}\right)^{*}\frac{\hbar}{im}\frac{d}{dx}\left(e^{+ikx}+S_{11}e^{-ikx}\right)\right] = \\
&= \mathrm{Re}\left[\left(e^{-ikx}+S_{11}^{*}e^{+ikx}\right)\frac{\hbar}{im}\left(ike^{+ikx}-ikS_{11}e^{-ikx}\right)\right] = \\
&= \frac{\hbar k}{m}\mathrm{Re}\Big[1-\underbrace{S_{11}e^{-2ikx}}_{\alpha}+\underbrace{S_{11}^{*}e^{+2ikx}}_{\alpha^{*}}-\underbrace{S_{11}^{*}S_{11}}_{|S_{11}|^{2}}\Big] = \\
&= \frac{\hbar k}{m}\mathrm{Re}\left[1-(\alpha-\alpha^{*})-|S_{11}|^{2}\right] \quad \Big| \ \alpha-\alpha^{*}=2i\,\mathrm{Im}(\alpha) \\
&= \frac{\hbar k}{m}\mathrm{Re}\left[1-2i\,\mathrm{Im}\,(\alpha)-|S_{11}|^{2}\right] = \\
&= \frac{\hbar k}{m}\left(1-|S_{11}|^{2}\right) \overset{(2.137)}{=} \frac{\hbar k}{m}(1-R_{\mathrm{L}})
\end{aligned} \tag{2.146}$$

On the right side of the barrier, with analogous calculations, we find:

$$\begin{aligned}
j_{\mathrm{III}}(x) &\overset{(2.143)}{=} \mathrm{Re}\left[\psi_{\mathrm{III}}^{*}\frac{\hbar}{im}\frac{d}{dx}\psi_{\mathrm{III}}\right] \overset{(2.145)}{=} \\
&= \mathrm{Re}\left[\left(S_{21}e^{+ikx}\right)^{*}\frac{\hbar}{im}\frac{d}{dx}\left(S_{21}e^{+ikx}\right)\right] = \\
&= \mathrm{Re}\left[S_{21}^{*}e^{-ikx}\frac{\hbar}{im}ikS_{21}e^{+ikx}\right] = \frac{\hbar k}{m}\mathrm{Re}\Big[S_{21}^{*}S_{21}\,e^{-ikx}e^{+ikx}\Big] = \\
&= \frac{\hbar k}{m}|S_{21}|^{2} \overset{(2.136)}{=} \frac{\hbar k}{m}T_{\mathrm{L}\to\mathrm{R}}
\end{aligned} \tag{2.147}$$

Since the probability current density must also be conserved in a scattering process due to particle conservation, it holds that $j_{\mathrm{I}}(x) = j_{\mathrm{III}}(x)$, and thus:

$$j_{\mathrm{I}}(x) = j_{\mathrm{III}}(x) \overset{\substack{(2.146)\\(2.147)}}{\Longrightarrow} \frac{\hbar k}{m}(1-R_{\mathrm{L}}) = \frac{\hbar k}{m}T_{\mathrm{L}\to\mathrm{R}} \implies R_{\mathrm{L}}+T_{\mathrm{L}\to\mathrm{R}} = 1 \tag{2.148}$$

The calculation carried out above can of course also be done for a particle incident exclusively from the right, leading to an analogous result for R_{R} and $T_{\mathrm{R}\to\mathrm{L}}$:

$$1 = R_{\mathrm{R}}+T_{\mathrm{R}\to\mathrm{L}} \tag{2.149}$$

Due to the functional dependence of R_{L} and $T_{\mathrm{L}\to\mathrm{R}}$ on the components M_{11} and M_{21} of the transfer matrix according to (2.136) and (2.137), it is possible in the case of resonance that $T_{\mathrm{L}\to\mathrm{R}} = 1$ and $R_{\mathrm{L}} = 0$, meaning the particle is transmitted with a hundred percent probability and not reflected. This is the case if $M_{21} = 0$. Contrary to classical intuition, however, for a finite potential barrier, it must *always* be the case that $T \neq 0$, because M_{11} must not vanish, since in this case T would diverge according to (2.136). Therefore, there is always a certain transmission probability with a finite potential barrier. In classical physics, however, the case $R = 1$ is indeed possible at $E < V_0$. The argument presented can be equally applied to R_{R} and $T_{\mathrm{R}\to\mathrm{L}}$.

From the conservation of probability current density, it follows also that the scattering matrix is unitary:

$$S^{-1} = S^{\dagger} \iff SS^{\dagger} = \mathbb{1} \tag{2.150}$$

The proof of (2.150) follows through the conservation of probability current density on the left and right side of the barrier, starting from the Ansatz (2.116). By substituting (2.116) into (2.143), it follows:

$$\begin{aligned}
j_{\mathrm{I}}(x) &\overset{(2.143)}{=} \mathrm{Re}\left[\psi_{\mathrm{I}}^{*}\frac{\hbar}{im}\frac{d}{dx}\psi_{\mathrm{I}}\right] \overset{(2.116)}{=} \\
&= \mathrm{Re}\left[\left(Ae^{+ikx}+Be^{-ikx}\right)^{*}\frac{\hbar}{im}\frac{d}{dx}\left(Ae^{+ikx}+Be^{-ikx}\right)\right] = \\
&= \mathrm{Re}\left[\left(A^{*}e^{-ikx}+B^{*}e^{+ikx}\right)\frac{\hbar}{im}\left(ikAe^{+ikx}-ikBe^{-ikx}\right)\right] = \\
&= \frac{\hbar k}{m}\mathrm{Re}\left[\left(A^{*}e^{-ikx}+B^{*}e^{+ikx}\right)\left(Ae^{+ikx}-Be^{-ikx}\right)\right] = \\
&= \frac{\hbar k}{m}\mathrm{Re}\Big[A^{*}A-B^{*}B+\underbrace{AB^{*}e^{+2ikx}}_{\alpha}-\underbrace{A^{*}Be^{-2ikx}}_{\alpha^{*}}\Big] = \\
&= \frac{\hbar k}{m}\mathrm{Re}\left[|A|^{2}-|B|^{2}+(\alpha-\alpha^{*})\right] = \quad \Big|\ \alpha-\alpha^{*}=2i\,\mathrm{Im}(\alpha) \\
&= \frac{\hbar k}{m}\mathrm{Re}\left[|A|^{2}-|B|^{2}+2i\,\mathrm{Im}\,(\alpha)\right] = \\
&= \frac{\hbar k}{m}\left(|A|^{2}-|B|^{2}\right)
\end{aligned} \tag{2.151}$$

$$\begin{aligned}
j_{\mathrm{III}}(x) &\overset{(2.143)}{=} \mathrm{Re}\left[\psi_{\mathrm{III}}^{*}\frac{\hbar}{im}\frac{d}{dx}\psi_{\mathrm{III}}\right] \overset{(2.116)}{=} \\
&= \mathrm{Re}\left[\left(Fe^{+ikx}+Ge^{-ikx}\right)^{*}\frac{\hbar}{im}\frac{d}{dx}\left(Fe^{+ikx}+Ge^{-ikx}\right)\right] = \\
&= \mathrm{Re}\left[\left(F^{*}e^{-ikx}+G^{*}e^{+ikx}\right)\frac{\hbar}{im}\left(ikFe^{+ikx}-ikGe^{-ikx}\right)\right] = \\
&= \frac{\hbar k}{m}\mathrm{Re}\left[\left(F^{*}e^{-ikx}+G^{*}e^{+ikx}\right)\left(Fe^{+ikx}-Ge^{-ikx}\right)\right] = \\
&= \frac{\hbar k}{m}\mathrm{Re}\Big[F^{*}F-G^{*}G+\underbrace{FG^{*}e^{+2ikx}}_{\alpha}-\underbrace{F^{*}Ge^{-2ikx}}_{\alpha^{*}}\Big] = \\
&= \frac{\hbar k}{m}\mathrm{Re}\left[|F|^{2}-|G|^{2}+(\alpha-\alpha^{*})\right] = \quad \Big|\ \alpha-\alpha^{*}=2i\,\mathrm{Im}(\alpha) \\
&= \frac{\hbar k}{m}\mathrm{Re}\left[|F|^{2}-|G|^{2}+2i\,\mathrm{Im}\,(\alpha)\right] = \\
&= \frac{\hbar k}{m}\left(|F|^{2}-|G|^{2}\right)
\end{aligned} \tag{2.152}$$

It must hold that $j_{\mathrm{I}}(x)=j_{\mathrm{III}}(x)$. Therefore, we find:

$$\begin{aligned}
j_{\mathrm{I}}(x) &= j_{\mathrm{III}}(x) \overset{(2.151)(2.152)}{\Longrightarrow} \\
\frac{\hbar k}{m}\left(|A|^{2}-|B|^{2}\right) &= \frac{\hbar k}{m}\left(|F|^{2}-|G|^{2}\right) \Longrightarrow \\
|B|^{2}+|F|^{2} &= |A|^{2}+|G|^{2}
\end{aligned} \tag{2.153}$$

The amplitudes A and G correspond to the waves incoming on the left and right sides of the barrier, and the amplitudes B and F correspond to the waves outgoing on the left and right sides. We can therefore define the amplitude vectors $\mathbf{a}_{\mathrm{in}}$ and $\mathbf{a}_{\mathrm{out}}$:

$$\mathbf{a}_{\mathrm{in}}=\begin{pmatrix}A\\G\end{pmatrix} \quad \text{and} \quad \mathbf{a}_{\mathrm{out}}=\begin{pmatrix}B\\F\end{pmatrix} \tag{2.154}$$

The left side of Equation (2.153) can thus be expressed as $\mathbf{a}_{\text{out}}^\dagger \mathbf{a}_{\text{out}}$, and the right side as $\mathbf{a}_{\text{in}}^\dagger \mathbf{a}_{\text{in}}$. Additionally, we can write the scattering matrix equation (2.124) compactly as follows:

$$\mathbf{a}_{\text{out}} = S\,\mathbf{a}_{\text{in}} \tag{2.155}$$

By expressing Equation (2.153) with the help of the amplitude vectors (2.155), and then using (2.155):

$$\begin{aligned} \mathbf{a}_{\text{out}}^\dagger \mathbf{a}_{\text{out}} &= \mathbf{a}_{\text{in}}^\dagger \mathbf{a}_{\text{in}} \quad \overset{(2.155)}{\Longrightarrow} \\ (S\mathbf{a}_{\text{in}})^\dagger\, S\mathbf{a}_{\text{in}} &= \mathbf{a}_{\text{in}}^\dagger \mathbf{a}_{\text{in}} \quad \Longrightarrow \\ \mathbf{a}_{\text{in}}^\dagger \underbrace{S^\dagger S}_{\mathbb{1}} \mathbf{a}_{\text{in}} &= \mathbf{a}_{\text{in}}^\dagger \mathbf{a}_{\text{in}} \end{aligned} \tag{2.156}$$

In order for Equation (2.156) to be satisfied, it must hold that:

$$S^\dagger S = \mathbb{1} \quad \square$$

Thus, we have proven the unitarity of the scattering matrix as postulated in (2.150).

2.4.5 Tunneling Effect

We have already shown in (2.123) that we can represent the transmission matrix M of a potential barrier as a product of the matrices P and Q, which we have already determined in (2.121) and (2.122). We now want to calculate the transmission probability through a barrier of length $\Delta x = L$. As already discussed, the penetration of the barrier (classically impossible) in the case $E < V_0$, as shown in Figure 12, is *not* forbidden. We call this phenomenon the *tunneling effect*. In the case of $E > V_0$, transmission is also classically permitted, resulting in a classical $T = 1$. In quantum physics, however, reflection $R > 0$ also occurs (except in resonance cases), resulting in transmission below one (Fabry-Pérot interferences).

Example: Transmission Probability through Barrier

The transmission probability of a particle incident from the left through a potential barrier is given by (2.136):

$$T_{\text{L}\to\text{R}} = \frac{1}{|M_{11}|^2}$$

The potential barrier has height V_0 and the particle has energy $E < V_0$. Since we are only interested in the entry M_{11} of the transfer matrix $M = PQ$ to be calculated according to

(2.123), we need not calculate the entire matrix M, but only the entry M_{11}:

$$\begin{aligned}
M_{11} &= P_{11}Q_{11} + P_{12}Q_{21} \overset{(2.121)(2.122)}{=} \\
&= \frac{k+i\kappa}{2k}\,\frac{\kappa-ik}{2\kappa}\,e^{L(ik+\kappa)} + \frac{k-i\kappa}{2k}\,\frac{\kappa+ik}{2\kappa}\,e^{L(ik-\kappa)} = \\
&= \frac{1}{4}\left[\left(1+i\frac{\kappa}{k}\right)\left(1-i\frac{k}{\kappa}\right)e^{L(ik+\kappa)} + \left(1-i\frac{\kappa}{k}\right)\left(1+i\frac{k}{\kappa}\right)e^{L(ik-\kappa)}\right] = \\
&= \frac{1}{4}\left[\left(1+i\frac{\kappa}{k}-i\frac{k}{\kappa}+1\right)e^{\kappa L} + \left(1-i\frac{\kappa}{k}+i\frac{k}{\kappa}+1\right)e^{-\kappa L}\right]e^{ikL} = \\
&= \frac{1}{4}\left[2\left(e^{\kappa L}+e^{-\kappa L}\right) + i\left(\frac{\kappa}{k}-\frac{k}{\kappa}\right)e^{\kappa L} - i\left(\frac{\kappa}{k}-\frac{k}{\kappa}\right)e^{-\kappa L}\right]e^{ikL} = \\
&= \frac{1}{4}\left[2\left(e^{\kappa L}+e^{-\kappa L}\right) + i\left(\frac{\kappa}{k}-\frac{k}{\kappa}\right)\left(e^{\kappa L}-e^{-\kappa L}\right)\right]e^{ikL} = \quad \Big|\ \varepsilon = \frac{\kappa}{k}-\frac{k}{\kappa} \\
&= \left[\frac{1}{2}\left(e^{\kappa L}+e^{-\kappa L}\right) + \frac{i\varepsilon}{4}\left(e^{+\kappa L}-e^{-\kappa L}\right)\right]e^{ikL} = \\
&= \left[\cosh(\kappa L) + \frac{i\varepsilon}{2}\sinh(\kappa L)\right]e^{ikL} = \frac{2\cosh(\kappa L)+i\varepsilon\sinh(\kappa L)}{2}e^{ikL}
\end{aligned} \tag{2.157}$$

The quantity $\varepsilon = \frac{\kappa}{k} - \frac{k}{\kappa}$ has been defined. By the magnitude square of $|M_{11}|^2$, one obtains the transmission probability $T_{\mathrm{L}\to\mathrm{R}}$ from (2.136):

$$\begin{aligned}
T_{\mathrm{L}\to\mathrm{R}} &= \frac{1}{|M_{11}|^2} \overset{(2.157)}{=} \\
&= \frac{4}{4\cosh^2(\kappa L) + \varepsilon^2\sinh^2(\kappa L)} \quad \Big|\ a\cosh^2(x) + b\sinh^2(x) = a + (b+a)\sinh^2(x) \\
&= \frac{4}{4 + (\varepsilon^2+4)\sinh^2(\kappa L)}
\end{aligned} \tag{2.158}$$

k and κ are equivalent to the expressions found in (2.98) and (2.99), but now κ should be represented in terms of the ratio of the energy of the incoming particle to the potential $x = E/V_0$:

$$k = \sqrt{\frac{2mE}{\hbar^2}} \quad \text{and} \quad \kappa = \sqrt{\frac{2m(V_0-E)}{\hbar^2}} = \sqrt{\frac{2mV_0(1-x)}{\hbar^2}} = \alpha\sqrt{1-x}$$

In α, all constants are summarized. ε can thus be written as a function of x:

$$\varepsilon = \frac{\kappa}{k} - \frac{k}{\kappa} = \sqrt{\frac{V_0(1-x)}{E}} - \sqrt{\frac{E}{V_0(1-x)}} = \sqrt{\frac{1-x}{x}} - \sqrt{\frac{x}{1-x}} = \frac{1-2x}{\sqrt{x(1-x)}}$$

For $E < V_0$, the already-derived formula for T still applies; for $E > V_0$, however, it must be noted that the root in the hyperbolic function becomes imaginary. Utilizing the relation $\sinh(ix) = -i\sin(x)$, one obtains a piecewise transmission probability as a function of the ratio $x = E/V_0$ in the following form:

$$T(x) = \begin{cases} \dfrac{4}{4+\left(\frac{(1-2x)^2}{x(1-x)}+4\right)\sinh^2(\alpha\sqrt{1-x})}, & x < 1 \\ \dfrac{4}{4-\left(\frac{(1-2x)^2}{x(1-x)}+4\right)\sin^2(\alpha\sqrt{x-1})}, & x > 1 \end{cases}$$

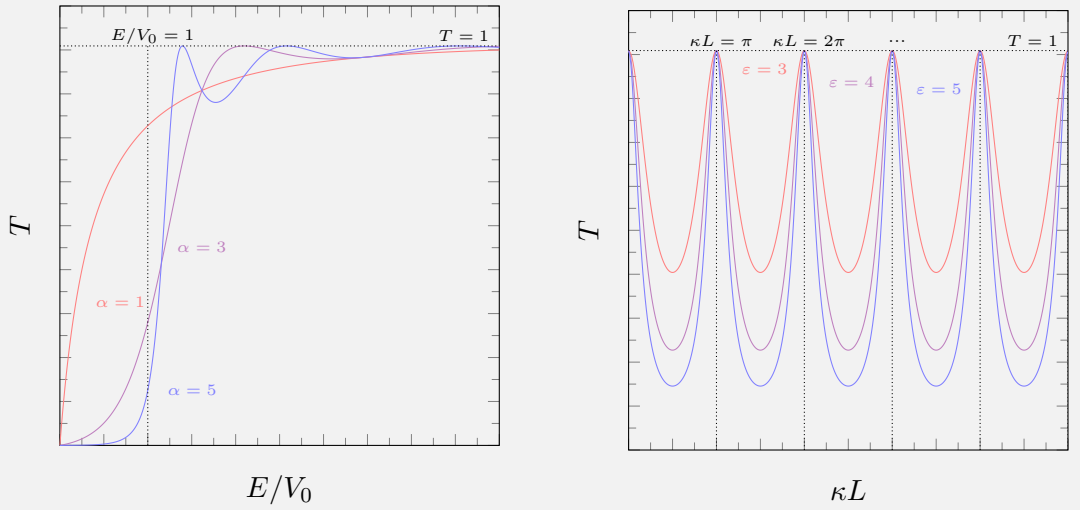


Fig. 12: (left) Representation of the transmission probability for different energy ratios E/V_0. (right) At constant potential characteristics V_0 and L, the $n\pi$-periodic structure of the resonances is apparent.

3 Formal Structure of Quantum Theory

Motivation: Dirac Notation and Hilbert Space

In the preceding chapters, we have shown that the state of a quantum system at a given time can be described with a normalized wave function $\psi(\mathbf{r})$. If we consider the space of all possible states a quantum system can take, then this is a space of wave functions $\psi_i(\mathbf{r})$, for which (due to normalization) the following holds: $\int d^3\mathbf{r}|\psi_N(\mathbf{r})|^2 = 1$ (the N stands for "normalized"). To achieve this, the (non-normalized) wave functions are provided with a normalization factor N, so that $\psi_N = N\psi$.

For a wave function to be normalizable at all, it must be square-integrable. Therefore, wave functions belong to the space L^2 of square-integrable functions in $\mathbb{R}^3$, where for all functions the following holds: $\int d^3\mathbf{r}|f(\mathbf{r})|^2 < \infty$. It can be shown that the L^2 – with certain additional conditions – forms a Hilbert space $\mathcal{H}$.

In the following, we will describe the properties of this Hilbert space, and show how one can use, among other things, the so-called *Dirac notation* to represent and compute quantum states very elegantly and compactly using the properties of the Hilbert space.

3.1 Dirac Notation

The origin of the Dirac notation introduced by Paul Dirac becomes clear when one imagines how the scalar product of two wave functions $\phi(\mathbf{r})$ and $\psi(\mathbf{r})$ can be written:

$$\langle\phi|\psi\rangle = \int d^3\mathbf{r}\ \phi^*(\mathbf{r})\psi(\mathbf{r}) \tag{3.1}$$

In the Dirac notation, one also uses the two halves of $\langle\phi|\psi\rangle$ outside of the scalar product: The quantum state corresponding to the wave function $\psi(\mathbf{r})$ is written as $|\psi\rangle$ and is called a **Ket vector**. The left part of $\langle\phi|\psi\rangle$ can also be written as $\langle\phi|$. This is then referred to as a **Bra vector**. Bra and Ket together form a complete bracket, English "**Bra**-c-**ket**". Therefore, the Dirac notation is also called "Braket notation". The reason one speaks of vectors – as we will see – is because all elements of a Hilbert space can be understood as vectors.

In this notation, the following must be considered: The wave function $\psi(\mathbf{r})$ explicitly depends on the location $\mathbf{r}$. However, one could also *represent exactly the same* quantum state, for instance, using a wave function $\tilde{\psi}(\mathbf{p})$, which depends on the momentum $\mathbf{p}$. The wave function then appears completely different. The necessary abstraction is provided by Dirac notation: The quantum state is written as $|\psi\rangle$, and this notation is independent of how the wave function is concretely represented in a further course.

3.1.1 Hilbert Space and Ket Vectors

The physical quantum state of a system is thus described by the state vector $|\psi\rangle$, which is interpreted as an element of the Hilbert space $\mathcal{H}$:

$$|\psi\rangle \in \mathcal{H} \tag{3.2}$$

Since a Hilbert space is a vector space, all properties of a vector space apply to a Ket state, such as associativity, commutativity, distributivity, and the existence of a neutral as well as an inverse element. The definition of a vector space essentially means that there must be vectors and scalars, that we can add the vectors "in a meaningful way", and that we can multiply scalars and vectors "in a meaningful way". We can now ascertain that the above-mentioned properties also apply to Ket state vectors and complex numbers in the Hilbert space $\mathcal{H}$.

In-Depth: Definition of a Vector Space

A vector space over a field $\mathbb{K}$ is the set $\mathcal{V}$, for which, together with addition $\mathcal{V} \times \mathcal{V} \to \mathcal{V}$ and scalar multiplication $\mathcal{V} \times \mathbb{K} \to \mathcal{V}$, the following properties must apply ($u, v, w \in \mathcal{V}$ and $\alpha, \beta \in \mathbb{K}$):

- $u + (v + w) = (u + v) + w$ (Associative law for vector addition)
- $u + v = v + u$ (Commutative law for vector addition)
- $u + \mathbf{0} = \mathbf{0} + u = u$ ($\exists$ neutral vector element $\mathbf{0}$)
- $u + (-u) = (-u) + u = 0$ ($\forall u : \exists$ inverse vector element $-u$)
- $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$ (1st Distributive law for scalar product)
- $(\alpha + \beta) \cdot u = \alpha \cdot u + \beta \cdot u$ (2nd Distributive law for scalar product)
- $(\alpha \cdot \beta) \cdot u = \alpha \cdot (\beta \cdot u)$ (Associative law for scalar product)
- $\mathbb{1} \cdot u = u$ (the neutral scalar $\mathbb{1}$ is "sensibly" processed by the scalar product)

For the Ket vectors $|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle \in \mathcal{H}$ and scalars $\alpha, \beta \in \mathbb{C}$, the following holds:

- $|\psi_1\rangle + (|\psi_2\rangle + |\psi_3\rangle) = (|\psi_1\rangle + (|\psi_2\rangle)) + |\psi_3\rangle$ (Associative law for vector addition)
- $|\psi_1\rangle + |\psi_2\rangle = |\psi_2\rangle + |\psi_1\rangle$ (Commutative law for vector addition)
- $|\psi_1\rangle + |0\rangle = |0\rangle + |\psi_1\rangle = |\psi_1\rangle$ ($\exists$ neutral state vector $|0\rangle$)
- $|\psi_1\rangle + (-|\psi_1\rangle) = (-|\psi_1\rangle) + |\psi_1\rangle = |0\rangle$ ($\forall |\psi\rangle : \exists$ inverse state vector $-|\psi\rangle$)
- $\alpha (|\psi_1\rangle + |\psi_2\rangle) = \alpha |\psi_1\rangle + \alpha |\psi_2\rangle$ (1st Distributive law for scalar product)
- $(\alpha + \beta) \cdot |\psi_1\rangle = \alpha |\psi_1\rangle + \beta |\psi_1\rangle$ (2nd Distributive law for scalar product)
- $(\alpha\beta) |\psi_1\rangle = \alpha (\beta |\psi_1\rangle)$ (Associative law for scalar product)
- $1 \cdot |\psi_1\rangle = |\psi_1\rangle$ (The multiplication of 1 with a vector leaves it unchanged)

That all these points are fulfilled becomes intuitively clear when you mentally replace the (abstract) Ket vectors with concrete wave functions: Of course, you can add multiple functions, where the order does not matter. Similarly, a function can be multiplied by a (complex) number, etc. And the "neutral state vector" $|0\rangle$ simply corresponds to the zero function, which returns zero everywhere in space, where we must deviate from the requirement of normalizability as an exception; but we are still in the L^2-space. Physically, $|0\rangle$ corresponds to a state "without particles". Therefore, it is occasionally acceptable that the probability of finding a particle somewhere in space is equal to zero.

In-Depth: Wave Functions as Vectors in Hilbert Space

We have thus just shown that Ket vectors and all wave functions that represent such Ket vectors, together with the complex numbers, satisfy the definition of a vector space. Ket vectors and wave functions therefore represent vectors in Hilbert space. How can one conceptualize this? What does it "really" (intuitively) mean for a Ket vector to be orthogonal to another Ket vector, especially when these Ket vectors represent wave functions?

To understand this, it helps to visualize the analogy between Ket vectors and simple vectors, for instance, in $\mathbb{R}^3$. You can think of a simple vector $\mathbf{v} \in \mathbb{R}^3$ as an arrow with

direction and length that exists independent of the chosen coordinate system: $\mathbf{v}$ thus represents the "vector in itself", without reference to any coordinate system. If the vector $\mathbf{v}$ is instead written as a triplet of numbers $\mathbf{v} = (v_1, v_2, v_3)^\intercal$, then one has decided on a particular coordinate system. You can represent the same vector in another coordinate system with a different triplet of numbers $\mathbf{v} = (\tilde{v}_1, \tilde{v}_2, \tilde{v}_3)^\intercal$. The Ket vector $|\psi\rangle$ in Hilbert space now corresponds to the vector $\mathbf{v} \in \mathbb{R}^3$. It represents the vector in Hilbert space "in itself", without committing to a coordinate system. A wave function $\psi(x)$ belonging to the Ket vector $|\psi\rangle$ corresponds to the representation of the Ket vector $|\psi\rangle$ in a particular coordinate system. We can also represent the same vector $|\psi\rangle$ in a different coordinate system, for instance with the wave function $\tilde{\psi}(p)$.

But how can a function $\psi(x)$ in the Hilbert space $\mathcal{H}$ be analogous to a triplet of numbers $(v_1, v_2, v_3)^\intercal$ in $\mathbb{R}^3$? To understand this, we first consider that we could replace the triplet of numbers with the three numbers v_1, v_2, and v_3 with a function $v(i)$, which would allow us to write the vector (in this coordinate system) as $(v(1), v(2), v(3))^\intercal$. For $v(i)$, i can only take on the values 1, 2, or 3. However, since we want to go beyond $\mathbb{R}^3$, we next assume an infinitely-dimensional space $\mathbb{R}^\infty$. By this point, it becomes clear that we can no longer represent a vector by simply writing triplets of numbers $(v_1, v_2, v_3, v_4, v_5, \dots)^\intercal$ with infinitely many entries. There is no sensible alternative but to represent the infinitely many values v_i by the function $v(i)$. In the final step of generalization, we assume that our vector space becomes so vast that not even the countably infinite number of values $v(i)$ with $i = 1, 2, 3, 4, 5, \dots$ is sufficient to represent the vector. There exists an infinite continuum of values between $v(1)$ and $v(2)$, for instance $v(1.5)$, $v(1.3333)$, $v(\sqrt{2})$, etc. Our vector is finally represented definitively by a function $v(i)$ with $i \in \mathbb{R}$!

The vector space was enlarged so that we can no longer manage with the countable values v_1, v_2, v_3, etc. Instead of $\mathbf{v}$, we now write $|\psi\rangle$ (to illustrate that we are located in a space of significantly higher dimensions – the Hilbert space $\mathcal{H}$), and we replace $v_i \equiv v(i)$ or $\tilde{v}_i \equiv \tilde{v}(i)$ with (for example) $\psi(x)$ or $\tilde{\psi}(p)$. With this visualization, it becomes easier to understand how two vectors in Hilbert space can be orthogonal, and how this can be calculated. We first visualize two vectors $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^3$, and then the inner product follows:

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{3} a_i b_i = \sum_{i=1}^{3} \tilde{a}_i \tilde{b}_i \tag{3.3}$$

If this value is equal to zero, then $\mathbf{a}$ and $\mathbf{b}$ are orthogonal. The value of the inner product is independent of the coordinate system used. If we denote the distance between the possible index values i as Δi, then $\Delta i = 1$, and for $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^3$ or in $\mathbb{R}^\infty$, the following holds:

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{3} \Delta i \; a_i b_i = \sum_{i=1}^{3} \Delta i \; \tilde{a}_i \tilde{b}_i \quad \text{and} \quad \mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{\infty} \Delta i \; a(i) b(i) = \sum_{i=1}^{\infty} \Delta i \; \tilde{a}(i) \tilde{b}(i) \tag{3.4}$$

In the case of $\mathbb{R}^\infty$, we no longer write the values a_i and b_i, rather express these solely with the functions $a(i)$ and $b(i)$. If we shift from $\mathbb{R}^\infty$ to $\mathbb{C}^\infty$, it must additionally be considered that one of the values must be conjugated complexly:

$$(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{\infty} \Delta i \; a^*(i) b(i) = \sum_{i=1}^{\infty} \Delta i \; \tilde{a}^*(i) \tilde{b}(i) \tag{3.5}$$

In the Hilbert space of square-integrable functions, where there are no countable indices

anymore, the inner product between the states $|a\rangle$ and $|b\rangle$ is finally calculated as follows:

$$\langle a|b\rangle = \int_{-\infty}^{\infty} dx\, a^*(x)b(x) = \int_{-\infty}^{\infty} dp\, \tilde{a}^*(p)\tilde{b}(p) \tag{3.6}$$

The "index" is now, for instance, x or p instead of i, the "index boundaries" now run continuously over the entire space; ultimately the sum becomes an integral.

3.1.2 Dual Space and Bra Vectors

We have described in the introduction that in the Dirac notation, the scalar product $\langle\phi|\psi\rangle$ is somewhat dissected and that we not only give the right part $|\psi\rangle$ as a "Ket vector" an independent identity but also the left part $\langle\phi|$ as a "Bra vector". For fundamental understanding, it is usually enough to simply imagine the Bra vector as the "left part of the scalar product". However, as we are indeed considering Bra vectors as independent objects, it makes sense to reflect on the space to which Bra vectors belong and what properties they have.
So, what is a Bra vector? The answer lies in the scalar product from which we originally "extracted" the Bra vector: Whenever we connect a Bra vector $\langle\phi|$ with a Ket vector $|\psi\rangle$, we obtain a scalar product $\langle\phi|\psi\rangle$, which gives us a (complex) number. You could say that the Bra vector is a mathematical object that takes a Ket vector as input and outputs a number. Such an object is called a *functional*. Specifically, a Bra vector is a *linear* functional because the following holds:

$$\begin{aligned}\langle\phi|\left(\alpha|\psi_1\rangle + \beta|\psi_2\rangle\right) &= \int d\mathbf{r}\,\phi^*(\mathbf{r})\left(\alpha\psi_1(\mathbf{r}) + \beta\psi_2(\mathbf{r})\right) = \\ &= \alpha\int d\mathbf{r}\,\phi^*(\mathbf{r})\psi_1(\mathbf{r}) + \beta\int d\mathbf{r}\,\phi^*(\mathbf{r})\psi_2(\mathbf{r}) = \\ &= \alpha\langle\phi|\psi_1\rangle + \beta\langle\phi|\psi_2\rangle\end{aligned} \tag{3.7}$$

Therefore, the space of all Bra vectors is the *space of all linear functionals* of the Hilbert space! This space is called the *dual space*. It can be shown that all properties of the vector space from the previous section are also satisfied for dual vectors, which is why the dual space also represents a vector space and thus justifies the label as Bra *vectors*. Through transposition and complex conjugation (briefly: "daggering"), a Ket vector can be rewritten into a Bra vector, and vice versa:

$$|\psi\rangle^\dagger = \langle\psi| \quad \text{and} \quad \langle\psi|^\dagger = |\psi\rangle \tag{3.8}$$

The usage of complex conjugation is easily understood due to the way in which the scalar product is calculated. The significance of the added transposition will only become clear later on.

3.1.3 The Scalar Product

In-Depth: Definition of Hilbert Space

A Hilbert space $\mathcal{H}$ is a (linear) vector space with a scalar product $\langle\psi|\psi\rangle$, which induces a corresponding norm $||\psi|| = \sqrt{\langle\psi|\psi\rangle}$. Additionally, the space must be complete, which means that for any convergent Cauchy sequence $\psi_i \in \mathcal{H}$, it holds that the limit $\lim_{i\to\infty}\psi_i$ is again in $\mathcal{H}$.

Let $|\psi\rangle$ now be an element of the Hilbert space $\mathcal{H}$, then for the scalar product $\langle\psi|\psi\rangle \geq 0$. Thus, the scalar product is positive semidefinite; equality $\langle\psi|\psi\rangle = 0$ is only fulfilled if $|\psi\rangle = 0$.

According to definition, we can only speak of a Hilbert space if it is a vector space equipped with a scalar product for the vectors, and we can define a norm via the scalar product. This should now be the case: A scalar product of the form $\langle\psi_1|\psi_2\rangle$ is already available as a starting point for our considerations.

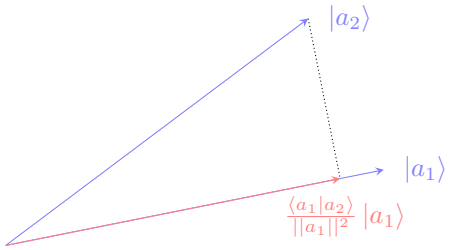


Fig. 13: Schematic representation of the scalar product as the projection of a vector $|a_2\rangle$ onto $|a_1\rangle$ in the two-dimensional space $\mathbb{R}^2$.

This scalar product $\langle\psi_1|\psi_2\rangle$ can be interpreted, quite analogous to the tangible position vectors in $\mathbb{R}^3$, as the projection of the state $|\psi_2\rangle$ onto $|\psi_1\rangle$. Just as with position vectors in $\mathbb{R}^3$, we can define with the abstract vectors in Hilbert space: If $|\psi_1\rangle$ and $|\psi_2\rangle$ are orthogonal with respect to each other, then the following holds:

$$\langle\psi_1|\psi_2\rangle = 0 \tag{3.9}$$

In this case, $|\psi_1\rangle$ contains no component parallel to $|\psi_2\rangle$. With the help of $\langle\psi|\psi\rangle$, one can now define the norm $||\psi||$ of a state vector $|\psi\rangle$:

$$||\psi|| = \sqrt{\langle\psi|\psi\rangle} \geq 0 \tag{3.10}$$

That the expression (3.10) always yields a real number greater than or equal to zero is a necessary condition for it to be a norm. Another requirement is also fulfilled: Expression (3.10) produces a value of zero if and only if $|\psi\rangle$ equals the zero vector $|0\rangle$:

$$\langle\psi|\psi\rangle = 0 \iff |\psi\rangle = |0\rangle \tag{3.11}$$

Another property of the scalar product is that it is linear in the second argument (this is according to our definition, in some fields of mathematics and physics, linearity in the first argument is also required):

$$\langle\psi_1|\alpha\psi_2 + \beta\psi_3\rangle = \langle\psi_1|\alpha\psi_2\rangle + \langle\psi_1|\beta\psi_3\rangle = \alpha\,\langle\psi_1|\psi_2\rangle + \beta\,\langle\psi_1|\psi_3\rangle \tag{3.12}$$

For the first argument, the scalar product is semi-linear, meaning the following:

$$\langle\alpha\psi_1 + \beta\psi_2|\psi_3\rangle = \langle\alpha\psi_1|\psi_3\rangle + \langle\beta\psi_2|\psi_3\rangle = \alpha^*\,\langle\psi_1|\psi_3\rangle + \beta^*\,\langle\psi_2|\psi_3\rangle \tag{3.13}$$

(3.12) and (3.13) together indicate that the scalar product is a so-called *sesquilinear form* with the following property:

$$\langle\psi_1|\psi_2\rangle^* = \langle\psi_2|\psi_1\rangle \tag{3.14}$$

Scalar quantities can therefore always be pulled out of the Ket-bracket:

$$|\alpha\psi\rangle = \alpha\,|\psi\rangle \tag{3.15}$$

For Bra vectors, we must note that scalars standing within the Bra-bracket are complex conjugated and transposed (briefly: "daggered"). For a scalar α, $\alpha^\dagger = \alpha^*$; for Bra vectors, therefore:

$$\langle\psi\alpha| = \langle\psi|\,\alpha^* \tag{3.16}$$

Furthermore, for a Ket vector due to the general rule for (complex-conjugated) transposition $(XY)^\dagger = Y^\dagger X^\dagger$:

$$(\alpha \left|\psi\right\rangle)^\dagger = \left|\psi\right\rangle^\dagger \alpha^\dagger = \left\langle\psi\right| \alpha^* \tag{3.17}$$

For the scalar product, we can state the *Cauchy-Schwarz inequality*:

$$|\left\langle\psi_1|\psi_2\right\rangle| \leq ||\psi_1|| \cdot ||\psi_2|| \tag{3.18}$$

With (3.18), the *triangle inequality* for arbitrary state vectors $|\psi_1\rangle$ and $|\psi_2\rangle$ can also be motivated:

$$||\psi_1 + \psi_2|| \leq ||\psi_1|| + ||\psi_2|| \tag{3.19}$$

Example: Triangle Inequality

From the Cauchy-Schwarz inequality, the validity of the triangle inequality can be easily demonstrated. The norm of the sum of two state vectors $|\psi\rangle$ and $|\phi\rangle$ is squared to eliminate roots, resulting in:

$$\begin{aligned} ||\phi + \psi||^2 &= ||\phi||^2 + ||\psi||^2 + \langle\phi|\psi\rangle + \langle\psi|\phi\rangle \overset{(3.18)}{\leq} \\ &\leq ||\phi||^2 + ||\psi||^2 + ||\phi|| \cdot ||\psi|| + ||\psi|| \cdot ||\phi|| = \\ &= ||\phi||^2 + ||\psi||^2 + 2||\phi|| \cdot ||\psi|| = \\ &= (||\phi|| + ||\psi||)^2 \end{aligned}$$

Indeed, this inequality follows from (3.18), allowing for an estimation of the magnitude of $\langle\phi|\psi\rangle$. Subtracting the square root from both sides exactly yields the triangle inequality (3.19):

$$||\phi + \psi|| \leq ||\phi|| + ||\psi||$$

3.1.4 Complete Orthonormal System and Dimension of the Hilbert Space

Just like with the tangible position vectors in $\mathbb{R}^3$, one can also represent every state vector $|\psi\rangle$ in an N-dimensional Hilbert space $\mathcal{H}$ as a linear combination of orthonormal basis states $\{|\phi_1\rangle, |\phi_2\rangle, \ldots |\phi_N\rangle\}$:

$$|\psi\rangle = \sum_i c_i |\phi_i\rangle \tag{3.20}$$

Unlike in $\mathbb{R}^3$, however, the expansion coefficients c_i are complex-valued: $c_i \in \mathbb{C}$. When all basis vectors $|\phi_i\rangle$ are orthogonal to each other and normalized to a length of 1, so that $\langle\phi_i|\phi_j\rangle = \delta_{ij}$ holds, one speaks of an *orthonormal system* (ONS). The dimension of a Hilbert space is given by the maximum possible number of linearly independent basis vectors. If an orthonormal basis in an N-dimensional Hilbert space $\mathcal{H}_N$ contains exactly N orthonormal vectors, then this basis is termed a *complete orthonormal system* (CONS). With a CONS as basis, any arbitrary vector in $\mathcal{H}_N$ can be represented with N expansion coefficients $c_1, \ldots, c_N$.
Depending on the type of quantum system being considered, the dimension of the underlying Hilbert space can be finite or even (countably) infinite. For many phenomena, it is even necessary to extend the Hilbert space to an uncountably infinite-dimensional space, the so-called *extended Hilbert space*. However, for now, we will focus on finite-dimensional Hilbert spaces.
Using the decomposition in basis vectors shown in (3.20), an abstract Ket vector $|\psi\rangle$ can be represented as a column vector in $\mathbb{C}^N$ that contains the expansion coefficients c_i:

$$|\psi\rangle = c_1 |\phi_1\rangle + c_2 |\phi_2\rangle + \cdots + c_N |\phi_N\rangle \xrightarrow{\{|\phi_i\rangle\}} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} \tag{3.21}$$

The corresponding Bra vector $\langle\psi| = |\psi\rangle^\dagger$ is then represented by the complex conjugated, transposed column vector of (3.21) – thus a row vector:

$$\langle\psi| = c_1^* \langle\phi_1| + c_2^* \langle\phi_2| + \cdots + c_N^* \langle\phi_N| \xrightarrow{\{\langle\phi_i|\}} \begin{pmatrix} c_1^* & c_2^* & \dots & c_N^* \end{pmatrix} \tag{3.22}$$

It should be noted that the numbers in the vectors (3.21) and (3.22) depend on the chosen basis. Therefore, one should use an arrow instead of an equals sign, which also specifies the underlying basis (in our case $\{\phi_i\}$). The scalar product $\langle\psi|\psi\rangle$ can be expressed as the multiplication of row and column vectors, and is *independent* of the chosen basis, which is why an equals sign can be used again:

$$\langle\psi|\psi\rangle = \begin{pmatrix} c_1^* & c_2^* & \dots & c_N^* \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} = c_1^* c_1 + c_2^* c_2 + \cdots + c_N^* c_N \tag{3.23}$$

3.1.5 Operators in the Braket Notation

There is also a compact notation for operators in the Dirac formalism. We have already learned operators as mathematical objects that "receive as input a wave function" (or, as typically phrased: "act on a wave function") and produce another function as a result. When we transfer this to the more abstract Dirac notation with vectors in Hilbert space, an operator $\hat{A}$ is a (in our case linear) mapping from the Hilbert space $\mathcal{H}_N$ into the same Hilbert space $\mathcal{H}_N$ ($\hat{A} : \mathcal{H}_N \to \mathcal{H}_N$). An operator $\hat{A}$ always acts only from the left on a Ket state $|\psi\rangle$; the action of the operator $\hat{A}$ from the right on the Ket vector $|\psi\rangle$ is not defined:

$$|\hat{A}\psi_1\rangle = \hat{A}\,|\psi_1\rangle = |\psi_2\rangle \tag{3.24}$$

However, the complex conjugated (adjoint) operator $\hat{A}^\dagger$ can indeed act from the right on a Bra state:

$$\langle\hat{A}\psi_1| = \langle\psi_1|\,\hat{A}^\dagger = \langle\psi_2| \tag{3.25}$$

That we can state $\langle\hat{A}\psi_1| = \langle\psi_1|\,\hat{A}^\dagger$ in (3.25) is due to the fact that the Bra-bracket $\langle\cdot|$ acts like a "dagger" operator on $\hat{A}$ and ψ, and the rule $(XY)^\dagger = Y^\dagger X^\dagger$ comes into effect.

Important Terms Regarding Operators In quantum mechanics, *hermitian* operators have particular significance. These are defined by the relation $\hat{A} = \hat{A}^\dagger$. Such operators act on a Ket vector (to the right) the same way they act on a Bra vector (to the left). *Unitary* operators, on the other hand, have the characteristic that the inverse and the adjoint of an operator $\hat{U}$ coincide:

$$\hat{U}^{-1} = \hat{U}^\dagger \quad \text{and} \quad \hat{U}\hat{U}^{-1} = \hat{U}\hat{U}^\dagger = \mathbb{1} \tag{3.26}$$

In general, an *adjoint* operator $\hat{A}^\dagger$ should be defined as follows for all $|\phi\rangle$ and $|\psi\rangle$:

$$\forall\ |\psi\rangle, |\phi\rangle :\ \langle\psi|\hat{A}^\dagger|\phi\rangle = [(\langle\psi|\hat{A}^\dagger|\phi\rangle)^\dagger]^* = \langle\phi|\hat{A}|\psi\rangle^* \tag{3.27}$$

An operator $\hat{A}$ is *homogeneous* in its action on a state $|\psi\rangle$, meaning that a scalar α and an operator $\hat{A}$ can always be interchanged:

$$\hat{A}\,|\alpha\psi\rangle = \hat{A}\alpha\,|\psi\rangle = \alpha\hat{A}\,|\psi\rangle \tag{3.28}$$

We only consider *linear* operators. This means that when $\hat{A}$ acts on a sum of states $|\psi_1\rangle$ and $|\psi_2\rangle$, it is the same as if the operator acts on each state individually, and then the sum is formed:

$$\hat{A}\,|\psi_1 + \psi_2\rangle = \hat{A}\,(|\psi_1\rangle + |\psi_2\rangle) = \hat{A}\,|\psi_1\rangle + \hat{A}\,|\psi_2\rangle \tag{3.29}$$

3.2 Hermitian Operators, Eigenfunctions, and Eigenvalues

When an operator $\hat{A}$ acts on a state $|\psi_1\rangle$, a new state is generally created, which we can call, for example, $|\psi_2\rangle$:

$$\hat{A}\,|\psi_1\rangle = |\psi_2\rangle$$

However, there are also states $|a_i\rangle$ for an operator $\hat{A}$ such that the following relation holds:

$$\hat{A}\,|a_i\rangle = a_i\,|a_i\rangle \tag{3.30}$$

The operator $\hat{A}$ maps the state $|a_i\rangle$ back onto itself, but with a scaling factor a_i (where a_i must be a scalar). Since there can be several such special states $|a_i\rangle$ for each operator, we number them with the subindex i. Generally, (3.30) reminds us of the eigenvalue problem from linear algebra, where for a matrix A, the corresponding eigenvectors $\mathbf{v}_i$ and scalars λ_i are determined, satisfying the eigenvalue equation $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$. The relationship (3.30) is thus an eigenvalue equation for the operator $\hat{A}$, with the eigenstates $|a_i\rangle$ and the eigenvalues a_i! The set of all eigenstates $|a_i\rangle$ of the operator $\hat{A}$ spans a subspace in the underlying Hilbert space, the so-called *eigenspace* of $\hat{A}$. As already mentioned, hermitian operators play a special role in quantum mechanics. To reiterate, for these operators the following holds:

$$\hat{A} = \hat{A}^\dagger \tag{3.31}$$

3.2.1 Real Eigenvalues of Hermitian Operators

And just as a hermitian matrix $A = A^\dagger$ has only real eigenvalues λ_i, a hermitian operator $\hat{A} = \hat{A}^\dagger$ has *only* real eigenvalues a_i. This can be confirmed as follows: If an operator $\hat{A}$ and a normalized eigenvector $|a_i\rangle$ are given, the associated eigenvalue a_i can be expressed as follows:

$$\underbrace{\langle a_i|\,\hat{A}\,|a_i\rangle}_{\hat{A}\text{ acts to the right}} = \underbrace{\langle a_i|a_i|a_i\rangle}_{\text{Scalar } a_i \text{ interchanged with } \langle a_i|} = a_i\ \underbrace{\langle a_i|a_i\rangle}_{\text{normalized}} = a_i \tag{3.32}$$

However, if the operator $\hat{A}$ is hermitian, you can also let $\hat{A}$ act to the left, and (3.32) is written as follows:

$$\underbrace{\langle a_i|\,\hat{A}\,|a_i\rangle}_{\hat{A}=\hat{A}^\dagger} = \underbrace{\langle a_i|\,\hat{A}^\dagger\,|a_i\rangle}_{\hat{A}^\dagger\text{ acts to the left}} = \underbrace{\langle a_i|a_i^*|a_i\rangle}_{\text{Scalar } a_i^* \text{ interchanged with } \langle a_i|} = a_i^*\ \underbrace{\langle a_i|a_i\rangle}_{\text{normalized}} = a_i^* \tag{3.33}$$

For a hermitian operator, it should make no difference whether it acts to the right on a Ket vector or to the left on a Bra vector. Thus, equations (3.32) and (3.33) must be simultaneously fulfilled. From this, it follows:

$$a_i = a_i^* \iff a_i \in \mathbb{R} \tag{3.34}$$

Thus, hermitian operators indeed possess real eigenvalues. But what does this mean physically? Recall that Ket vectors like $|a_i\rangle$ correspond to *states* of quantum systems. The application of a hermitian operator $\hat{A}$ to a state is linked in quantum theory to the *measurement* of the corresponding observable. The fact that there are corresponding eigenstates $|a_i\rangle$ with eigenvalues a_i for each operator $\hat{A}$ means the following initially: If the system is in the eigenstate $|a_i\rangle$, we *always* measure the measurement value a_i.
This may seem trivial, but recall: Measurements on quantum systems are usually accompanied by some degree of uncertainty. Even if an experiment is prepared the same way every time, the measurement outcome can deviate from the previous one every time. Therefore, a quantum state that yields the same measurement result each time is rather an exception than the rule in the quantum world and is thus something special!

Alongside, we now understand why we need hermitian operators (with real eigenvalues) to represent observables in the first place: The eigenvalues of the operators correspond to the *measurement values* of eigenstates, and measurement values are always represented by real numbers.

3.2.2 Orthogonality of Eigenstates of Hermitian Operators

We have shown in (3.34) that hermitian operators have real eigenvalues $a_i \in \mathbb{R}$. Thankfully, this means that we don't need to worry about whether a hermitian operator $\hat{A} = \hat{A}^\dagger$ acts to the right on a Ket vector or to the left on a Bra vector. Due to the hermiticity of $\hat{A}$, the following holds equally for the eigenvectors $|a_i\rangle$ and $\langle a_i|$:

$$\hat{A}\,|a_i\rangle = a_i\,|a_i\rangle \quad \text{and} \quad \langle a_i|\,\hat{A} = \langle a_i|\,a_i$$

With this, we can show that for a hermitian operator $\hat{A}$, any two arbitrary eigenstates $|a_i\rangle$ and $|a_j\rangle$ $(i \neq j)$ are always *orthogonal* to each other.

Non-degenerate System First, we let $\hat{A}$ act on a state $|a_i\rangle$ and project the result onto $\langle a_j|$ – due to the hermiticity of the operator, we could also act on $\langle a_j|$ with $\hat{A}$ and project that result onto $|a_i\rangle$. We obtain two equations:

$$\begin{aligned} &\text{(I)}| \quad \langle a_j|\hat{A}|a_i\rangle = a_i\,\langle a_j|a_i\rangle \quad | \text{ Acts to the right} \\ &\text{(II)}| \quad \langle a_j|\hat{A}|a_i\rangle = a_j\,\langle a_j|a_i\rangle \quad | \text{ Acts to the left} \end{aligned} \tag{3.35}$$

Subtracting the lower equation from the upper in (3.35) gives us:

$$\text{(I)} - \text{(II)}| \quad 0 = (a_i - a_j)\,\langle a_j|a_i\rangle \tag{3.36}$$

(3.36) is fulfilled when either the eigenvalues a_i and a_j are equal, or $\langle a_j|a_i\rangle$ vanishes; the latter occurs when an eigenstate equals the zero vector or both eigenstates are orthogonal. If the eigenvalues all satisfy $a_i \neq a_j$, then they are termed *non-degenerate*, and the only non-trivial way to satisfy (3.36) is as follows:

$$\langle a_i|a_j\rangle = 0 \tag{3.37}$$

This means that the eigenstates of hermitian operators with non-degenerate eigenvalues are orthogonal to each other! If the eigenvalues of two or more (linearly independent) eigenvectors are the same $(a_i = a_j)$, then one speaks of *degeneracy*, and the eigenfunctions are not necessarily orthogonal to each other: $\langle \psi_i|\psi_j\rangle \neq 0$. If n linearly independent states belong to the same eigenvalue, one speaks of n-fold degeneracy.

System with Degeneracy The case of degeneracy shall now be considered in more detail: Suppose we have $a_1 = a_2$ for the first two eigenvalues. The associated eigenfunctions $|a_1\rangle$ and $|a_2\rangle$ are normalized but not orthogonal due to degeneracy: $\langle a_1|a_2\rangle \neq 0$. By forming linear combinations of $|a_1\rangle$ and $|a_2\rangle$, orthogonal states can always be constructed – this can be done through a simple algorithm within the *Gram-Schmidt orthogonalization process*.
The goal of the Gram-Schmidt process is to find two new (normalized) states $|\tilde{a}_1\rangle$ and $|\tilde{a}_2\rangle$ for which now applies: $\langle \tilde{a}_1|\tilde{a}_2\rangle = 0$. For simplicity, the first new state is chosen as:

$$|\tilde{a}_1\rangle = |a_1\rangle \tag{3.38}$$

For $|\tilde{a}_2\rangle$, we initially project $|a_2\rangle$ onto $|a_1\rangle$, thus obtaining the parallel component of the two eigenstates – this is depicted by the red vector $\langle a_1|a_2\rangle\,|a_1\rangle$ in Figure 14. Subtracting the parallel component from the original state $|a_2\rangle$ leaves only the component $|\tilde{a}_2\rangle$ orthogonal to $|a_1\rangle$. Therefore, it holds:

$$|\tilde{a}_2\rangle = |a_2\rangle - \langle a_1|a_2\rangle\,|a_1\rangle \tag{3.39}$$

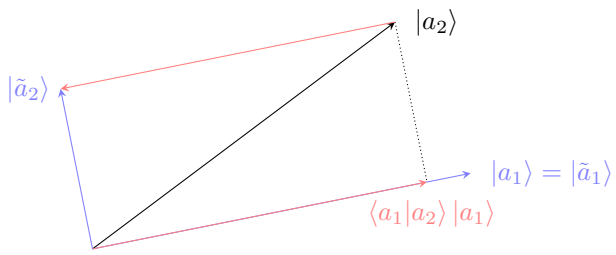


Fig. 14: Construction of a state $|\tilde{a}_2\rangle$ that satisfies the condition $\langle\tilde{a}_1|\tilde{a}_2\rangle = \langle a_1|\tilde{a}_2\rangle = 0$. Note that $|\tilde{a}_2\rangle$ must be normalized first.

Orthogonality was required between the newly constructed states, which can be easily checked:

$$\begin{aligned}
\langle\tilde{a}_1|\tilde{a}_2\rangle &= \langle a_1|\tilde{a}_2\rangle \overset{(3.39)}{=} \\
&= \langle a_1|\left(|a_2\rangle - \langle a_1|a_2\rangle\,|a_1\rangle\right) = \\
&= \langle a_1|a_2\rangle - \langle a_1|\,\langle a_1|a_2\rangle\,|a_1\rangle = \quad |\,\langle a_1|a_2\rangle \text{ is a scalar} \Rightarrow \text{factor out} \\
&= \langle a_1|a_2\rangle - \langle a_1|a_2\rangle\,\langle a_1|a_1\rangle = \quad |\,\langle a_1|a_1\rangle = 1 \\
&= \langle a_1|a_2\rangle - \langle a_1|a_2\rangle = 0 \;\square
\end{aligned} \tag{3.40}$$

For general number n of eigenfunctions to be orthogonalized, the Gram-Schmidt process states the following:

$$|\tilde{a}_i\rangle = |a_i\rangle - \sum_{j=1}^{i-1} \frac{\langle a_j|a_i\rangle}{\langle a_j|a_j\rangle}\,|a_j\rangle \tag{3.41}$$

Since the parallel component is always subtracted between the affected vectors, in the end only the desired orthogonal state vector remains.

3.3 Matrix Representation of an Operator

We now consider a hermitian operator $\hat{A}$ whose eigenfunctions $|a_j\rangle$ corresponding to the discrete eigenvalues a_j are known, fulfilling the eigenvalue equation $\hat{A}\,|a_j\rangle = a_j\,|a_j\rangle$. If this eigenvalue equation is projected onto another eigenstate a_i, we obtain:

$$A_{ij} = \langle a_i|\hat{A}|a_j\rangle = \langle a_i|a_j|a_j\rangle = a_j\,\langle a_i|a_j\rangle \tag{3.42}$$

Since a_j is a scalar quantity, we could factor it out of the Braket-bracket in the last step. A_{ij} is now the *matrix representation* of the operator $\hat{A}$ in the space of its eigenfunctions, where the index i corresponds to the row index and index j to the column index. Since i and j share the same range of values, A_{ij} is a square matrix.

Non-degenerate System If no degeneracies occur, further statements can be made about A_{ij}: Since each eigenvalue a_i is associated with a single eigenfunction $|a_i\rangle$, these must be orthogonal to each other, and (3.42) simplifies to:

$$A_{ij} = \langle a_i|\hat{A}|a_j\rangle = a_i\,\langle a_i|a_j\rangle = a_i\delta_{ij} \tag{3.43}$$

Thus, the matrix is diagonal without degeneracy and can be written in the space of eigenfunctions (in the so-called eigenbasis) as:

$$A_{ij} = \begin{pmatrix} a_1 & 0 & \vdots \\ 0 & a_2 & \\ \vdots & & \ddots \end{pmatrix} = \mathrm{diag}\{a_1, a_2, \dots\} \tag{3.44}$$

System with Degeneracy If the assumption is dropped that no degeneracies occur, strict diagonal format is lost, though block-diagonal format remains. Assuming for example two eigenfunctions $|a_k\rangle$ and $|a_l\rangle$ have the same eigenvalue a_k, then $\langle a_k|a_l\rangle \neq \delta_{kl}$ holds. The matrix is ultimately block-diagonal:

$$A_{ij} = \begin{pmatrix} a_1 & 0 & \dots & & \\ 0 & a_2 & & & \\ \vdots & & \ddots & & \\ & & & \begin{pmatrix} \ddots & \dots \\ \vdots & \ddots \end{pmatrix} & \end{pmatrix} \tag{3.45}$$

The form of the block matrix can be quickly written in the special case (twofold degeneracy of the eigenstates ($|a_3\rangle$ and $|a_4\rangle$ with a common eigenvalue a_3 from a total of four possible eigenstates): For the off-diagonal elements, it holds $\langle a_3|\hat{A}|a_4\rangle = a_3 \langle a_3|a_4\rangle$ (we let $\hat{A}$ act to the left here), and $\langle a_4|\hat{A}|a_3\rangle = a_3 \langle a_4|a_3\rangle$ (here it acts to the right), which allows us to write the block in matrix representation as:

$$A_{34}^{\text{Block}} = a_3 \begin{pmatrix} 1 & \langle a_3|a_4\rangle \\ \langle a_4|a_3\rangle & 1 \end{pmatrix} = a_3 \begin{pmatrix} 1 & \langle a_3|a_4\rangle \\ \langle a_3|a_4\rangle^* & 1 \end{pmatrix} \tag{3.46}$$

3.3.1 Projections and Spectral Representation

According to (3.20), any arbitrary state $|\psi\rangle$ in an N-dimensional Hilbert space $\mathcal{H}_N$ can always be represented by a set of basis states $\{|\phi_i\rangle\}$. It is especially practical to use normalized, orthogonal basis vectors (an orthonormal system, briefly ONS) for this. Then it holds:

$$|\psi\rangle = \sum_i c_i |\phi_i\rangle \quad \text{with} \quad \langle\phi_i|\phi_j\rangle = \delta_{ij} \tag{3.47}$$

Each expansion coefficient c_i in (3.47) can be easily calculated for an ONS – just as with position vectors in $\mathbb{R}^3$ – by projecting onto the respective basis vector $|\phi_i\rangle$:

$$c_i = \langle\phi_i|\psi\rangle \tag{3.48}$$

From the completeness of the Hilbert space $\mathcal{H}_N$, it follows that a complete orthonormal basis system (briefly: CONS) can be found. By forming the tensor product $|\phi_i\rangle \langle\phi_i|$ from the basis vectors and summing over all possible states i, a "completeness relation" (i.e., a unity operator $\mathbb{1}$ that does not change the state) can be constructed:

$$\mathbb{1} = \sum_i |\phi_i\rangle \langle\phi_i| \tag{3.49}$$

If we apply a completeness relation to a wave function $|\psi\rangle$, we naturally obtain $|\psi\rangle$ again, hence $\mathbb{1}|\psi\rangle = |\psi\rangle$. Let us take a closer look at what happens when we apply the completeness relation, as defined in (3.49), to $|\psi\rangle$:

$$|\psi\rangle = \mathbb{1}|\psi\rangle \overset{(3.49)}{=} \sum_i |\phi_i\rangle \langle\phi_i|\psi\rangle = \sum_i \langle\phi_i|\psi\rangle |\phi_i\rangle \overset{(3.48)}{=} \sum_i c_i |\phi_i\rangle \overset{(3.47)}{=} |\psi\rangle \quad \square$$

Hence, we can regard $|\phi_i\rangle \langle\phi_i|$ as an operator $\hat{P}_i$ that projects $|\psi\rangle$ onto the ith eigenstate $|\phi_i\rangle$, scaling with c_i. The effect of $\hat{P}_i$ is defined as follows:

$$\hat{P}_i |\psi\rangle = c_i |\phi_i\rangle \quad \text{with} \quad \hat{P}_i = |\phi_i\rangle \langle\phi_i| \tag{3.50}$$

$\hat{P}_i$ is a *projection operator*, which is recognizable by its idempotence: $\hat{P}_i^2 = \hat{P}_i$ (and accordingly holds over complete induction: $\hat{P}_i^n = \hat{P}_i$). That this is so can be easily demonstrated:

$$\hat{P}_i^2 \overset{(3.50)}{=} |\phi_i\rangle \langle\phi_i|\phi_i\rangle \langle\phi_i| = |\phi_i\rangle \langle\phi_i| \overset{(3.50)}{=} \hat{P}_i \quad \square$$

The trick that one can insert the unity operator $\mathbb{1}$ at any time is also used to represent an operator $\hat{A}$ in its eigenbasis $\{|a_i\rangle\}$:

$$\begin{aligned}
\hat{A} &= \mathbb{1}\hat{A}\mathbb{1} \overset{(3.49)}{=} \\
&= \left(\sum_i |a_i\rangle \langle a_i|\right)\hat{A}\left(\sum_j |a_j\rangle \langle a_j|\right) = \\
&= \sum_i \sum_j |a_i\rangle \langle a_i|\hat{A}|a_j\rangle \langle a_j| = \quad | \; \hat{A}\,|a_j\rangle = a_j\,|a_j\rangle \\
&= \sum_i \sum_j |a_i\rangle \langle a_i|a_j|a_j\rangle \langle a_j| = \quad | \; a_j \text{ is a scalar} \Rightarrow \text{pull out} \\
&= \sum_i \sum_j a_j\,|a_i\rangle \langle a_i|a_j\rangle \langle a_j| = \quad | \; \langle a_i|a_j\rangle = \delta_{ij} \\
&= \sum_i \sum_j a_j\,|a_i\rangle \delta_{ij} \langle a_j| = \\
&= \sum_i a_i\,|a_i\rangle \langle a_i|
\end{aligned} \tag{3.51}$$

This representation of an operator is called the *spectral representation* in the eigenbasis. In this eigenbasis, $\hat{A}$ is then representable as a diagonal matrix with the corresponding eigenvalues in the diagonal:

$$\hat{A} \xrightarrow{\{|a_i\rangle\}} \begin{pmatrix} a_1 & 0 & \dots & \\ 0 & a_2 & & \\ \vdots & & \ddots & \\ & & & a_N \end{pmatrix} \tag{3.52}$$

If the operator $\hat{A}$ is hermitian and thus satisfies $\hat{A} = \hat{A}^\dagger$, all eigenvalues $a_i = a_i^*$ are real quantities. This can be utilized in the following relation:

$$\hat{A}^\dagger = \sum_i (a_i\,|\phi_i\rangle \langle\phi_i|)^\dagger = \sum_i a_i^*\,\langle\phi_i|^\dagger\,|\phi_i\rangle^\dagger = \sum_i a_i\,|\phi_i\rangle \langle\phi_i| = \hat{A} \quad \square \tag{3.53}$$

3.4 Distribution Space

Motivation: Continuous Observables

In the previous chapters, we have (more or less tacitly) assumed that for an operator $\hat{A}$ there is always a *countable* number of eigenstates. The eigenvalue equation $\hat{A}\,|a_i\rangle = a_i\,|a_i\rangle$ expresses this through the discrete index i. The physical meaning was that with every single measurement, we can only measure one value from the list of possible operator eigenvalues $\{a_i\}$. The corresponding eigenstates $\{|a_i\rangle\}$ form an ONB, which is why we can represent general states $|\psi\rangle$ as a superposition of the eigenstates $|a_i\rangle$: $|\psi\rangle = \sum_i c_i\,|a_i\rangle$ with $c_i = \langle a_i|\psi\rangle$. Even more simply, we can represent $|\psi\rangle$ in the basis $\{|a_i\rangle\}$ as a column vector and operators as square $N \times N$ matrices.

This makes physical sense in many problems: For instance, if one measures the energy of a particle in an infinitely deep potential well, the measuring device will display a value from the list of eigenenergies belonging to the Hamiltonian operator with every energy measurement. There is a lowest energy E_1 (ground state energy), then a next higher energy E_2 (energy of the first excited state), etc. Even if we assume infinitely many energy values: The energy basis remains (in this example) always countable.

Apparently, there are observables that behave differently: If we measure, for example, the position of a particle in a one-dimensional system along the x-axis, this is expressed by the corresponding position operator $\hat{x}$. However, in general, we will not expect that there exists only a countable set $\{x_i\}$ of places where we can find the particle during a measurement. The notation $\hat{x}\,|x_i\rangle = x_i\,|x_i\rangle$, which implies discrete eigenvalues $\{x_i\}$ and position eigenvectors $\{|x_i\rangle\}$, therefore no longer makes sense.

In the following chapters, we will deal with how to solve this problem. Obviously, one must move from a countable set of basis vectors to an uncountable set; accordingly, a general state $|\psi\rangle$ can no longer be represented as a discrete superposition of basis vectors, but only through a corresponding integral. Related to this, we will also reveal the secret of how to transition from the abstract ket representation $|\psi\rangle$ to the more concrete wave function $\psi(x)$, or, for example, to $\tilde{\psi}(p)$.

3.4.1 Position and Momentum Eigenvalues and Eigenfunctions

In a one-dimensional task, we will generally not expect that there exists only a countable set $\{x_i\}$ of places where a particle can be located during a measurement. The notation $\hat{x}\,|x_i\rangle = x_i\,|x_i\rangle$ is therefore meaningless if one assumes that i may only take integer values. To fix the problem, one could, of course, agree that the index i can take not only integer values but any real values. For example, $x_{1.234}$ would be the eigenvalue corresponding to a measurement where we find the particle at position $x = 1.234$. However, from this example, one can already see that we no longer need the index i in this case. If x can already take any real value, it makes sense in our notation to use x instead of x_i. And since for every real eigenvalue x, there is a corresponding eigenvector, we can also write this with $|x\rangle$ instead of $|x_i\rangle$. This allows us to write the (uncountably infinite) set of eigenvalue equations for the position operator $\hat{x}$ in one dimension as follows:

$$\hat{x}\,|x\rangle = x\,|x\rangle \tag{3.54}$$

Correspondingly, we write for the three-dimensional position operator $\hat{\mathbf{r}}$ a similar eigenvalue equation:

$$\hat{\mathbf{r}}\,|\mathbf{r}\rangle = \mathbf{r}\,|\mathbf{r}\rangle \tag{3.55}$$

The (uncountably infinite) set of position eigenvectors $\{|\mathbf{r}\rangle\}$ is called the *position eigenbasis*. It should be noted that in the three-dimensional case, the corresponding eigenvalues $\mathbf{r}$ are, of course, position vectors that must be fully expressed with three numerical values.
Analogously, one can define the *momentum eigenbasis* $\{|\mathbf{p}\rangle\}$ for the three-dimensional momentum operator $\hat{\mathbf{p}}$. For every momentum eigenvector $|\mathbf{p}\rangle$, there exists an eigenvalue $\mathbf{p}$, so the following eigenvalue equation holds:

$$\hat{\mathbf{p}}\,|\mathbf{p}\rangle = \mathbf{p}\,|\mathbf{p}\rangle \tag{3.56}$$

Of course, an eigenvalue equation for each component of the momentum operator $\hat{p}_i$ can also be written. Because $\mathbf{p} = \hbar\mathbf{k}$, one can equivalently also write for the wave number operator $\hat{\mathbf{k}}$:

$$\hat{\mathbf{k}}\,|\mathbf{k}\rangle = \mathbf{k}\,|\mathbf{k}\rangle \tag{3.57}$$

At this point, we must note: The action of $\hat{\mathbf{p}}$ or $\hat{\mathbf{k}}$ from (3.56) and (3.57) is so simple only because we write the eigenvalue problem in the respective eigenbasis $\{|\mathbf{p}\rangle\}$ or $\{|\mathbf{k}\rangle\}$. What the action of the momentum operator in the position basis $\{|\mathbf{r}\rangle\}$ looks like, we will discuss in a later section.

3.4.2 From the Abstract Ket Vector to the Concrete Wave Function

Let us recap how a wave function can be decomposed with a *countable* basis $\{|a_i\rangle\}$ and how the normalization appears:

$$|\psi\rangle = \mathbb{1}\,|\psi\rangle = \sum_i |a_i\rangle \underbrace{\langle a_i|\psi\rangle}_{c_i} = \sum_i c_i\,|a_i\rangle \quad \text{and} \quad \langle\psi|\psi\rangle = \sum_i c_i^* c_i = 1$$

If we are dealing with a continuous, *uncountable* basis such as the one-dimensional position basis $\{|x\rangle\}$, then we must replace the sum with an integral:

$$|\psi\rangle = \mathbb{1}\,|\psi\rangle = \int dx\ |x\rangle \underbrace{\langle x|\psi\rangle}_{c(x)} = \int dx\, c(x)\,|x\rangle \tag{3.58}$$

The normalization can also be represented in a form similar to the discrete case, which we will later justify more precisely:

$$\langle\psi|\psi\rangle = \int dx\, c^*(x)c(x) = 1 \tag{3.59}$$

We also know that the scalar product $\langle\psi|\psi\rangle$ is independent of the representation of the vectors; therefore, the following relation must also be valid:

$$\langle\psi|\psi\rangle = \int dx\, \psi^*(x)\psi(x) = 1 \tag{3.60}$$

A comparison of equations (3.59) and (3.60) shows that $\psi(x)$ and $c(x)$ are apparently identical, i.e., $\psi(x) \equiv c(x)$. From (3.58) we additionally know that $c(x) = \langle x|\psi\rangle$. Thus, we have now discovered how to formally transition from the abstract vector $|\psi\rangle$ to the wave function $\psi(x)$:

$$\psi(x) = \langle x|\psi\rangle \tag{3.61}$$

The abstract wave function $|\psi\rangle$ is thus projected into the position basis $\{|x\rangle\}$ to enable a meaningful representation of a state. In three-dimensional space, for a wave function, the same reasoning as in one dimension applies analogously:

$$\psi(\mathbf{r}) = \langle \mathbf{r}|\psi\rangle \tag{3.62}$$

If one wants a wave function $\tilde{\psi}(\mathbf{p})$, which does not have the position as a function argument but a momentum, one can project the abstract ket vector $|\psi\rangle$ onto the momentum eigenbasis analogously:

$$\tilde{\psi}(\mathbf{p}) = \langle \mathbf{p}|\psi\rangle \tag{3.63}$$

Because of the relation $\mathbf{p} = \hbar\mathbf{k}$ we can write equivalently for a wave function $\bar{\psi}(\mathbf{k})$ with a wave vector as argument:

$$\bar{\psi}(\mathbf{k}) = \langle \mathbf{k}|\psi\rangle \tag{3.64}$$

So one recognizes the following: Just as, for example, the same position vector $\mathbf{r}$ in Euclidean space can be represented by completely different numeric values (expansion coefficients c_i), the same state vector $|\psi\rangle$ in the Hilbert space $\mathcal{H}$ can be represented by entirely different mathematical functions $\psi(\mathbf{r})$ or $\tilde{\psi}(\mathbf{p})$. These ultimately represent nothing more than an uncountably large set of "expansion coefficients" for different bases.

3.4.3 Position and Momentum Eigenfunctions in Position and Momentum Basis

For orthogonal, normalized base systems with *countably many* basis vectors $\{|a_i\rangle\}$, we have already stated that (due to orthogonality and normalization) it must trivially hold:

$$\langle a_i|a_j\rangle = \delta_{ij}$$

In the transition to orthogonal, normalized base systems with *uncountably many* basis vectors, such as the position basis $\{|x\rangle\}$, it holds analogously:

$$\langle x|x'\rangle = \delta(x - x') \tag{3.65}$$

With $\delta(x - x')$ is meant the Dirac delta function (or more precisely: delta distribution), which is defined by:

$$\int_{-\infty}^{\infty} dx\, \delta(x - x')\, f(x) = f(x') \tag{3.66}$$

In-Depth: Delta Distribution

The delta distribution $\delta(x)$ assigns the evaluation of the function at $x = 0$ to every arbitrarily often differentiable function $f(x)$. One can write the functional:

$$\langle \delta, f \rangle = f(0) \tag{3.67}$$

This pairing takes place formally in an integral of the following form:

$$\langle \delta, f \rangle = \int_{-\infty}^{\infty} dx\, \delta(x) f(x) = f(0) \tag{3.68}$$

This shows that the delta distribution $\delta(x)$ must have the property of taking the value 0 everywhere for $x \neq 0$. Around the value $x = 0$, $\delta(x)$ must – graphically speaking – form an "infinitely narrow" peak that has the exact area of 1, so that it holds:

$$\int_{-\infty}^{\infty} dx\, \delta(x) = 1 \tag{3.69}$$

Although in physics one often speaks of the "delta function," strictly speaking, there is no function that meets this definition. $\delta(x)$ is only defined by (3.67) and (3.68) and is thus more accurately called delta distribution. However, $\delta(x)$ can be represented as the limit of a sequence of functions, whereby there are many possibilities for representation here. One of these uses an ever narrower Gaussian function or sinc function:

$$\begin{aligned} \delta(x) &= \lim_{\sigma \to 0} \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/(2\sigma^2)} \\ \delta(x) &= \lim_{n \to 0} \frac{1}{n} \mathrm{sinc}(x/n) \end{aligned} \tag{3.70}$$

The most common calculation rules with the δ function are given below (noting that we *always* have to integrate over $\delta(x)$):

$$f(x)\delta(x - x_0) = f(x_0)\delta(x - x_0) \tag{3.71}$$

$$\delta(x) = \delta(-x) \tag{3.72}$$

$$x\delta(x) = 0 \tag{3.73}$$

$$\delta(ax) = \delta(x)/|a| \tag{3.74}$$

$$|x|\, \delta(x^2) = \delta(x) \tag{3.75}$$

In the case of the following relation (3.76), x_i are the *simple* zeros of the function $f(x)$.

$$\delta(f(x)) = \sum_i \frac{\delta(x - x_i)}{|f'(x_i)|} \tag{3.76}$$

In (3.61), we have learned that $\langle x|\psi\rangle = \psi(x)$ holds. Thus, the expression $\langle x|x'\rangle$ in equation (3.65) is nothing other than the eigenfunction $|x'\rangle$ of the position operator (to the eigenvalue x') represented in the position basis $\{|x\rangle\}$ – in other words, the projection of one position state $|x'\rangle$ onto another position state $|x\rangle$. Because the eigenstates of the position operator form an eigenbasis, different position states are orthogonal to each other, and it follows:

$$\langle x|x'\rangle = \phi_{x'}(x) = \delta(x - x') \tag{3.77}$$

The scalar product only does not vanish if both position states coincide ($x' = x$). Completely

analogously for k-space or momentum space, one can write:

$$\begin{aligned}\langle p|p'\rangle &= \phi_{p'}(p) = \delta(p-p')\\ \langle k|k'\rangle &= \phi_{k'}(k) = \delta(k-k')\end{aligned} \tag{3.78}$$

In three-dimensional space, similar relations hold, but at this point, we have to use the three-dimensional delta distribution:

$$\begin{aligned}\langle \mathbf{r}|\mathbf{r}'\rangle &= \phi_{\mathbf{r}'}(\mathbf{r}) = \delta^{(3)}(\mathbf{r}-\mathbf{r}')\\ \langle \mathbf{p}|\mathbf{p}'\rangle &= \phi_{\mathbf{p}'}(\mathbf{p}) = \delta^{(3)}(\mathbf{p}-\mathbf{p}')\\ \langle \mathbf{k}|\mathbf{k}'\rangle &= \phi_{\mathbf{k}'}(\mathbf{k}) = \delta^{(3)}(\mathbf{k}-\mathbf{k}')\end{aligned} \tag{3.79}$$

Representation in Eigen Space We now want to determine the effect of the position operator $\hat{x}$ in the position basis (its eigenbasis). The eigenvalue equation to be fulfilled is:

$$\hat{x}\,|x'\rangle = x'\,|x'\rangle$$

To exploit relationship (3.77), we project this relation from the left onto $\langle x|$:

$$\langle x|\hat{x}|x'\rangle = \begin{cases}\langle x|\hat{x}x'\rangle = \langle x|x'|x'\rangle = x'\,\langle x|x'\rangle \overset{(3.77)}{=} x'\,\delta(x-x')\\ \langle x\hat{x}|x'\rangle = \langle x|x|x'\rangle = x\,\langle x|x'\rangle \overset{(3.77)}{=} x\,\delta(x-x')\end{cases} \tag{3.80}$$

Because the position operator is Hermitian $\hat{x} = \hat{x}^\dagger$, we can let it act both to the right and to the left – from this, it can be concluded that $x\delta(x-x') = x'\,\delta(x-x')$ holds. In the one- and three-dimensional position space, we recognize through this the (very simple) form of the position operator $\hat{x}$ or $\hat{\mathbf{r}}$:

$$\hat{x} \xrightarrow{\{|x\rangle\}} x \tag{3.81}$$

$$\hat{\mathbf{r}} \xrightarrow{\{|\mathbf{r}\rangle\}} \mathbf{r} \tag{3.82}$$

The same applies correspondingly for the momentum operator $\hat{p}_i \equiv \hat{p}$ or $\hat{\mathbf{p}}$ in momentum space, where we can argue analogously to position space:

$$\hat{p} \xrightarrow{\{|p\rangle\}} p \tag{3.83}$$

$$\hat{\mathbf{p}} \xrightarrow{\{|\mathbf{p}\rangle\}} \mathbf{p} \tag{3.84}$$

Next, we determine the momentum eigenfunctions $|p_x\rangle \equiv |p\rangle$ in the one-dimensional position basis $\phi_p(x) = \langle x|p\rangle$. The eigenvalue equation to be fulfilled for the operator $\hat{p}_x \equiv \hat{p}$ reads:

$$\hat{p}\,|p\rangle = p\,|p\rangle$$

Since we are interested in the representation in the position basis, we project the equation onto a position eigenstate $\langle x|$. In this case, we can use the already motivated position representation of the momentum operator from (2.15):

$$\langle x|\hat{p}|p\rangle = \langle x|\hat{p}p\rangle = \langle x|p|p\rangle = p\,\langle x|p\rangle = p\,\phi_p(x) \xrightarrow{(2.14)} -i\hbar\frac{\partial}{\partial x}\phi_p(x) = p\,\phi_p(x) \tag{3.85}$$

We obtain a differential equation for the function $\phi_p(x)$. For the momentum eigenfunctions in the position basis, we assume plane waves, which we know solve the Schrödinger equation:

$$\phi_p(x) = \frac{1}{\sqrt{2\pi\hbar}}e^{ikx} = \frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar} \tag{3.86}$$

It should be noted that $\phi_p(x)$ is not an element of the Hilbert space because plane waves are not spatially limited and therefore not square-integrable (the integral over the entire space approaches infinity). However, by superimposing plane waves, as shown in Chapter 2.2, wave packets with limited extension can be formed, which then do form an element of $\mathcal{H}$. To show that the Ansatz (3.86) solves the eigenvalue problem, we must substitute into (3.85):

$$-i\hbar\frac{\partial}{\partial x}\left(\frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar}\right) = -i\hbar\left(-\frac{1}{i\hbar}p\right)\left(\frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar}\right) = p\left(\frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar}\right) \quad \square$$

Thus, we have shown that the expression (3.86) represents the momentum eigenfunctions in the position basis. We can therefore write the eigenfunction $|p\rangle$ of the momentum operator, expressed in the position basis, as follows:

$$\langle x|p\rangle = \phi_p(x) = \frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar} \tag{3.87}$$

Because $p = \hbar k$, we can choose, up to constants, the same ansatz for $\phi_k(x)$ (which solves the eigenvalue equation $\hat{p}\,\phi_k(x) = \hbar k\phi_k(x)$):

$$\langle x|k\rangle = \phi_k(x) = \frac{1}{\sqrt{2\pi}}e^{ikx} \tag{3.88}$$

In three-dimensional space, we can choose a similar Ansatz for each:

$$\begin{aligned}\langle \mathbf{r}|\mathbf{p}\rangle &= \phi_{\mathbf{p}}(\mathbf{r}) = \frac{1}{\sqrt{2\pi\hbar}}e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}\\ \langle \mathbf{r}|\mathbf{k}\rangle &= \phi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{2\pi}}e^{i\mathbf{k}\cdot\mathbf{r}}\end{aligned} \tag{3.89}$$

As the last point in this section, the eigenfunction $|x\rangle$ of the position operator, expressed in the momentum or k-basis, is still missing. Because $\langle p|x\rangle = \langle x|p\rangle^*$, we can immediately write this with a glance at (3.87) and (3.88):

$$\begin{aligned}\langle p|x\rangle &= \phi_x(p) = \frac{1}{\sqrt{2\pi\hbar}}e^{-ipx/\hbar}\\ \langle k|x\rangle &= \phi_x(k) = \frac{1}{\sqrt{2\pi}}e^{-ikx}\end{aligned} \tag{3.90}$$

In three-dimensional space, it holds once again analogously:

$$\begin{aligned}\langle \mathbf{p}|\mathbf{r}\rangle &= \phi_{\mathbf{r}}(\mathbf{p}) = \frac{1}{\sqrt{2\pi\hbar}}e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar}\\ \langle \mathbf{k}|\mathbf{r}\rangle &= \phi_{\mathbf{r}}(\mathbf{k}) = \frac{1}{\sqrt{2\pi}}e^{-i\mathbf{k}\cdot\mathbf{r}}\end{aligned} \tag{3.91}$$

We can already recognize that the momentum eigenfunctions in position space have a great similarity to the Fourier transformation between momentum and position space. In the next section, we want to take advantage of this to enable a continuous transformation between the two eigenbases.

3.4.4 Continuous Spectrum and Fourier Transformation

In-Depth: Fourier Transformation

For Fourier transformations (and the corresponding inverse transformations), there are different conventions regarding the prefactors and the sign in the exponent of the exponential function. In this script, we follow the following conventions:

Transformation from Position Space to Momentum Space (1D):

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx\, \psi(x)\, e^{-ipx/\hbar} \quad \text{and} \quad \bar{\psi}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx\, \psi(x)\, e^{-ikx} \tag{3.92}$$

Transformation from Position Space to Momentum Space (3D):

$$\tilde{\psi}(\mathbf{p}) = \frac{1}{(2\pi\hbar)^{3/2}} \int_{V} d^3r\, \psi(\mathbf{r})\, e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar} \quad \text{and} \quad \bar{\psi}(\mathbf{k}) = \frac{1}{(2\pi)^{3/2}} \int_{V} d^3r\, \psi(\mathbf{r})\, e^{-i\mathbf{k}\cdot\mathbf{r}} \tag{3.93}$$

Inverse Transformation from Momentum Space to Position Space (1D):

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp\, \tilde{\psi}(p)\, e^{+ipx/\hbar} \quad \text{and} \quad \psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk\, \bar{\psi}(k)\, e^{+ikx} \tag{3.94}$$

Inverse Transformation from Momentum Space to Position Space (3D):

$$\psi(\mathbf{r}) = \frac{1}{(2\pi\hbar)^{3/2}} \int_{p^3} d^3p\, \tilde{\psi}(\mathbf{p})\, e^{+i\mathbf{p}\cdot\mathbf{r}/\hbar} \quad \text{and} \quad \psi(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \int_{k^3} d^3k\, \bar{\psi}(\mathbf{k})\, e^{+i\mathbf{k}\cdot\mathbf{r}} \tag{3.95}$$

Transformation of the Delta Distribution from Position Space to Momentum Space:
To determine which result the transformation of the Delta distribution into momentum space yields, we simply set $\psi(x) = \delta(x - x_0)$, and perform the corresponding integrals from (3.92):

$$\tilde{\delta}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx\, \delta(x - x_0)\, e^{-ip(x-x_0)/\hbar} = \frac{1}{\sqrt{2\pi\hbar}} \tag{3.96}$$

$$\bar{\delta}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx\, \delta(x - x_0)\, e^{-ik(x-x_0)} = \frac{1}{\sqrt{2\pi}} \tag{3.97}$$

The Delta distribution transformed into momentum space $\delta(x - x_0)$ therefore yields a constant!

Integral Representation of the Delta Distribution:
By substituting the transformed delta distributions $\tilde{\delta}(p) = 1/\sqrt{2\pi\hbar}$ and $\bar{\delta}(k) = 1/\sqrt{2\pi}$ into the corresponding inverse transformations (3.94), we obtain integral expressions for the delta distribution that are often used in practice:

$$\delta(x - x_0) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp\, \frac{1}{\sqrt{2\pi\hbar}}\, e^{ip(x-x_0)/\hbar} = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dp\, e^{ip(x-x_0)/\hbar} \tag{3.98}$$

$$\delta(x - x_0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk\, \frac{1}{\sqrt{2\pi}}\, e^{ik(x-x_0)} = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{ik(x-x_0)} \tag{3.99}$$

Just as in section 3.3.1, a "completeness relation" $\mathbb{1}$ was formed by summing over all basis vectors of a discrete CONS – see equation (3.49) – one can also construct a "completeness relation" with position or momentum eigenfunctions. Since the set of eigenvectors is uncountable in this case,

the sum must be replaced by an integral:

$$\mathbb{1} = \int dx\ |x\rangle\langle x| = \int dp\ |p\rangle\langle p| = \int dk\ |k\rangle\langle k| \tag{3.100}$$

Similar to the case of discrete eigenfunctions, one can again express an arbitrary state $|\psi\rangle$ in a certain basis by "inserting a one":

$$|\psi\rangle = \mathbb{1}\,|\psi\rangle = \begin{cases} \int dx\ |x\rangle\langle x|\psi\rangle\,, & \text{(Position space)} \\ \int dp\ |p\rangle\langle p|\psi\rangle\,, & \text{(Momentum space)} \\ \int dk\ |k\rangle\langle k|\psi\rangle\,, & (k\text{-space}) \end{cases} \tag{3.101}$$

The expressions $\hat{P}_x = |x\rangle\langle x|$, $\hat{P}_p = |p\rangle\langle p|$, and $\hat{P}_k = |k\rangle\langle k|$ can again be regarded as projection operators that project the state $|\psi\rangle$ onto the respective basis vector.

In the following, we briefly demonstrate how one can, for example, prove the identity $\langle x|x'\rangle = \delta(x - x')$ using the Fourier integral representation of the delta function (3.98):

$$\begin{aligned} \langle x|x'\rangle = \langle x|\mathbb{1}|x'\rangle &\overset{(3.100)}{=} \int_{-\infty}^{+\infty} dk\ \langle x|k\rangle\langle k|x'\rangle = \int_{-\infty}^{+\infty} dk\ \langle k|x\rangle^*\langle k|x'\rangle = \\ &= \int_{-\infty}^{+\infty} dk\,\phi_x^*(k)\phi_{x'}(k) \overset{(3.87)}{=} \int_{-\infty}^{+\infty} dk \frac{1}{\sqrt{2\pi}} e^{ikx} \frac{1}{\sqrt{2\pi}} e^{-ikx'} = \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} dk\, e^{ik(x-x')} \overset{(3.98)}{=} \delta(x - x') \end{aligned} \tag{3.102}$$

At this point, it can also be easily shown that a general state $\bar{\psi}(k)$ is nothing other than the Fourier transform of $\psi(x)$:

$$\begin{aligned} \bar{\psi}(k) = \langle k|\psi\rangle = \langle k|\mathbb{1}|\psi\rangle &\overset{(3.100)}{=} \int_{-\infty}^{+\infty} dx\ \langle k|x\rangle\langle x|\psi\rangle = \int_{-\infty}^{+\infty} dx\,\phi_x(k)\psi(x) = \\ &= \int_{-\infty}^{+\infty} dx\,\psi(x)\phi_x(k) \overset{(3.87)}{=} \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} dx\,\psi(x)e^{-ikx} \end{aligned} \tag{3.103}$$

Conversely, $\psi(x)$ is of course nothing other than the inverse Fourier transform of $\bar{\psi}(k)$:

$$\begin{aligned} \psi(x) = \langle x|\psi\rangle = \langle x|\mathbb{1}|\psi\rangle &\overset{(3.100)}{=} \int_{-\infty}^{+\infty} dk\ \langle x|k\rangle\langle k|\psi\rangle = \int_{-\infty}^{+\infty} dk\,\phi_k(x)\bar{\psi}(k) = \\ &= \int_{-\infty}^{+\infty} dk\,\bar{\psi}(k)\phi_k(x) \overset{(3.87)}{=} \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} dk\,\bar{\psi}(k)e^{ikx} \end{aligned} \tag{3.104}$$

In-Depth: Action of the Momentum Operator in Position Space

Already in (2.15), the 1D momentum operator $\hat{p}$, which acts on wave functions in position space $\psi(x)$, was derived. Here we show another derivation, based on the methods of the last chapters:

$$\hat{p}\psi(x) = \langle x|\hat{p}|\psi\rangle = \langle x|\hat{p}\mathbb{1}|\psi\rangle \overset{(3.100)}{=} \int_{-\infty}^{\infty} dp\, \langle x|\hat{p}|p\rangle \langle p|\psi\rangle = \int_{-\infty}^{\infty} dp\, \langle x|p|p\rangle \langle p|\psi\rangle =$$
$$= \int_{-\infty}^{\infty} dp\, p \langle x|p\rangle \langle p|\psi\rangle \overset{3.86}{=} \int_{-\infty}^{\infty} dp\, \frac{1}{\sqrt{2\pi\hbar}}\, p\, e^{ipx/\hbar} \langle p|\psi\rangle \tag{3.105}$$

We conduct a side calculation here to simplify the expression from (3.105):

$$\frac{\partial}{\partial x} e^{ipx/\hbar} = \frac{ip}{\hbar} e^{ipx/\hbar} \implies p\, e^{ipx/\hbar} = \frac{\hbar}{i}\frac{\partial}{\partial x} e^{ipx/\hbar} \equiv -i\hbar \frac{\partial}{\partial x} e^{ipx/\hbar} \tag{3.106}$$

The momentum p can thus simply be replaced by a derivative with respect to position! We substitute (3.106) into (3.105):

$$\hat{p}\psi(x) = -i\hbar \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp\, \frac{\partial}{\partial x} e^{ipx/\hbar} \langle p|\psi\rangle = -i\hbar \frac{1}{\sqrt{2\pi\hbar}} \frac{\partial}{\partial x} \int_{-\infty}^{\infty} dp\, e^{ipx/\hbar} \langle p|\psi\rangle =$$
$$= -i\hbar \frac{\partial}{\partial x} \int_{-\infty}^{\infty} dp\, \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \langle p|\psi\rangle \overset{3.86}{=} -i\hbar \frac{\partial}{\partial x} \int_{-\infty}^{\infty} dp\, \langle x|p\rangle \langle p|\psi\rangle \overset{(3.100)}{=}$$
$$= -i\hbar \frac{\partial}{\partial x} \langle x|\mathbb{1}|\psi\rangle = -i\hbar \frac{\partial}{\partial x} \langle x|\psi\rangle = -i\hbar \frac{\partial}{\partial x} \psi(x) \tag{3.107}$$

In three dimensions, this expression can be generalized to the already known expression:

$$\hat{\mathbf{p}} = -i\hbar \boldsymbol{\nabla} \tag{3.108}$$

3.4.5 General Spectral Representation

There are physical systems for which both discrete spectral decomposition and continuous spectral decomposition are needed. For example, consider the possible eigenenergies of an electron in the Coulomb potential of the nucleus of a hydrogen atom: the bound states correspond to the discrete eigenvectors of the Hamiltonian operator $\hat{H}$. The unbound (scattering) states correspond to the continuous eigenstates of the Hamiltonian operator. In such a case, to consider the whole, we must break down the Hamiltonian operator into two parts:

$$\hat{H} = \sum_n E_n |n\rangle \langle n| + \int dE\, E\, |\phi_E\rangle \langle \phi_E| \tag{3.109}$$

$|n\rangle$ represents the (countable) bound energy eigenstates, while $|\phi_E\rangle$ represents the eigenstates in the continuous spectrum region, thus describing the unbound states.

3.4.6 Direct Sum and Tensor Product of Vector Spaces

In the present case, an operator was decomposed into its discrete spectrum (with an N-dimensional Hilbert space $\mathcal{H}_N$) and its continuous spectrum (with the distribution space $\mathcal{D}$ of continuous states). To understand how the entire Hilbert space can be described, we look at the generalized case of two vector spaces $\mathcal{V}_1$ and $\mathcal{V}_2$, which are supposed to span a new vector space via the direct sum:

$$\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2 \tag{3.110}$$

The dimension of $\mathcal{V}$ is simply obtained via the sum of the dimensions of the individual vector spaces:

$$\dim \mathcal{V} = \dim \mathcal{V}_1 + \dim \mathcal{V}_2 \tag{3.111}$$

Now, consider two elements from $\mathcal{V}$, namely

$$u = \begin{pmatrix} u_1 \in \mathcal{V}_1 \\ u_2 \in \mathcal{V}_2 \end{pmatrix} \quad \text{and} \quad w = \begin{pmatrix} w_1 \in \mathcal{V}_1 \\ w_2 \in \mathcal{V}_2 \end{pmatrix}.$$

When forming the scalar product between u and w, we must consider the affiliation to the respective vector space:

$$\langle u|w\rangle = \underbrace{\langle u_1|w_1\rangle}_{\in\mathcal{V}_1} + \underbrace{\langle u_2|w_2\rangle}_{\in\mathcal{V}_2} \tag{3.112}$$

Alternatively, a vector space $\mathcal{V}$ can also be formed using the so-called tensor product between two vector spaces:

$$\mathcal{V} = \mathcal{V}_1 \otimes \mathcal{V}_2 \tag{3.113}$$

As a concrete example for the use of the tensor product, separable wave functions $\psi(x_1, x_2)$ can be used, for which: $\psi(x_1, x_2) = \psi_1(x_1)\psi_2(x_2)$. If one wants to calculate the dimension of the vector space $\mathcal{V}$ in this case, it follows:

$$\dim \mathcal{V} = \dim \mathcal{V}_1 \cdot \dim \mathcal{V}_2 \tag{3.114}$$

When calculating the scalar product between u and w, we must form the product between the elements from the respective vector spaces:

$$\langle u|w\rangle = \underbrace{\langle u_1|w_1\rangle}_{\in\mathcal{V}_1} \cdot \underbrace{\langle u_2|w_2\rangle}_{\in\mathcal{V}_2} \tag{3.115}$$

This can be easily illustrated with the example of a separable function $\psi(x_1, x_2) = \psi_1(x_1)\psi_2(x_2)$. For the projection onto itself, we get:

$$\begin{aligned}
\langle\psi|\psi\rangle &= \iint dx_1 dx_2\, \psi^*(x_1, x_2)\psi(x_1, x_2) = \\
&= \iint dx_1 dx_2\, \psi_1^*(x_1)\psi_2^*(x_2)\psi_1(x_1)\psi_2(x_2) = \\
&= \int dx_1\, \psi_1^*(x_1)\psi_1(x_1) \int dx_2\, \psi_2^*(x_2)\psi_2(x_2) = \\
&= \langle\psi_1|\psi_1\rangle\, \langle\psi_2|\psi_2\rangle
\end{aligned} \tag{3.116}$$

We recognize that in the case of the separable function, the elements of the respective sub-vector space can be naturally separated.

3.5 Operator Algebra

We have already learned about the meaning and functioning of operators, as well as discussed their properties. Fundamentally, operators can be understood as the mapping of one state to another state, with both lying in the same vector space. This chapter will explain how to handle multiple operators and delve deeper into *complementary* or *compatible* operators.

Let a general state $|\psi\rangle$ be given, and two arbitrary operators $\hat{A}$ and $\hat{B}$ whose effect on $|\psi\rangle$ we do not know in detail. If we let both operators act on the state, the following notation should apply:

$$\hat{A}\hat{B}\,|\psi\rangle = \hat{A}(\hat{B}\,|\psi\rangle) \quad \text{and} \quad \hat{B}\hat{A}\,|\psi\rangle = \hat{B}(\hat{A}\,|\psi\rangle)$$

The first equation means that $\hat{A}$ acts on a state that has already been changed by the action of $\hat{B}$. The second equation similarly means that $\hat{B}$ acts on a state that has already been changed by the action of $\hat{A}$. In general, it holds that $\hat{A}\hat{B} \neq \hat{B}\hat{A}$, and thus the two equations above will yield different results!

Example: Commutativity of Position and Momentum Operator

Let a state $\psi(x)$ be given, which is represented in the one-dimensional position space. Likewise, we know the effect of the position and momentum operators $\hat{x} = x$ and $\hat{p} \equiv \hat{p}_x = -i\hbar\frac{\partial}{\partial x}$ in the position basis. If $\hat{x}$ acts on $\psi(x)$, we obtain the eigenvalue x, while $\hat{p}$ is first characterized by the derivative. If we let $\hat{p}$ act first and then $\hat{x}$, it results in:

$$\hat{x}\hat{p}\psi(x) = -i\hbar x\frac{\partial}{\partial x}\psi(x)$$

The effect of $\hat{x}$ on the spatial derivative of $\psi(x)$ corresponds again to x. However, in the reverse order, we obtain:

$$\hat{p}\hat{x}\psi(x) = -i\hbar\frac{\partial}{\partial x}(x\psi(x)) = -i\hbar\left(\psi(x) + x\frac{\partial}{\partial x}\psi(x)\right) = -i\hbar\left(1 + x\frac{\partial}{\partial x}\right)\psi(x)$$

The results are obviously different. Calculating the difference $\hat{x}\hat{p} - \hat{p}\hat{x}$ and their effect on $\psi(x)$ yields the important relation:

$$(\hat{x}\hat{p} - \hat{p}\hat{x})\psi(x) = -i\hbar\left(x\frac{\partial}{\partial x} - 1 - x\frac{\partial}{\partial x}\right)\psi(x) = i\hbar\psi(x)$$

We will later use this to identify position and momentum operator as complementary operators. If we omit the state $\psi(x)$ on which the operators act in the notation, we obtain a so-called commutator relation:

$$\hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar \tag{3.117}$$

The fact that two operators $\hat{A}$ and $\hat{B}$ do not commute in general ($\hat{A}\hat{B} - \hat{B}\hat{A} \neq 0$) leads us in the following section to the commutator relationships of operators.

3.5.1 Commutator

That operators generally do not commute leads to the introduction of the *commutator*. This is defined over two operators $\hat{A}$ and $\hat{B}$:

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \tag{3.118}$$

If the commutator between two operators vanishes ($[\hat{A}, \hat{B}] = 0$), $\hat{A}$ and $\hat{B}$ fulfill the commutativity relation and they can henceforth be referred to as "commutative". This is equivalent to the statement that the two operators $\hat{A}$ and $\hat{B}$ commute. The commutator is an antisymmetric function, which can be easily shown by:

$$[\hat{B}, \hat{A}] = \hat{B}\hat{A} - \hat{A}\hat{B} = -(\hat{A}\hat{B} - \hat{B}\hat{A}) = -[\hat{A}, \hat{B}] \tag{3.119}$$

The commutator between two identical operators necessarily leads to the zero operator:

$$[\hat{A}, \hat{A}] = \hat{A}\hat{A} - \hat{A}\hat{A} = 0 \tag{3.120}$$

With a third operator $\hat{C}$ and the two scalars β and γ, it can also be shown that the commutator satisfies the distributive law:

$$\begin{aligned}[\hat{A}, \beta\hat{B} + \gamma\hat{C}] &= \beta\hat{A}\hat{B} + \gamma\hat{A}\hat{C} - \beta\hat{B}\hat{A} - \gamma\hat{C}\hat{A} = \\ &= \beta(\hat{A}\hat{B} - \hat{B}\hat{A}) + \gamma(\hat{A}\hat{C} - \hat{C}\hat{A}) = \\ &= \beta[\hat{A}, \hat{B}] + \gamma[\hat{A}, \hat{C}]\end{aligned} \tag{3.121}$$

Another useful property of distributivity arises in the commutator of $\hat{A}$ and the product of two operators $\hat{B}\hat{C}$:

$$\begin{aligned}[\hat{A}, \hat{B}\hat{C}] &= \hat{A}\hat{B}\hat{C} - \hat{B}\hat{C}\hat{A} = \\ &= (\hat{A}\hat{B} - \hat{B}\hat{A} + \hat{B}\hat{A})\hat{C} - \hat{B}(\hat{C}\hat{A} - \hat{A}\hat{C} + \hat{A}\hat{C}) = \\ &= ([\hat{A}, \hat{B}] + \hat{B}\hat{A})\hat{C} - \hat{B}([\hat{C}, \hat{A}] + \hat{A}\hat{C}) = \\ &= [\hat{A}, \hat{B}]\hat{C} + \hat{B}\hat{A}\hat{C} - \hat{B}[\hat{C}, \hat{A}] - \hat{B}\hat{A}\hat{C} = \\ &= \hat{B}[\hat{A}, \hat{C}] + [\hat{A}, \hat{B}]\hat{C} \end{aligned} \tag{3.122}$$

Commutators can also be nested and in the simplest case can even be cyclically permuted in this way – the *Jacobi identity* plays an important role here:

$$[\hat{A}, [\hat{B}, \hat{C}]] + [\hat{C}, [\hat{A}, \hat{B}]] + [\hat{B}, [\hat{C}, \hat{A}]] = 0 \tag{3.123}$$

Example: Proof of the Jacobi Identity

It should be shown that the Jacobi identity from (3.123) is fulfilled. All commutators are evaluated explicitly with (3.122), otherwise no special assumptions are made:

$$\begin{aligned}\text{(I)}|\quad [\hat{B}, [\hat{C}, \hat{A}]] &= [\hat{B}, \hat{C}\hat{A} - \hat{A}\hat{C}] = [\hat{B}, \hat{C}\hat{A}] - [\hat{B}, \hat{A}\hat{C}] \overset{(3.122)}{=} \\ &= \hat{C}[\hat{B}, \hat{A}] + [\hat{B}, \hat{C}]\hat{A} - \hat{A}[\hat{B}, \hat{C}] - [\hat{B}, \hat{A}]\hat{C} = \\ &= -([\hat{C}, \hat{B}]\hat{A} + \hat{A}[\hat{B}, \hat{C}] + [\hat{B}, \hat{A}]\hat{C} + \hat{C}[\hat{A}, \hat{B}]) \\ \text{(II)}|\quad [\hat{C}, [\hat{A}, \hat{B}]] &= [\hat{C}, \hat{A}\hat{B} - \hat{B}\hat{A}] = [\hat{C}, \hat{A}\hat{B}] - [\hat{C}, \hat{B}\hat{A}] \overset{(3.122)}{=} \\ &= \hat{A}[\hat{C}, \hat{B}] + [\hat{C}, \hat{A}]\hat{B} - \hat{B}[\hat{C}, \hat{A}] - [\hat{C}, \hat{B}]\hat{A} = \\ &= \hat{A}\hat{C}\hat{B} - \hat{A}\hat{B}\hat{C} + \hat{C}\hat{A}\hat{B} - \hat{A}\hat{C}\hat{B} - \hat{B}\hat{C}\hat{A} + \hat{B}\hat{A}\hat{C} - \hat{C}\hat{B}\hat{A} + \hat{B}\hat{C}\hat{A} = \\ &= -\hat{A}\hat{B}\hat{C} + \hat{C}\hat{A}\hat{B} + \hat{B}\hat{A}\hat{C} - \hat{C}\hat{B}\hat{A} = \\ &= [\hat{B}, \hat{A}]\hat{C} + \hat{C}[\hat{A}, \hat{B}] \\ \text{(III)}|\quad [\hat{A}, [\hat{B}, \hat{C}]] &= [\hat{A}, \hat{B}\hat{C} - \hat{C}\hat{B}] = [\hat{A}, \hat{B}\hat{C}] - [\hat{A}, \hat{C}\hat{B}] \overset{(3.122)}{=} \\ &= \hat{B}[\hat{A}, \hat{C}] + [\hat{A}, \hat{B}]\hat{C} - \hat{C}[\hat{A}, \hat{B}] - [\hat{A}, \hat{C}]\hat{B} = \\ &= \hat{B}\hat{A}\hat{C} - \hat{B}\hat{C}\hat{A} + \hat{A}\hat{B}\hat{C} - \hat{B}\hat{A}\hat{C} - \hat{C}\hat{A}\hat{B} + \hat{C}\hat{B}\hat{A} - \hat{A}\hat{C}\hat{B} + \hat{C}\hat{A}\hat{B} = \\ &= -\hat{B}\hat{C}\hat{A} + \hat{A}\hat{B}\hat{C} + \hat{C}\hat{B}\hat{A} - \hat{A}\hat{C}\hat{B} = \\ &= [\hat{C}, \hat{B}]\hat{A} + \hat{A}[\hat{B}, \hat{C}] \end{aligned}$$

From the results of the individual terms, it is quickly apparent that the Jacobi identity is fulfilled, as (I) is compensated by (II) and (III) and it must therefore hold that:

$$[\hat{A}, [\hat{B}, \hat{C}]] + [\hat{C}, [\hat{A}, \hat{B}]] + [\hat{B}, [\hat{C}, \hat{A}]] = 0$$

It is also possible to define a function called the *anticommutator*. It differs from the commutator by the internal sign:

$$\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A} \tag{3.124}$$

The notation varies in the literature; note that $\{\hat{A}, \hat{B}\}$ and $[\hat{A}, \hat{B}]_+$ both symbolize the anticommutator. Unlike the commutator, the anticommutator is a symmetric function:

$$\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A} = \hat{B}\hat{A} + \hat{A}\hat{B} = \{\hat{B}, \hat{A}\} \tag{3.125}$$

Together with the commutator relation, for an arbitrary product of two operators $\hat{A}$ and $\hat{B}$, the following decomposition can be found:

$$\hat{A}\hat{B} = \frac{1}{2}\left(2\hat{A}\hat{B} + \hat{B}\hat{A} - \hat{B}\hat{A}\right) = \frac{1}{2}\left(\hat{A}\hat{B} - \hat{B}\hat{A} + \hat{A}\hat{B} + \hat{B}\hat{A}\right) = \frac{1}{2}[\hat{A},\hat{B}] + \frac{1}{2}\{\hat{A},\hat{B}\} \tag{3.126}$$

Restricting ourselves again to Hermitian operators $\hat{A} = \hat{A}^\dagger$ and $\hat{B} = \hat{B}^\dagger$, we can find new relationships:

$$[\hat{A},\hat{B}]^\dagger = (\hat{A}\hat{B})^\dagger - (\hat{B}\hat{A})^\dagger = \hat{B}^\dagger\hat{A}^\dagger - \hat{A}^\dagger\hat{B}^\dagger = \hat{B}\hat{A} - \hat{A}\hat{B} = [\hat{B},\hat{A}] = -[\hat{A},\hat{B}] \tag{3.127}$$

By taking the adjoint of the commutator, we obtain the same commutator except for a sign. We refer to this property as *anti-Hermiticity* – unlike a Hermitian operator, an anti-Hermitian operator only has imaginary eigenvalues.

Proof: Imaginary Eigenvalues of an Anti-Hermitian Operator

Let an anti-Hermitian operator $\hat{C} = -\hat{C}^\dagger$ be given. As we have just seen, a commutator of two Hermitian operators $\hat{A}$ and $\hat{B}$ is anti-Hermitian, so we could define such an operator $\hat{C}$, for example, by the relationship $\hat{C} = [\hat{A},\hat{B}]$. The operator $\hat{C}$ is related to the state $|\phi\rangle$ in the eigenvalue relationship $\hat{C}\,|\phi\rangle = c\,|\phi\rangle$. The eigenvalue c can be calculated directly by:

$$c = \langle\phi|\hat{C}|\phi\rangle\,. \tag{3.128}$$

Conjugating and transposing equation (3.128) gives:

$$\begin{aligned} c^* &\overset{(3.128)}{=} \langle\psi|\hat{C}|\psi\rangle^\dagger = \langle\psi|\hat{C}^\dagger|\psi\rangle = \quad |\hat{C}^\dagger = -\hat{C} \text{ (anti-Hermitian)} \\ &= -\langle\psi|\hat{C}|\psi\rangle \overset{(3.128)}{=} -c \end{aligned} \tag{3.129}$$

For the eigenvalue c it follows that: $c^* = -c$. This is only fulfilled if c is a purely imaginary quantity. This can be verified by assuming that c is a general complex number $c = \alpha + i\beta$ with real part α and imaginary part β:

$$\begin{gathered} c^* \overset{!}{=} -c \\ (\alpha + i\beta)^* = -(\alpha + i\beta) \\ \alpha - i\beta = -\alpha - i\beta \\ \text{only possible if } \alpha = 0 \text{ and } c = i\beta \text{ is purely imaginary. } \square \end{gathered} \tag{3.130}$$

It can also be easily shown that the anticommutator is Hermitian:

$$\{\hat{A},\hat{B}\}^\dagger = (\hat{A}\hat{B})^\dagger + (\hat{B}\hat{A})^\dagger = \hat{B}^\dagger\hat{A}^\dagger + \hat{A}^\dagger\hat{B}^\dagger = \hat{B}\hat{A} + \hat{A}\hat{B} = \{\hat{B},\hat{A}\} = \{\hat{A},\hat{B}\} \tag{3.131}$$

This implies that we only obtain real eigenvalues for the anticommutator. We have already provided the corresponding proof in (3.32–3.34).

3.5.2 Complementarity, Canonical Commutation Relation, and Uncertainty Principle

The commutator between the position and momentum operator $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$ yields an important relationship, referred to as the *canonical commutation relation*. Already in (3.117) the canonical commutation relation $[\hat{x},\hat{p}] = i\hbar$ in one-dimensional position space was shown. This can be generalized to three dimensions:

$$[\hat{\mathbf{r}},\hat{\mathbf{p}}] = i\hbar \tag{3.132}$$

If the commutator contains operators that act along the direction of the basis vectors of an orthonormal system (for example, $\hat{r}_1 = \hat{x}$, $\hat{r}_2 = \hat{y}$, $\hat{r}_3 = \hat{z}$ or $\hat{p}_1 = \hat{p}_x$, $\hat{p}_2 = \hat{p}_y$, $\hat{p}_3 = \hat{p}_z$), then the expression in (3.132) can be further generalized to:

$$[\hat{r}_i, \hat{p}_j] = i\hbar\delta_{ij} \tag{3.133}$$

Operators $\hat{A}$ and $\hat{B}$ that fulfill the relation (3.133) are referred to as *complementary operators*:

$$[\hat{A}, \hat{B}] = i\hbar \tag{3.134}$$

Here the statement in (3.127), that the commutator carries purely imaginary eigenvalues, can be confirmed, as the effect is obviously characterized by the imaginary quantity $i\hbar$. A similar relationship can also be found in classical mechanics, where instead of the commutator, the Poisson bracket is used.

Motivation: Poisson Brackets in Classical Mechanics

Let an arbitrary observable of the form $f(q_k, p_k, t)$ be given with $k = 1, \ldots, N$ degrees of freedom, where q_k describe the generalized position and p_k the generalized momentum. The time derivatives of these quantities lead to Hamilton's equations of motion, which are introduced in (10.2) and have the following form:

$$\dot{q}_k = \frac{\partial H}{\partial p_k} \quad \text{and} \quad \dot{p}_k = -\frac{\partial H}{\partial q_k} \quad \text{with} \quad k = 1, \ldots, N$$

These relationships can be used to calculate the total time differential of $f(q_k, p_k, t)$. It holds:

$$\begin{aligned} \frac{df}{dt} &= \sum_{k=1}^{N} \left(\frac{\partial f}{\partial q_k}\frac{\partial q_k}{\partial t} + \frac{\partial f}{\partial p_k}\frac{\partial p_k}{\partial t} \right) + \frac{\partial f}{\partial t} = \\ &= \sum_{k=1}^{N} \left(\frac{\partial f}{\partial q_k}\frac{\partial H}{\partial p_k} - \frac{\partial f}{\partial p_k}\frac{\partial H}{\partial q_k} \right) + \frac{\partial f}{\partial t} = \\ &= \{f, H\} + \frac{\partial f}{\partial t} \end{aligned} \tag{3.135}$$

We denote $\{f, H\}$ as the *Poisson bracket* between the observable and the Hamiltonian. If we additionally assume that $f(q_k, p_k, t) \equiv f(q_k, p_k)$ does not explicitly depend on time, the last term $\frac{\partial f}{\partial t}$ vanishes. If it further holds that the observable is a conserved quantity, the total differential must also vanish, leading to: $\{f, H\} = 0$. If this relationship is fulfilled for any $f(q_k, p_k)$, we can, by reverse, conclude that it is a conserved quantity.

The Poisson bracket can generally be formed between any quantities. We now specifically consider the generalized position and momentum and find the following relationship:

$$\{q_i, p_l\} = \sum_{k=1}^{N} \left(\frac{\partial q_i}{\partial q_k}\frac{\partial p_l}{\partial p_k} - \frac{\partial q_i}{\partial p_k}\frac{\partial p_l}{\partial q_k} \right) = \sum_{k=1}^{N} \delta_{ik}\delta_{lk} = \delta_{il} \tag{3.136}$$

We thus obtain a similar form to the canonical commutation relation from (3.133). Furthermore, the analogy also holds for Poisson brackets between identical quantities, which can be easily compared in the case of commutators with (3.120):

$$\{q_i, q_l\} = \{p_i, p_l\} = 0 \tag{3.137}$$

Let us examine the commutator relationship from (3.134) more closely and restrict ourselves to two Hermitian operators $\hat{A} = \hat{A}^\dagger$ and $\hat{B} = \hat{B}^\dagger$, then the Heisenberg uncertainty relation shown in (2.80) can now be derived for any operators $\hat{A}$ and $\hat{B}$. First, we want to introduce a fluctuation

operator $\Delta\hat{A}$, which describes the deviation of the effect from the expected value (mean of many measurements):

$$\Delta\hat{A} = \hat{A} - \langle\hat{A}\rangle \tag{3.138}$$

$\langle\hat{A}\rangle = \langle\psi|\hat{A}|\psi\rangle$ is the expected value of $\hat{A}$ for a given, normalized wave function $|\psi\rangle$. The spread (or variance) can be calculated as the expected value of the squared fluctuation operator $(\Delta\hat{A})^2$; the standard deviation is described by the square root of the variance:

$$\sigma_A = \sqrt{\langle(\Delta\hat{A})^2\rangle} \tag{3.139}$$

If we now want to obtain the expected value of the variance $(\Delta\hat{A})^2$, we can use that the following general calculation rules apply to expectation values: $\langle\hat{X}+\hat{Y}\rangle = \langle\hat{X}\rangle + \langle\hat{Y}\rangle$ and $\langle\hat{X}\hat{Y}\rangle = \langle\hat{X}\rangle\langle\hat{Y}\rangle$. Furthermore, for the expectation value of an expectation value, it holds: $\langle\langle\hat{X}\rangle\rangle = \langle\hat{X}\rangle$.

$$\begin{aligned}
\langle(\Delta\hat{A})^2\rangle &\overset{(3.138)}{=} \langle(\hat{A}-\langle\hat{A}\rangle)^2\rangle = \langle(\hat{A}-\langle\hat{A}\rangle)(\hat{A}-\langle\hat{A}\rangle)\rangle = \langle\hat{A}^2 - \langle\hat{A}\rangle\hat{A} - \hat{A}\langle\hat{A}\rangle + \langle\hat{A}\rangle^2\rangle = \\
&= \langle\hat{A}^2 - 2\langle\hat{A}\rangle\hat{A} + \langle A\rangle^2\rangle = \langle\hat{A}^2\rangle - \langle 2\langle\hat{A}\rangle\hat{A}\rangle + \langle\langle\hat{A}\rangle^2\rangle = \\
&= \langle\hat{A}^2\rangle - \langle 2\rangle\langle\langle\hat{A}\rangle\rangle\langle\hat{A}\rangle + \langle\hat{A}\rangle^2 = \langle\hat{A}^2\rangle - 2\langle\hat{A}\rangle\langle\hat{A}\rangle + \langle\hat{A}\rangle^2 = \\
&= \langle\hat{A}^2\rangle - 2\langle\hat{A}\rangle^2 + \langle\hat{A}\rangle^2 = \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2
\end{aligned} \tag{3.140}$$

$\langle\hat{A}^2\rangle$ is called the second moment of $\hat{A}$. Generally, $\langle\hat{A}^x\rangle$ is called the x-th moment of $\hat{A}$, so $\langle\hat{A}\rangle = \langle\hat{A}^1\rangle$ is the first moment. Without loss of generality, let us now assume an expected value of 0 for simplification; namely $\langle\hat{A}\rangle = \langle\hat{B}\rangle = 0$. Then the variance $\langle(\Delta\hat{A})^2\rangle$ is:

$$\langle(\Delta\hat{A})^2\rangle = \langle\psi|\hat{A}^2|\psi\rangle = \underbrace{\langle\hat{A}\psi|\hat{A}\psi\rangle}_{\substack{\text{Square of} \\ \text{length of } \hat{A}|\psi\rangle}} = ||\hat{A}|\psi\rangle||^2 \tag{3.141}$$

Similar calculation steps can also be carried out for an operator $\hat{B}$, so that we obtain an expected variance of $\langle(\Delta\hat{B})^2\rangle = ||\hat{B}|\psi\rangle||^2$. Let us now try to estimate the product of the variances $\langle(\Delta\hat{A})^2\rangle$ and $\langle(\Delta\hat{B})^2\rangle$:

$$\begin{aligned}
\langle(\Delta\hat{A})^2\rangle\langle(\Delta\hat{B})^2\rangle &= ||\hat{A}|\psi\rangle||^2\,||\hat{B}|\psi\rangle||^2 \overset{(3.18)}{\geq} \\
&\geq |\langle\hat{A}\psi|\hat{B}\psi\rangle|^2 = |\langle\psi|\hat{A}\hat{B}|\psi\rangle|^2 \overset{(3.126)}{=} \\
&= \left|\left\langle\psi\right|\frac{1}{2}[\hat{A},\hat{B}] + \frac{1}{2}\{\hat{A},\hat{B}\}\left|\psi\right\rangle\right|^2 = \\
&= \frac{1}{4}\Big|\langle\psi|\underbrace{[\hat{A},\hat{B}]}_{\text{anti-Herm.}}|\psi\rangle + \langle\psi|\underbrace{\{\hat{A},\hat{B}\}}_{\text{Hermitian}}|\psi\rangle\Big|^2 = \ |\ \alpha = \langle[\hat{A},\hat{B}]\rangle\,,\ \beta = \langle\{\hat{A},\hat{B}\}\rangle \\
&= \frac{1}{4}\left|\alpha + i\beta\right|^2 = \frac{1}{4}\left(\alpha^2+\beta^2\right) \geq \\
&\geq \frac{1}{4}\,\beta^2 = \frac{1}{4}\left|\langle\psi|\,[\hat{A},\hat{B}]\,|\psi\rangle\right|^2
\end{aligned} \tag{3.142}$$

In the first line of the above derivation, we used the squared Cauchy-Schwarz inequality from (3.18). In the fourth line, we use that the commutator of Hermitian operators is anti-Hermitian, and thus according to (3.130) only provides imaginary eigenvalues, while the anticommutator of Hermitian operators is itself a Hermitian quantity and thus according to (3.131) only provides real eigenvalues. Taking the square root of the expression, we obtain the uncertainty relation for initially any Hermitian operators $\hat{A}$ and $\hat{B}$:

$$\sigma_A \sigma_B \geq \frac{1}{2} \left| \langle\psi| \left[\hat{A}, \hat{B}\right] |\psi\rangle \right| \tag{3.143}$$

In the case of operators that represent complementary observables, the relation $[\hat{A}, \hat{B}] = i\hbar$ holds. We can substitute this into (3.143), and we get:

$$\langle(\Delta\hat{A})^2\rangle \langle(\Delta\hat{B})^2\rangle \geq \frac{1}{4} |\langle\psi|i\hbar|\psi\rangle|^2 = \frac{1}{4} |i\hbar\langle\psi|\psi\rangle|^2 = \frac{\hbar^2}{4} \tag{3.144}$$

By taking the square root on both sides, we finally obtain the Heisenberg uncertainty relation for complementary observables:

$$[\hat{A}, \hat{B}] = i\hbar \implies \sigma_A \sigma_B \geq \frac{\hbar}{2} \tag{3.145}$$

3.5.3 Compatibility

Another special relationship that two operators $\hat{A}$ and $\hat{B}$ can satisfy in the commutator is:

$$[\hat{A}, \hat{B}] = 0 \iff \hat{A}\hat{B} = \hat{B}\hat{A} \tag{3.146}$$

In this case, $\hat{A}$ and $\hat{B}$ are interchangeable and are therefore called *compatible operators*. Operators that satisfy these relations possess a common, complete orthonormal system, and this statement is also valid in the reverse direction (a common orthonormal system leads to interchangeable operators). Compatible operators play an outstanding role in quantum physics and should therefore be discussed in more detail.

Before we turn to the proof that compatible operators and a common orthonormal system mutually imply each other, a brief excursion into the notation using an example: Suppose we have the compatible operators $\hat{A}$ and $\hat{B}$ and a common orthonormal system with the basis vectors $\{|\phi_1\rangle, |\phi_2\rangle, |\phi_3\rangle, |\phi_4\rangle, |\phi_5\rangle, |\phi_6\rangle\}$. If we let the operator $\hat{A}$ or $\hat{B}$ act separately on each of the basis vectors, we get the following results in our example:

$$\begin{aligned}
\hat{A}|\phi_1\rangle &= a_1|\phi_1\rangle & \hat{B}|\phi_1\rangle &= b_1|\phi_1\rangle \\
\hat{A}|\phi_2\rangle &= a_2|\phi_2\rangle & \hat{B}|\phi_2\rangle &= b_1|\phi_2\rangle \\
\hat{A}|\phi_3\rangle &= a_2|\phi_3\rangle & \hat{B}|\phi_3\rangle &= b_2|\phi_3\rangle \\
\hat{A}|\phi_4\rangle &= a_3|\phi_4\rangle & \hat{B}|\phi_4\rangle &= b_1|\phi_4\rangle \\
\hat{A}|\phi_5\rangle &= a_3|\phi_5\rangle & \hat{B}|\phi_5\rangle &= b_2|\phi_5\rangle \\
\hat{A}|\phi_6\rangle &= a_3|\phi_6\rangle & \hat{B}|\phi_6\rangle &= b_3|\phi_6\rangle
\end{aligned}$$

Most of the eigenvalues of both operators are thus degenerate. Neither an eigenvalue a_i, nor an eigenvalue b_j uniquely identifies an eigenvector, but the combination of a_i and b_j does. Because the operators $\hat{A}$ and $\hat{B}$ are compatible (commute), we can let them act consecutively in any order:

$$\begin{aligned}
\hat{A}\hat{B}|\phi_1\rangle &= \hat{B}\hat{A}|\phi_1\rangle = a_1 b_1 |\phi_1\rangle \\
\hat{A}\hat{B}|\phi_2\rangle &= \hat{B}\hat{A}|\phi_2\rangle = a_2 b_1 |\phi_2\rangle \\
\hat{A}\hat{B}|\phi_3\rangle &= \hat{B}\hat{A}|\phi_3\rangle = a_2 b_2 |\phi_3\rangle \\
\hat{A}\hat{B}|\phi_4\rangle &= \hat{B}\hat{A}|\phi_4\rangle = a_3 b_1 |\phi_4\rangle \\
\hat{A}\hat{B}|\phi_5\rangle &= \hat{B}\hat{A}|\phi_5\rangle = a_3 b_2 |\phi_5\rangle \\
\hat{A}\hat{B}|\phi_6\rangle &= \hat{B}\hat{A}|\phi_6\rangle = a_3 b_3 |\phi_6\rangle
\end{aligned}$$

Since certain eigenvalues a_i and b_j are characteristic for each basis vector, we now want to introduce the following notation:

$$|\phi_1\rangle \equiv |a_1 b_1\rangle\,, \quad |\phi_2\rangle \equiv |a_2 b_1\rangle\,, \quad |\phi_3\rangle \equiv |a_2 b_2\rangle\,, \quad |\phi_4\rangle \equiv |a_3 b_1\rangle\,, \quad |\phi_5\rangle \equiv |a_3 b_2\rangle\,, \quad |\phi_6\rangle \equiv |a_3 b_3\rangle$$

That means, we now generally refer to a basis vector no longer as $|\phi_i\rangle$, but as $|a_i b_j\rangle$. By this, we mean that the following eigenvalue equations should be satisfied for such a basis vector:

$$\hat{A}\,|a_i b_j\rangle = a_i\,|a_i b_j\rangle \quad \text{and} \quad \hat{B}\,|a_i b_j\rangle = b_j\,|a_i b_j\rangle \tag{3.147}$$

Just as in our example, the following generally applies to compatible operators: If we let both operators act consecutively on $|a_i b_j\rangle$, the eigenstate does not change (because of the compatibility of the operators):

$$\begin{aligned} \hat{A}\hat{B}\,|a_i b_j\rangle &= \hat{A} b_j\,|a_i b_j\rangle = b_j \hat{A}\,|a_i b_j\rangle = b_j a_i\,|a_i b_j\rangle = a_i b_j\,|a_i b_j\rangle \\ \hat{B}\hat{A}\,|a_i b_j\rangle &= \hat{B} a_i\,|a_i b_j\rangle = a_i \hat{B}\,|a_i b_j\rangle = a_i b_j\,|a_i b_j\rangle \end{aligned} \tag{3.148}$$

We must only be careful with the summation indices in this notation when we want to represent a sum over basis vectors. For example, we would simply write a superposition over all basis vectors $\{|\phi_i\rangle\}$ like this:

$$|\Psi\rangle = c_1\,|\phi_1\rangle + c_2\,|\phi_2\rangle + c_3\,|\phi_3\rangle + c_4\,|\phi_4\rangle + c_5\,|\phi_5\rangle + c_6\,|\phi_6\rangle = \sum_{i=1}^{6} c_i\,|\phi_i\rangle$$

In the newly introduced notation for a common orthonormal system $\{|a_i b_j\rangle\}$, we must instead write this sum as follows:

$$|\Psi\rangle = c_{11}\,|a_1 b_1\rangle + c_{21}\,|a_2 b_1\rangle + c_{22}\,|a_2 b_2\rangle + c_{31}\,|a_3 b_1\rangle + c_{32}\,|a_3 b_2\rangle + c_{33}\,|a_3 b_3\rangle = \sum_{i=1}^{3} \sum_{j=1}^{g_a(i)} c_{ij}\,|a_i b_j\rangle$$

Here, $g_a(i)$ is the degeneracy of the eigenvalue a_i. Having clarified this, we want to prove the statement

$$[\hat{A}, \hat{B}] = 0 \iff \exists \text{ common orthonormal system}$$

first "from right to left." We want to show: "If the operators $\hat{A}$ and $\hat{B}$ have a common, complete orthonormal system (CONS), then $\hat{A}$ and $\hat{B}$ are compatible." To do this, we write $\hat{A}$ and $\hat{B}$ in their spectral representation according to (3.147):

$$\hat{A} = \sum_i \sum_j^{g_a(i)} |a_i b_j\rangle\, a_i\, \langle a_i b_j| \quad \text{and} \quad \hat{B} = \sum_i \sum_j^{g_a(i)} |a_i b_j\rangle\, b_j\, \langle a_i b_j| \tag{3.149}$$

We now want to specify the products $\hat{A}\hat{B}$ and $\hat{B}\hat{A}$ in this common eigenbasis in the spectral representation according to (3.149):

$$\begin{aligned} \hat{A}\hat{B} &\overset{(3.149)}{=} \left[\sum_i \sum_j^{g_a(i)} |a_i b_j\rangle\, a_i\, \langle a_i b_j|\right] \left[\sum_k \sum_j^{g_a(l)} |a_k b_l\rangle\, b_l\, \langle a_k b_l|\right] = \\ &= \sum_i \sum_j^{g_a(i)} \sum_k \sum_j^{g_a(l)} |a_i b_j\rangle\, a_i \underbrace{\langle a_i b_j | a_k b_l\rangle}_{\delta_{ik}\delta_{jl}} b_l\, \langle a_k b_l| = \sum_i \sum_j^{g_a(i)} |a_i b_j\rangle\, a_i b_j\, \langle a_i b_j| \end{aligned} \tag{3.150}$$

$$\begin{aligned} \hat{B}\hat{A} &\overset{(3.149)}{=} \left[\sum_i \sum_j^{g_a(i)} |a_i b_j\rangle\, b_j\, \langle a_i b_j|\right] \left[\sum_k \sum_j^{g_a(l)} |a_k b_l\rangle\, a_k\, \langle a_k b_l|\right] = \\ &= \sum_i \sum_j^{g_a(i)} \sum_k \sum_j^{g_a(l)} |a_i b_j\rangle\, b_l \underbrace{\langle a_i b_j | a_k b_l\rangle}_{\delta_{ik}\delta_{jl}} a_k\, \langle a_k b_l| = \sum_i \sum_j^{g_a(i)} |a_i b_j\rangle\, a_i b_j\, \langle a_i b_j| \end{aligned} \tag{3.151}$$

Comparing (3.150) and (3.151), we see: If we assume that there is a common orthonormal system for the operators $\hat{A}$ and $\hat{B}$, then there is no difference between $\hat{A}\hat{B}$ and $\hat{B}\hat{A}$, and therefore: $[\hat{A}, \hat{B}] = 0$.
The proof should now also be carried out in "forward direction". We want to prove: "If $\hat{A}$ and $\hat{B}$ are compatible (commute), then a common, complete orthonormal basis can be found". So, we initially assume only that there is an ONBS $\{|a_i\rangle\}$ for the operator $\hat{A}$ and independently an ONBS $\{|b_j\rangle\}$ for the operator $\hat{B}$, such that the following eigenvalue relations hold:

$$\hat{A}\,|a_i\rangle = a_i\,|a_i\rangle \quad \text{and} \quad \hat{B}\,|b_j\rangle = b_j\,|b_j\rangle \tag{3.152}$$

For the time being, we do not make any claim about whether these two basis systems are identical or not. Let's now consider, starting from the eigenvalue equation for $\hat{A}$, the following simple calculation:

$$\begin{aligned}
\hat{A}\,|a_i\rangle &= a_i\,|a_i\rangle && \mid \text{let } \hat{B} \text{ act from the left} \\
\hat{B}\hat{A}\,|a_i\rangle &= \hat{B}a_i\,|a_i\rangle && \mid \hat{B}\hat{A} = \hat{A}\hat{B},\ \hat{B}a_i = a_i\hat{B} \\
\hat{A}\hat{B}\,|a_i\rangle &= a_i\hat{B}\,|a_i\rangle && \Longrightarrow \\
\hat{A}(\hat{B}\,|a_i\rangle) &= a_i(\hat{B}\,|a_i\rangle)
\end{aligned} \tag{3.153}$$

In the last line, we have added (actually unnecessary) parentheses to clarify the following: Clearly, not only is $|a_i\rangle$ an eigenstate of $\hat{A}$ for the eigenvalue a_i, but also $\hat{B}\,|a_i\rangle$ is an eigenstate of $\hat{A}$, and indeed for the same eigenvalue a_i! We must now discuss two cases:

- ***a_i* is non-degenerate**: If a_i is non-degenerate, there is a unique association between the eigenvalue a_i and the eigenvector $|a_i\rangle$. Therefore, if $\hat{B}\,|a_i\rangle$ is *also* an eigenvector of $\hat{A}$ with the eigenvalue a_i, this is only possible if $\hat{B}\,|a_i\rangle$ represents a (complex-valued) stretched or contracted version of $|a_i\rangle$, so that:

  $$\hat{B}\,|a_i\rangle = c\,|a_i\rangle \tag{3.154}$$

  However, this is a new eigenvalue equation showing that $|a_i\rangle$ is thus also an eigenstate of the operator $\hat{B}$ for the eigenvalue c! The state $|a_i\rangle$ is therefore a *common* eigenvector of both $\hat{A}$ and $\hat{B}$ (which was to be proven).

- ***a_i* is degenerate**: If a_i is degenerate, there are two or more eigenvectors for the same eigenvalue a_i. The association between the eigenvalue a_i and the corresponding eigenvector is no longer unique. In other words: There are multiple linearly independent eigenstates, all lying in the degenerate subspace for the eigenvalue a_i. Therefore, the earlier argumentation no longer works. The vector $\hat{B}\,|a_i\rangle$ indeed lies in this degenerate subspace, but it is no longer necessarily just a stretched or contracted version of $|a_i\rangle$.

  Thus, we need a different line of proof. First, let's represent the eigenvector $|a_i\rangle$, which is part of the degenerate subspace, in the eigenstates of $\hat{B}$:

  $$|a_i\rangle = \sum_j \beta_j\,|b_j\rangle \tag{3.155}$$

  Here, β_j are the expansion coefficients. Let's consider the eigenvalue equation (3.153) again:

  $$\hat{A}(\hat{B}\,|a_i\rangle) = a_i(\hat{B}\,|a_i\rangle)$$

Simple transformations now allow us to achieve a clearer picture:

$$\begin{aligned}
0 &= \hat{A}(\hat{B}\,|a_i\rangle) - a_i(\hat{B}\,|a_i\rangle) \\
&= (\hat{A} - a_i)(\hat{B}\,|a_i\rangle) \overset{(3.155)}{=} \\
&= (\hat{A} - a_i)\hat{B}\sum_j \beta_j\,|b_j\rangle = \\
&= (\hat{A} - a_i)\sum_j \beta_j \hat{B}\,|b_j\rangle = \quad |\ \hat{B}\,|b_j\rangle = b_j\,|b_j\rangle \\
&= (\hat{A} - a_i)\sum_j \beta_j b_j\,|b_j\rangle \\
&= \sum_j \beta_j \underbrace{b_j(\hat{A} - a_i)\,|b_j\rangle}_{|b'_j\rangle}
\end{aligned}$$

The superposition of the vectors $|b'_j\rangle$ results in a null vector, even though all $|b_j\rangle$ are linearly independent. To fulfill this equation, the following must hold true: For every value of the summation index j, either the expansion coefficient β_j or the vector $|b'_j\rangle$ must vanish. However, we can be certain that *not all* β_j vanish since in such a case, due to (3.155), $|a_i\rangle = 0$ would result, which is certainly not true. Thus, there are terms in the sum in which (for certain values of j) β_j does *not* vanish. In these terms, the vector $|b'_j\rangle$ must vanish. This leads us directly to the desired result:

$$|b_j\rangle = 0 \implies (\hat{A} - a_i)\,|b_j\rangle = 0 \implies \hat{A}\,|b_j\rangle = a_i\,|b_j\rangle \implies \hat{A}\,|b_j\rangle = a_i\,|b_j\rangle$$

Thus, we have found an eigenvalue equation showing that even in the degenerate subspace, the basis vector $|b_j\rangle$ of the operator $\hat{B}$ is simultaneously an eigenvector of the operator $\hat{A}$. The assertion that for compatible operators a common, complete orthonormal system can be found is thus confirmed in full generality.

These considerations, which at first glance are very theoretical, are of great practical relevance. When one solves the Schrödinger equation for a specific problem (e.g., the hydrogen atom), one generally tries to find a complete orthonormal system that can describe a maximum number of mutually compatible observables:

$$[\hat{H}, \hat{A}_i] = 0 \quad \text{and} \quad [\hat{A}_i, \hat{A}_j] = 0,\ \forall i, j$$

The maximum number of compatible observables determines the maximum amount of information that can be determined in a quantum measurement in accordance with the Heisenberg uncertainty principle. All observables that commute with each other can also be measured *precisely* simultaneously!

3.6 Measurements in Quantum Theory and Collapse of the Wave Function

Motivation: The Mystery of Measurement in Quantum Mechanics

We have already shown what happens when a measurement is performed on a system that is in an eigenstate $|a_i\rangle$ of the operator $\hat{A}$ associated with the measurement. It is rather unspectacular: every time, we measure the eigenvalue associated with the eigenstate with absolute certainty. But what if the system is *not* in an eigenstate of the operator? Something strange happens.

Let's look at a concrete example: Suppose we measure the energy of a state $|\psi\rangle$. The operator associated with energy measurement is the Hamiltonian operator $\hat{H}$. The eigenvalues (eigenenergies) of our operator are $\{E_1, E_2, \ldots, E_N\}$ and the corresponding energy eigenstates are denoted as $\{|e_1\rangle, |e_2\rangle, \ldots, |e_N\rangle\}$. The state to be measured $|\psi\rangle$ is now a superposition of the first two energy eigenstates: $|\psi\rangle = c_1 |e_1\rangle + c_2 |e_2\rangle$. Classically, one would probably expect to measure a total energy E in each measurement, which somehow lies between the energies E_1 and E_2. After all, the state $|\psi\rangle$ is a "mixture" of the corresponding states $|e_1\rangle$ and $|e_2\rangle$.

But this is not the case! In each individual measurement on $|\psi\rangle$, the measuring device only shows either the energy E_1 or the energy E_2! The fact that the state $|\psi\rangle$ is a superposition of the energy eigenstates E_1 and E_2 is only revealed when numerous measurements are performed: only after averaging over many measurement results does an *average* measurement result emerge, which approaches the quantum mechanically calculable expectation value. One can imagine it as if the system has not yet decided, as long as we have not measured, which state it is "really" in. Only our measurement creates this reality.

But that's not all: with the measurement, we have irreversibly destroyed the original state $|\psi\rangle$! If we (randomly) obtained the measurement value E_1 in a measurement process, the system is no longer in the state $|\psi\rangle$ after the measurement, but in the state $|e_1\rangle$. If we happened to obtain the measurement value E_2, the system is in the state $|e_2\rangle$ after the measurement. Our quantum system, which was in a superposition of two eigenstates before the measurement, is forced by the measurement to (randomly) decide on one of the eigenstates. And our calculations cannot predict which of the two variants will occur. All that quantum mechanics can do in this case is predict the *probability* that our measuring device will display E_1 or E_2 during a measurement. This reduction of the wave function is called the *collapse of the wave function* during the measurement. It is not something that automatically follows from the calculation, but must be inserted into the calculation separately. The collapsed wave function must also be renormalized.

Finally, this example also shows the following: In quantum mechanics, measurement and preparation of an experiment cannot be distinguished. What we have presented above as a measurement can also be used to prepare a state. If we want to prepare the state $|e_2\rangle$ in our example, we simply measure the state $|\psi\rangle$ until our measuring device displays the measurement value E_2, and then use only this state as the prepared state!

It is important to understand the following: If we start from an operator $\hat{A}$ that corresponds to a certain physical observable, and also from a quantum system in a general state $|\psi\rangle$, then quantum mechanics can generally *not* predict the result of a single measurement. However, it is possible to calculate the expectation value of the measurement value to which the average of many measurements on the state $|\psi\rangle$ will converge. The expectation value of the observable associated with the operator $\hat{A}$ is usually denoted by $\langle\hat{A}\rangle$. Note: Indicating $\langle\hat{A}\rangle$ by itself only

makes sense if it is clear to which state $|\psi\rangle$ this value relates! The expectation value $\langle\hat{A}\rangle$ can be calculated as follows:

$$\langle\hat{A}\rangle = \langle\psi|\hat{A}|\psi\rangle \tag{3.156}$$

To prove that $\langle\hat{A}\rangle$ is a real quantity, one can use the same proof as in (3.32–3.34), replacing $|a_i\rangle$ with $|\psi\rangle$, and $\langle a_i|$ with $\langle\psi|$.

In a measurement on a quantum system, only eigenvalues a_i of the operator $\hat{A}$ associated with the observable are ever measured, even if the system is in a state that can be represented as a superposition of these eigenstates. The eigenvalues thus describe possible measurement values of the respective observables, and must therefore be real. The operator $\hat{A}$ associated with the observable is therefore Hermitian: $\hat{A} = \hat{A}^\dagger$ or $a_i \in \mathbb{R}$. We assume that there is a bound state and thus a discrete spectrum of $\hat{A}$:

$$\hat{A}\,|\phi_i\rangle = a_i\,|\phi_i\rangle$$

Now let the state of the system be measured at time t_0. Again, we must distinguish between the two cases of a degenerate and non-degenerate system. For simplicity, we will first examine the non-degenerate case.

3.6.1 Measurement on a Non-Degenerate System

We assume that before the measurement, there is a general wave function $|\psi(t)\rangle$, which we can represent as a *coherent superposition* of the eigenstates of any operator. For practical purposes, we choose the eigensystem of the measurement operator $\hat{A}$:

$$|\psi(t)\rangle = \sum_i |\phi_i\rangle\,\langle\phi_i|\psi(t)\rangle = \sum_i c_i(t)\,|\phi_i\rangle \tag{3.157}$$

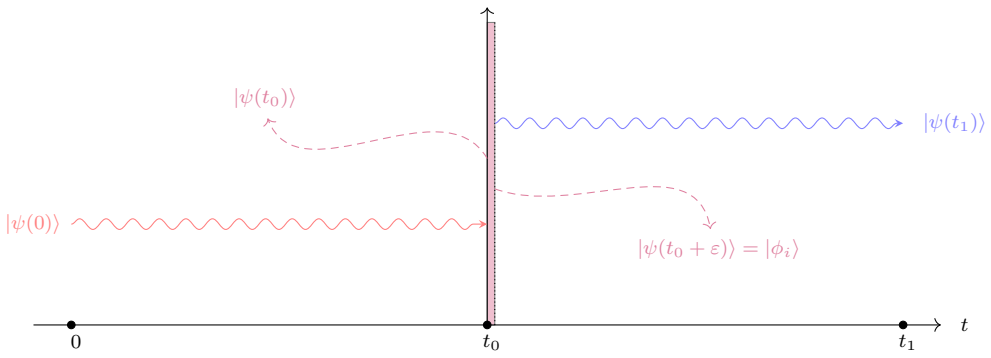


Fig. 15: A measurement of the observable $\hat{A}$ is performed on a system in the state $|\psi(t)\rangle$ at time t_0. Owing to the "intervention" of the measurement process, $|\psi(t)\rangle$ collapses into an eigenstate $|\phi_i\rangle$ of the measurement operator. The Schrödinger equation describes the propagation of the state up to time $t = t_0$. After the measurement, the "collapsed" state can continue to propagate.

At the time of measurement t_0, this state of the superposition "collapses." The measurement results in the *collapse of the wave function* to an eigenstate $|\phi_i\rangle$ of the measurement operator. Therefore, it holds:

$$|\psi(t_0)\rangle = |\phi_i\rangle$$

In general, it is not possible to predict to which eigenstate $|\phi_i\rangle$ the wave function will collapse in a single experiment! One can always only calculate the *probability* for a particular eigenstate $|\phi_i\rangle$ for a given experiment.

If the measurement operator is non-degenerate, each possible eigenstate of $\hat{A}$ is associated with a unique eigenvalue (i.e., measurement value), namely $a_i = \langle\phi_i|\hat{A}|\phi_i\rangle = \langle\psi(t_0)|\hat{A}|\psi(t_0)\rangle$. There is *no* uncertainty immediately after the measurement; the state after the collapse of the wave function can be uniquely determined by the eigenvalue a_i. If the measurement process lasts a short duration $\Delta t = \varepsilon$, we can continue to assume after this short duration:

$$|\psi(t_0 + \varepsilon)\rangle = |\phi_i\rangle$$

However, the state is changed again by the influence of the system after the measurement. Mathematically speaking, the Schrödinger equation, according to its Hamiltonian operator $\hat{H}$, dictates the time evolution of the state.
For conducting a quantum-mechanically significant measurement experiment, it is necessary that we can perform a large number of measurements on the same initial state $|\psi(t)\rangle$. This is possible by repeatedly performing the same experiment (with the same preparation of the initial state). This allows an interpretation of the measured observables in a statistical sense. If we were to measure only once, we could claim that our superposed state from (3.157) was already in the measured state before the measurement! Only repeated measurements allow for the experimental determination of the probability p_i, that the general state $|\psi(t)\rangle$ collapses into the eigenstate $|\phi_i\rangle$ at the time of measurement t_0. The probability is given by:

$$p_i = |\langle\phi_i|\psi\rangle|^2 = |c_i(t)|^2 \tag{3.158}$$

Measuring once is therefore generally not meaningful. Only repeated measurements enable the application of statistical methods so that we can, for example, assign a higher probability to measurement values that occur more frequently than those observed less frequently.

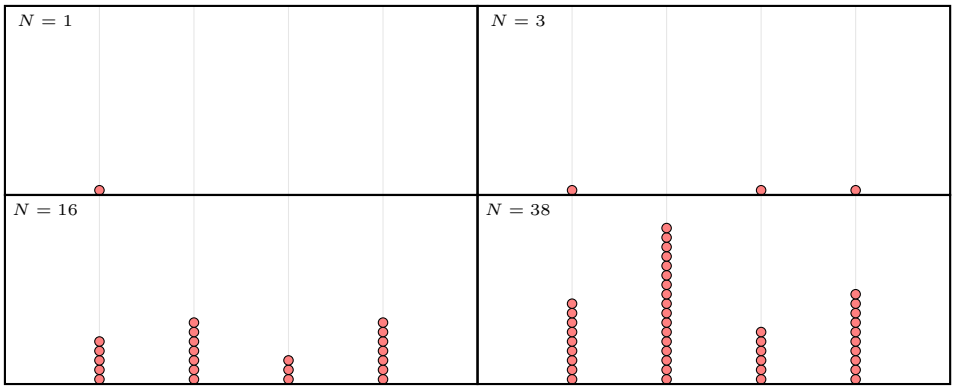


Fig. 16: An ensemble of similar states is measured: Each measurement of an observable leads to the collapse of the wave function and the identification of the state. Only repeated execution of the same measurement (by multiple repetitions of the experiment) allows for a statistical interpretation of the measurement.

If we know the probabilities with which a state $|\phi_i\rangle$ will occur during a measurement, we can calculate the *expectation value* of the observable $\hat{A}$:

$$\begin{aligned}\langle\psi|\hat{A}|\psi\rangle &= \langle\psi|\mathbb{1}\hat{A}\mathbb{1}|\psi\rangle = \\ &= \sum_{ij} \langle\psi|\phi_i\rangle \langle\phi_i|\hat{A}|\phi_j\rangle \langle\phi_j|\psi\rangle = \\ &= \sum_{ij} c_i^* c_j a_j \langle\phi_i|\phi_j\rangle = \\ &= \sum_{i} c_i^* c_j a_j \delta_{ij} = \\ &= \sum_{i} |c_i|^2 a_i \overset{(3.158)}{=} \\ &= \sum_{i} p_i a_i \end{aligned} \tag{3.159}$$

The expectation value is therefore a weighted average of the eigenvalues a_i of the observable $\hat{A}$. It is clearly distinct from the measured value in each individual measurement: While the measurement result in each individual measurement can only be one of the eigenvalues $\{a_i\}$ of the measurement operator $\hat{A}$, the expectation value arises from a multitude of measurements and can numerically lie between the discrete eigenvalues.

(3.159) can be written more concisely if we directly include beforehand that $\hat{A}$ is Hermitian and it necessarily follows that $\langle\phi_i|\phi_j\rangle = 0$. We can thus directly neglect the sum over j and simply state the operator in its spectral representation:

$$\langle\psi|\hat{A}|\psi\rangle = \sum_i \langle\psi|a_i|\phi_i\rangle \langle\phi_i|\psi\rangle = \sum_i a_i \langle\psi|\hat{P}_i|\psi\rangle = \sum_i a_i p_i$$

Using the projection operator $\langle\psi|\hat{P}_i|\psi\rangle = \langle\psi|\phi_i\rangle \langle\phi_i|\psi\rangle = |c_i|^2$, the expectation value of all possible eigenvalues (measurement values) at time t_0 can be briefly written as:

$$\langle\psi|\hat{A}|\psi\rangle = \sum_i \langle\psi|a_i\hat{P}_i|\psi\rangle = \sum_i p_i a_i \tag{3.160}$$

The expectation value of many measurements therefore corresponds to the averaging of the possible eigenvalues a_i, weighted with the probabilities p_i for each eigenvalue.

3.6.2 Measurement on a Degenerate System

Now consider the more complex case with additional degeneracy, this degree of degeneracy g must be taken into account (i.e., how often a state is degenerate). The eigenvalue equation for the states $|\phi_i\rangle$ must therefore be extended by the degenerate wave functions:

$$\hat{A}\,|\phi_i^n\rangle = a_i\,|\phi_i^n\rangle \tag{3.161}$$

For an eigenvalue a_i, there exist several eigenstates $|\phi_i^n\rangle$, whose degeneracy is indicated by the additional index $n = 1, \ldots, g$. Even in the projector, the degeneracy must be considered, resulting in a new definition:

$$\hat{P}_i^n = |\phi_i^n\rangle \langle\phi_i^n| \quad \text{where} \quad \sum_{n=1}^{g} \hat{P}_i^n = \hat{P}_i \tag{3.162}$$

Applying this to a state $|\psi\rangle$ yields: $\hat{P}_i^n\,|\psi\rangle = c_i^n\,|\phi_i^n\rangle$. If we expand the wave function $|\psi(t_0)\rangle \equiv |\psi\rangle$ into the eigenstates of the observable $\hat{A}$, we obtain an expression analogous to (3.159):

$$|\psi(t_0)\rangle \equiv |\psi\rangle = \mathbb{1}\,|\psi\rangle = \sum_i \sum_{n=1}^{g} \hat{P}_i^n\,|\psi\rangle = \sum_i \sum_{n=1}^{g} c_i^n\,|\phi_i^n\rangle \tag{3.163}$$

The expectation value of $\hat{P}_i$ at time $t = t_0$ corresponds to the probability that the system is in a state with eigenvalue a_i. Accounting for degeneracy, we obtain:

$$\langle\psi|\hat{P}_i|\psi\rangle = \sum_i \sum_{n=1}^{g} |c_i^n|^2$$

If we again measure the expectation value of the observable $\hat{A}$, we can proceed as in (3.160), but must take into account the additional summation over all degeneracies:

$$\langle\psi|\hat{A}|\psi\rangle = \sum_i \sum_{n=1}^{g} \langle\psi|a_i\hat{P}_i^n|\psi\rangle = \sum_i \sum_{n=1}^{g} a_i|c_i^n|^2 = \sum_i \sum_{n=1}^{g} a_i p_i^n \tag{3.164}$$

p_i^n is intended to express the probability that the system is in a state a_i with the degeneracy n.

3.6.3 Measurement on a Continuous System

Let's briefly step away from discrete states and turn to a continuous system. For example, for the one-dimensional position operator $\hat{x}$, the eigenvalue equation $\hat{x}\,|x\rangle = x\,|x\rangle$ from (3.80) holds. Again, there is initially a general superposition state $|\psi\rangle$ that is to be measured at a certain time. The considered observable is the position $\hat{x}$, whereby the expectation value after many measurements is determined by:

$$\begin{aligned}
\langle\hat{x}\rangle = \langle\psi|\hat{x}|\psi\rangle = \langle\psi|\mathbb{1}\hat{x}\mathbb{1}|\psi\rangle &\overset{(3.100)}{=} \\
&= \int dx \int dx'\ \langle\psi|x\rangle\,\langle x|\hat{x}|x'\rangle\,\langle x'|\psi\rangle = \quad |\,\hat{x}\,|x'\rangle = x' \\
&= \int dx \int dx'\, x'\,\langle\psi|x\rangle\,\langle x|x'\rangle\,\langle x'|\psi\rangle \overset{(3.77)}{=} \\
&= \int dx \int dx'\, x'\delta(x-x')\,\langle\psi|x\rangle\,\langle x'|\psi\rangle = \\
&= \int dx\, x\,\langle\psi|x\rangle\,\langle x|\psi\rangle = \int dx\, x\,|\,\langle x|\psi\rangle\,|^2 = \\
&= \int dx\, x\,|\psi(x)|^2
\end{aligned} \tag{3.165}$$

At the end of the first line, (3.100) was used. The expectation value of $\hat{x}$ thus corresponds to the weighted mean of x with the probability density $|\psi(x)|^2$.

3.6.4 Collapse of the Wave Function and Compatible Operators

In the case of a non-degenerate, discrete system, the state $|\psi\rangle$ changes to the collapsed state $|\phi_i\rangle$ through measurement:

$$|\psi\rangle \xrightarrow{\text{measurement}} |\phi_i\rangle = \frac{\hat{P}_i\,|\psi\rangle}{\sqrt{\langle\psi|\hat{P}_i^2|\psi\rangle}} = \frac{\hat{P}_i\,|\psi\rangle}{\sqrt{\langle\psi|\hat{P}_i|\psi\rangle}} \quad \text{with} \quad \hat{P}_i = |\phi_i\rangle\,\langle\phi_i|$$

To obtain a normalized wave function, one divides by the norm, exploiting the fact that $\hat{P}_i$ is idempotent (i.e., it holds that: $\hat{P}_i{}^2 = \hat{P}_i$).
For the degenerate case, a measurement leads to a projection onto the multidimensional (degenerate) subspace associated with the measurement value a_i:

$$|\psi\rangle \xrightarrow{\text{measurement}} \frac{\hat{P}_i\,|\psi\rangle}{\sqrt{\langle\psi|\hat{P}_i|\psi\rangle}} \quad \text{with} \quad \hat{P}_i = \sum_{n=1}^{g} \hat{P}_i^n$$

Let's consider the case where two compatible operators $\hat{A}$ and $\hat{B}$ exist, which satisfy the relation $[\hat{A}, \hat{B}] = 0$. It follows that for both observables, a common, complete orthonormal system $\{|a_i b_j\rangle\}$ must exist. A new projection operator, the product projector $\hat{P}_{ij}$, can thus be defined again:

$$\hat{P}_{ij} = |a_i b_j\rangle \langle a_i b_j| \tag{3.166}$$

The eigenvalues of $\hat{P}_{ij}$ are determined by $p_{ij} = \langle\psi|\hat{P}_{ij}|\psi\rangle = \langle\psi|a_i b_j\rangle \langle a_i b_j|\psi\rangle$, where p_{ij} can be interpreted as the product probability. Furthermore, it must hold that due to completeness, the following relation is fulfilled:

$$\mathbb{1} = \sum_{ij} \hat{P}_{ij} = \sum_{ij} |a_i b_j\rangle \langle a_i b_j| \tag{3.167}$$

The summation indices represent the respective affiliation to an eigenvalue of the corresponding observable. Due to the compatibility of $\hat{A}$ and $\hat{B}$, both observables can be measured simultaneously, and the independent eigenvalues a_i and b_j are obtained. According to (3.160), the expectation value of the observable $\hat{A}$ can be quickly written as:

$$\langle\hat{A}\rangle = \sum_{ij} a_i p_{ij} = \sum_i a_i \Big(\sum_j^{g_j} p_{ij} \Big) = \sum_i a_i p_i \tag{3.168}$$

a_i is independent of the sum over j and can therefore be factored out. If the product probability p_{ij} is summed over all states of $\hat{B}$, a "new" variable p_i can be defined as a *reduced probability* which implicitly already contains that sum. The same applies to the expectation value of $\hat{B}$:

$$\langle\hat{B}\rangle = \sum_{ij} b_i p_{ij} = \sum_j b_j \Big(\sum_i^{g_i} p_{ij} \Big) = \sum_j b_j p_j \tag{3.169}$$

Here the corresponding summation index changes, so that in the end only the summation over the reduced probability p_j remains.

4 Harmonic Oscillator

Motivation: The fundamental role of the harmonic oscillator in physics

The concept of the harmonic oscillator, where the restoring force increases linearly with the displacement for arbitrarily large displacements, is of fundamental importance for physics. Many potentials that have a local minimum can be well approximated by a harmonic potential and thus described analytically. This allows us to approximate many systems very well if only small displacements from the equilibrium position are considered. The concept of the harmonic oscillator also has central importance in the quantization of the electromagnetic field in quantum electrodynamics (QED).

4.1 Analytical Solution

4.1.1 The classical harmonic oscillator

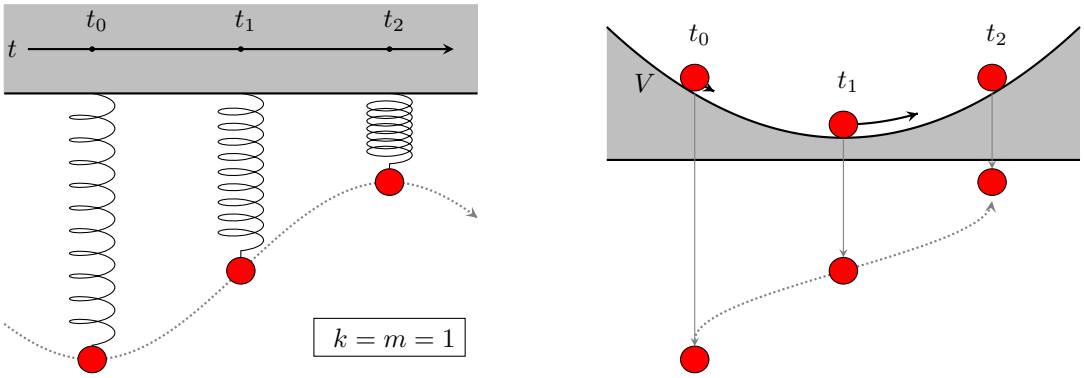


Fig. 17: (left) Model for the spring pendulum: As time progresses, the mass m describes a trajectory $x(t)$ from (4.3). (right) The motion in the quadratic oscillator potential (4.10) leads to the same trajectory (in the illustration, this sine curve corresponds to the height in the potential).

As an introduction, we recap the classical harmonic oscillator using the example of the one-dimensional spring pendulum, which is displaced in the x-direction. Its motion is determined by the second Newton's law $F = ma$:

$$F(t) = m\frac{d^2x(t)}{dt^2} \tag{4.1}$$

The assumed restoring force $F(t) = -kx(t)$ increases linearly with the displacement $x(t)$. The linearity factor k is the spring constant and always acts *against* the current displacement (hence the negative sign). Thus the equation of motion is:

$$-kx(t) = m\frac{d^2x(t)}{dt^2} \tag{4.2}$$

We can define the angular frequency ω as $\omega = \sqrt{k/m}$. The differential equation (4.2) can be solved with the following ansatz:

$$x(t) = A\sin(\omega t) + B\cos(\omega t) \tag{4.3}$$

If we define the amplitude $C = \sqrt{A^2 + B^2}$, Then the phase angle $\varphi = \arctan(A/B)$, the relation (4.3) is equivalent to the ansatz:

$$x(t) = C\cos(\omega t - \varphi) \tag{4.4}$$

But for now, we stick with the ansatz (4.3). Based on this, the velocity $v(t)$ follows for the first time derivative:

$$v(t) = \dot{x}(t) = \omega \left[A\cos(\omega t) - B\sin(\omega t) \right] \tag{4.5}$$

Choosing as initial conditions the lower turning point $-x_u$ of the oscillation ($x_u > 0$), so that $x(0) = -x_u$ and $v(0) = 0$. Then (4.3) and (4.5) simplify to:

$$x(t) = -x_u \cos(\omega t) \tag{4.6}$$

$$v(t) = \omega x_u \sin(\omega t) \tag{4.7}$$

Equation (4.7) can be rewritten to generate a dependence on the positional coordinate:

$$\begin{aligned} v(t) &= \omega x_u \sqrt{\sin^2(\omega t)} = \\ &= \omega x_u \sqrt{1 - \cos^2(\omega t)} = \\ &= \omega \sqrt{x_u^2 - x_u^2 \cos^2(\omega t)} = \\ &= \omega \sqrt{x_u^2 - (-x_u \cos(\omega t))^2} \overset{(4.6)}{=} \\ &= \omega \sqrt{x_u^2 - x^2(t)} \end{aligned}$$

Considering a large number of such harmonic oscillators, the statistical probability $P(x)$ of finding an oscillator at a certain position x can be obtained. The trick here is to convert the corresponding probability $P(t)$ of finding the oscillator at a certain time t within a period between $t = 0$ and $t = T$ into $P(x)$. In particular, since $P(t) = 1/T$, we obtain:

$$P(t)dt = \frac{dt}{T} = \frac{dx}{v(x)T} = \frac{dx}{2\pi\sqrt{x_u^2 - x^2}} = P(x)dx \tag{4.8}$$

Thus, we find the result that the probability of finding the particle at a specific location x is inversely proportional to the speed of the particle – the probability of measuring the particle is greatest at the two turning points. Later we will see that the probabilities of finding calculated wave functions in the classical limit will approximate the form of $P(x)$!

4.1.2 Solution of the Schrödinger equation with the potential of the harmonic oscillator

To solve the Schrödinger equation, we need, instead of the restoring force $F = -kx$, an equivalent potential. This leads us to the one-dimensional harmonic oscillator, whose potential $V(x)$ is:

$$V(x) = \frac{1}{2} m\omega^2 x^2 \tag{4.9}$$

The force corresponds to the negative gradient of the potential – we recognize that from $V(x)$ the spring force $F = -kx$ follows:

$$F = -\frac{dV}{dx} = -\frac{d}{dx}\left(\frac{1}{2} m\omega^2 x^2\right) = -m\omega^2 x = -m\frac{k}{m}x = -kx \ \square \tag{4.10}$$

If we insert this potential into the time-independent, one-dimensional Schrödinger equation (2.8) in the position representation, we get:

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2} m\omega^2 x^2 \right) \psi(x) = E\psi(x) \tag{4.11}$$

Transformation to the dimensionless Schrödinger equation To simplify the Schrödinger equation, we introduce the following dimensionless quantities („reduced units“) ε and y for the energy E and the position coordinate x:

$$E = \hbar\omega\varepsilon \tag{4.12}$$

$$x = x_0 y \tag{4.13}$$

The purpose of this transformation is to bring the Schrödinger equation to a form that is independent of all units. For position and momentum, we need the conversion factors x_0 and p_0:

$$x_0 = \sqrt{\frac{\hbar}{m\omega}} \tag{4.14}$$

$$p_0 = \frac{\hbar}{x_0} = \sqrt{m\hbar\omega} \tag{4.15}$$

Since we substitute x with y, we must adjust both the differentials dx and the wave function $\psi(x)$. From (4.13), it follows:

$$\frac{dy}{dx} = \frac{1}{x_0} \tag{4.16}$$

The substitution of the wave function $\psi(x)$ is performed by changing the dependency; we swap x to $y(x)$, so that $\psi(x) \to \overline{\psi}(y(x)) \equiv \overline{\psi}(y)$. It follows:

$$\begin{aligned}
\frac{d^2}{dx^2}\overline{\psi}(y) &= \frac{d}{dx}\underbrace{\frac{d\overline{\psi}(y)}{dx}}_{\mathrm{f}(y)} = \frac{\mathrm{df}(y)}{dx} = \quad \Big| \ \frac{\mathrm{df}(y)}{dx} = \frac{\mathrm{df}(y)}{dy}\frac{dy}{dx} \\
&= \frac{\mathrm{df}(y)}{dy}\underbrace{\frac{dy}{dx}}_{1/x_0} \overset{(4.16)}{=} \frac{1}{x_0}\frac{\mathrm{df}(y)}{dy} = \frac{1}{x_0}\frac{d}{dy}\frac{d\overline{\psi}(y)}{dx} = \quad \Big| \ \frac{d\overline{\psi}(y)}{dx} = \frac{d\overline{\psi}(y)}{dy}\frac{dy}{dx} \\
&= \frac{1}{x_0}\frac{d}{dy}\Big(\frac{d\overline{\psi}(y)}{dy}\underbrace{\frac{dy}{dx}}_{1/x_0}\Big) \overset{(4.16)}{=} \frac{1}{x_0}\frac{d}{dy}\Big(\frac{d\overline{\psi}(y)}{dy}\frac{1}{x_0}\Big) = \frac{1}{x_0^2}\frac{d^2}{dy^2}\overline{\psi}(y)
\end{aligned} \tag{4.17}$$

We can now insert (4.12), (4.13), and (4.17) into the Schrödinger equation from (4.11):

$$\begin{aligned}
\hbar\omega\varepsilon\overline{\psi}(y) &= \left(-\frac{\hbar^2}{2m}\frac{1}{x_0^2}\frac{d^2}{dy^2} + \frac{1}{2}m\omega^2 x_0^2 y^2\right)\overline{\psi}(y) \overset{(4.14)}{=} \\
&= \left(-\frac{\hbar^2}{2m}\frac{m\omega}{\hbar}\frac{d^2}{dy^2} + \frac{1}{2}m\omega^2\frac{\hbar}{m\omega}y^2\right)\overline{\psi}(y) = \\
&= \left(-\frac{\hbar\omega}{2}\frac{d^2}{dy^2} + \frac{\hbar\omega}{2}y^2\right)\overline{\psi}(y)
\end{aligned}$$

On both sides $\hbar\omega$ can be cancelled – thus obtaining a compact, dimensionless representation of the Schrödinger equation:

$$\frac{1}{2}\left(-\frac{d^2}{dy^2} + y^2\right)\overline{\psi}(y) = \varepsilon\overline{\psi}(y) \tag{4.18}$$

Before we actively solve this equation, we attempt to learn something about the structure of the solution in advance: The potential of the harmonic oscillator (4.9) is symmetric around $x = 0$. Thus, we can expect the wave functions of the stationary (bound) states of the system to exhibit the same symmetry behavior. Formally, this follows from the fact that the parity operator $\hat{\Pi}$ commutes with the Hamilton operator $\hat{H}$, causing the solution functions of $\hat{H}$ to

also be eigenstates of $\hat{\Pi}$. According to the node rule, we can again arrange the eigenstates $\overline{\psi}_n(y)$ and eigenenergies ε_n by the number of "nodes" n.
In the limit $x \to \pm\infty$, the $V(x)$ is infinitely high – as with the finite deep potential well, the wave functions will also be able to penetrate into the "classically forbidden region" of the potential for arbitrary, smaller values of x, but they will decrease exponentially there.

Solution of the Ground State With these considerations, we can motivate the following approach for the dimensionless Schrödinger equation (4.18):

$$\overline{\psi}(y) = N e^{-\frac{1}{2}y^2}\, h(y) \tag{4.19}$$

We assume a Gaussian function that meets the required symmetry and decay behavior for $y \to \pm\infty$. For $h(y)$ we initially assume an arbitrary power series:

$$h(y) = \sum_{i=0}^{\infty} a_i y^i \tag{4.20}$$

Perhaps the simplest case is the special case $h(y) = 1$ (with $a_0 = 1$ and all other $a_i = 0$), which results in a pure Gaussian function:

$$\overline{\psi}_0(y) = N e^{-\frac{1}{2}y^2} \tag{4.21}$$

Substituting this wave function approach into the Schrödinger equation, we can determine the associated eigenvalue ε_0. Since the wave function (4.21) has no nodes (the Gaussian function is positive for all x, hence $n = 0$), we can associate ε_0 with the ground state energy. It follows:

$$\begin{aligned}
\varepsilon_0 N e^{-\frac{1}{2}y^2} &= \frac{N}{2}\left(-\frac{d^2}{dy^2} + y^2\right) e^{-\frac{1}{2}y^2} = \\
&= \frac{N}{2}\left[-\frac{d}{dy}\left(-y\, e^{-\frac{1}{2}y^2}\right) + y^2 e^{-\frac{1}{2}y^2}\right] = \\
&= \frac{N}{2}\left(e^{-\frac{1}{2}y^2} - y^2\, e^{-\frac{1}{2}y^2} + y^2 e^{-\frac{1}{2}y^2}\right) = \frac{1}{2} N e^{-\frac{1}{2}y^2}
\end{aligned}$$

The wave function $\overline{\psi}_0(y)$ cancels on both sides. We thus obtain the reduced ground state energy $\varepsilon_0 = 1/2$. With (4.12) we can back-substitute to the actual energy value and find:

$$E_0 = \frac{1}{2}\hbar\omega \tag{4.22}$$

Unlike in classical physics, the harmonic oscillator in quantum physics has a ground state energy that lies above the minimum of the potential $V(0) = 0$.

In-Depth: Ground State with Zero Energy $E_0 = 0$

If the harmonic quantum oscillator could have a ground state with zero energy, then the position of the particle "at the bottom of the potential" would be precisely defined. However, without energy, the momentum must also vanish, meaning that both position and momentum would be sharply defined. This, however, is fundamentally impossible according to Heisenberg's uncertainty principle. Therefore, in quantum physics, there can be no harmonic oscillator with a ground state energy of zero!

General Solution Approach To obtain all solutions, we must substitute the general solution approach (4.19) into the dimensionless Schrödinger equation (4.18). To do this, we first determine the second derivative of the general approach for $\overline{\psi}(y)$:

$$\begin{aligned}\frac{1}{N}\frac{d^2}{dy^2}\overline{\psi}(y) &= \frac{d^2}{dy^2}\left[e^{-\frac{1}{2}y^2}\,h(y)\right] = \frac{d}{dy}\left[\left(h'(y)-yh(y)\right)e^{-\frac{1}{2}y^2}\right] = \\ &= \Big[h''(y)-yh'(y)-h(y)\Big]e^{-\frac{1}{2}y^2} - \Big[h'(y)-yh(y)\Big]ye^{-\frac{1}{2}y^2} = \\ &= \Big[h''(y)-2yh'(y)-(1-y^2)h(y)\Big]\,e^{-\frac{1}{2}y^2} \end{aligned} \tag{4.23}$$

Now we substitute the approach (4.19) and the result (4.23) into the Schrödinger equation (4.18) to obtain a new differential equation of the function $h(y)$:

$$\begin{aligned}\varepsilon N e^{-\frac{1}{2}y^2}\,h(y) &= \frac{N}{2}\left[-\Big(h''(y)-2yh'(y)-(1-y^2)h(y)\Big)\,e^{-\frac{1}{2}y^2} + y^2h(y)e^{-y^2/2}\right] = \\ &= \frac{N}{2}\left[-h''(y)+2yh'(y)+(1-y^2)h(y)+y^2h(y)\right]e^{-y^2/2} = \\ &= \frac{N}{2}\left[-h''(y)+2yh'(y)+h(y)\right]e^{-y^2/2}\end{aligned}$$

Reducing the Gaussian function terms on both sides of the equation gives us the so-called *Hermite differential equation*:

$$-h''(y)+2yh'(y)+(1-2\varepsilon)h(y)=0 \tag{4.24}$$

It is now up to us to solve this Hermite differential equation (4.24). We already know the development coefficient for $n=0$ based on our assumption that the ground state is described by a Gaussian function: $a_0=1$. As mentioned in (4.20), we use a power series approach for the solution of the Hermite differential equation for the function $h(y)$. To substitute this approach into equation (4.24), we must first compute both derivatives $h'(y)$ and $h''(y)$:

$$\begin{aligned}h'(y) &= \frac{d}{dy}\sum_{i=0}^{\infty}a_iy^i = \frac{d}{dy}\Big(a_0+a_1y+a_2y^2+a_3y^3+\dots\Big) = \\ &= 1a_1+2a_2y+3a_3y^2\cdots = \sum_{i=1}^{\infty}na_iy^{i-1} \end{aligned} \tag{4.25}$$

$$\begin{aligned}h''(y) &= \frac{d}{dy}h'(y) \overset{(4.25)}{=} \frac{d}{dy}\Big(1a_1+2a_2y+3a_3y^2+\dots\Big) = \\ &= 2\cdot 1a_2+3\cdot 2a_3y+\cdots = \sum_{n=i}^{\infty}i(i-1)a_iy^{i-2}\end{aligned} \tag{4.26}$$

Special attention must be paid to the initial summand and its index due to changes depending on the degree of differentiation, avoiding negative exponents (here to $i=1$ or $i=2$). Now we substitute our Ansatz (4.20) and the derived power series (4.25) and (4.26) into the Hermite differential equation (4.24):

$$\underbrace{-\sum_{i=2}^{\infty}i(i-1)a_iy^{i-2}}_{\text{(I) in (4.28)}} + \underbrace{2y\sum_{i=1}^{\infty}na_iy^{i-1}}_{\text{(II) in (4.29)}} + (1-2\varepsilon)\sum_{i=0}^{\infty}a_iy^i = 0 \tag{4.27}$$

We want to modify all y terms in the sums, especially in the (I) and (II) terms, to have the same power y^i to combine everything under a single sum starting at $i=0$. In the second sum (II), this can be easily achieved by multiplying the term y inside the sum. The sum can now

begin at $i = 0$ instead of $i = 1$, as the factor of i results in each summand being naturally zero, avoiding negative exponents:

$$\text{(II)}|\quad 2y\sum_{i=1}^{\infty} i a_i y^{i-1} = 2\sum_{i=1}^{\infty} n a_i y^i = 2\left(1\cdot a_1 y + 2\cdot a_2 y^2 + \dots\right) = \\ = 2\left(0\cdot a_0 + 1\cdot a_1 y + 2\cdot a_2 y^2 + \dots\right) = 2\sum_{i=0}^{\infty} i a_i y^i \tag{4.28}$$

Note that the inclusion of i ensures the term a_0 vanishing at $i = 0$. We can similarly modify the first sum (4.29):

$$\text{(I)}|\quad \sum_{i=2}^{\infty} i(i-1)a_i y^{i-2} = 2\cdot 1\, a_2\, y^0 + 3\cdot 2\, a_3 + \dots = \sum_{j=0}^{\infty}(j+2)(j+1)\, a_{j+2}\, y^j \tag{4.29}$$

For clarification, the index i starting at zero can be reintroduced on the right-hand side of equation (4.29), allowing us to re-express it in equation (4.27):

$$0 = -\sum_{i=0}^{\infty}(i+2)(i+1)a_{i+2}y^i + 2\sum_{i=0}^{\infty} i a_i y^i + (1-2\varepsilon)\sum_{i=0}^{\infty} a_i y^i = \\ = \sum_{i=0}^{\infty}\underbrace{\left[-(i+2)(i+1)a_{i+2} + (2i+1-2\varepsilon)a_i\right]}_{P_i} y^i = \sum_{i=0}^{\infty} P_i y^i \tag{4.30}$$

Equation (4.30) holds only if each individual polynomial P_i is zero, as all separate powers y^i are linearly independent. By setting $P_i = 0$ in (4.30) and solving for a_{i+2}, we obtain the following recursion relation:

$$a_{i+2} = a_i \frac{2i + (1-2\varepsilon)}{(i+2)(i+1)} \tag{4.31}$$

With this recursion, we can determine the polynomials $h(y)$. For example, assuming (without normalization): $a_0 = 1$. Then we can determine all even expansion coefficients $(a_2, a_4, a_6, \dots)$ using (4.31). If a_1 is known, then the odd expansion coefficients $(a_3, a_5, a_7, \dots)$ can be determined.
Due to the symmetry properties of the oscillator potential $V(x)$, we expect $\overline{\psi}(y)$ to have both even and odd solutions (but no solutions with mixed symmetry). Due to the definition of $\overline{\psi}(y)$ in (4.21), and the fact that $e^{-y^2/2}$ has even parity, the polynomial $h(y)$ determines whether $\overline{\psi}(y)$ is even or odd: If $a_0 = 0$, then all a_i with even i are zero, making $h(y)$ and $\overline{\psi}(y)$ have odd parity. If $a_1 = 0$, then all a_i with odd i are zero, making $h(y)$ and $\overline{\psi}(y)$ have even parity.

We now consider whether the series in $h(y)$ can truly have infinitely many terms without causing wave function divergence for y. We examine the asymptotic behavior of a_i as $i \to \infty$. The ratio of coefficients a_{i+2}/a_i initially yields:

$$\frac{a_{i+2}}{a_i} \overset{(4.31)}{=} \frac{2i}{(i+2)(i+1)} + \frac{1-2\varepsilon}{(i+2)(i+1)} = \frac{2i}{i^2+3i+2} + \frac{1-2\varepsilon}{(i+2)(i+1)}$$

While the second term will clearly approach zero as $n \to \infty$, this is not true for the first term. Here, the quadratic term dominates in the denominator:

$$\frac{a_{i+2}}{a_i} \xrightarrow{n\to\infty} \frac{2i}{i^2} = \frac{2}{i} \tag{4.32}$$

Finding a function having the same behavior in the limit $i \to \infty$ allows us to associate this with $h(y)$ and analyze whether the entire, dimensionless wave function $\overline{\psi}(y)$ converges or diverges for

very large y. As y increases indefinitely, the series term y^i with larger i dominates. To verify this, consider a Gaussian function expressed as a series:

$$e^{y^2} = \sum_{k=0}^{\infty} \frac{1}{k!} \left(y^2\right)^k = \sum_{k=0}^{\infty} \frac{1}{k!} y^{2k} = \sum_{i=0,2,4,\ldots}^{\infty} \underbrace{\frac{1}{(i/2)!}}_{b_i} y^i = \sum_{i=0,2,4,\ldots}^{\infty} b_i y^i \tag{4.33}$$

The counting variable i in the last sum only takes even number values. Examining the convergence behavior:

$$\frac{b_{i+2}}{b_i} = \frac{(i/2)!}{(i/2+1)!} = \frac{(i/2)!}{(i/2+1)(i/2)!} = \frac{1}{(i/2+1)} = \frac{2}{i+2} \xrightarrow{i\to\infty} \frac{2}{i}$$

It becomes clear that the Gaussian function demonstrates the same convergence behavior for large i as the recursion found in (4.32). If the series does not cease, it indicates that $h(y)$ behaves like a Gaussian function from (4.33) for large i. This would imply the following problem:

$$\overline{\psi}(y) = e^{-y^2/2}\, h(y) \xrightarrow{i\to\infty} e^{-y^2/2}\, e^{y^2} = e^{y^2/2} \tag{4.34}$$

Yet, this function is *not* normalizable, implying divergence for $y \to \pm\infty$ (contrary to our initial analysis of the potential). This necessitates finding solutions where the infinite series in the Ansatz (4.20) terminates!

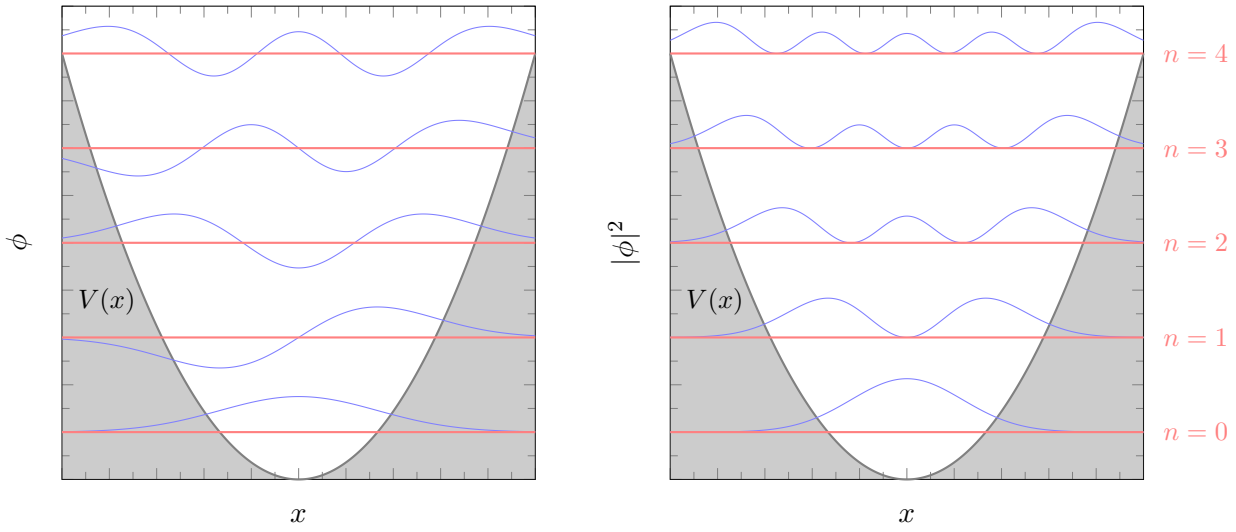


Fig. 18: Representation of the wave functions (left) and the probabilities (right) of the harmonic quantum oscillator. The first five energy states are shown; the symmetry of the eigenstates and the correlation between the number of nodes and energy level are clearly evident.

This circumvents the divergence issue because a finite series can no longer be equated to an analytic function like (4.33). If the series (4.20) terminates, there must be a largest value n for the index i. Setting this value n into the recursion relation (4.31) shows $a_{n+2} = 0$ to ensure the series stops. Consequently, all subsequent a_i vanish when $i > n$. Setting this n in the recursion relation (4.31), where $a_{n+2} = 0$, defines the (dimensionless) eigenenergy ε_n such that the series terminates at index $i = n$:

$$a_{i+2} \overset{!}{=} 0 \overset{(4.31)}{\Longrightarrow} \frac{2n + (1 - 2\varepsilon_n)}{(n+2)(n+1)} = 0 \implies 2n + 1 - 2\varepsilon_n = 0 \implies \varepsilon_n = n + \frac{1}{2} \tag{4.35}$$

Using the substitution $E = \hbar\omega$ from (4.12), the eigenenergies E_n arise:

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) \tag{4.36}$$

Thus, the harmonic quantum oscillator has equidistant eigenenergies! Using the energy quantum number n, we can hierarchically organize individual eigenstates and energies. Recall this applies to the nodes of each wave function as much as energy levels; the greater the energy, the more nodes a wave function manifests (thus stronger oscillation).

Eigenfunctions of the Harmonic Oscillator The obtained expression for the possible (reduced) eigenenergies (4.35) can now be substituted into the differential equation (4.24):

$$\begin{aligned} 0 &= -h''(y) + 2yh'(y) + (1 - 2\varepsilon_n)h(y) \overset{(4.35)}{=} \\ &= -h''(y) + 2yh'(y) + \left[1 - 2\left(n + \frac{1}{2}\right)\right] h(y) = \\ &= -h''(y) + 2yh'(y) + (1 - 2n - 1)\, h(y) = \ \mid \cdot(-1) \\ &= h''(y) - 2yh'(y) + 2nh(y) \end{aligned} \tag{4.37}$$

This differential equation (4.37) is solved by our power series Ansatz (4.20), assuming a specific (reduced) eigenenergy ε_n of the n-th excited state based on (4.35), and then determining the coefficients a_i of the power series using the recursion relation (4.31) (normalization not included). This is equivalent to the solution $h(y) = H_n(y)$ for the n-th eigenenergy ε_n, where $H_n(y)$ denotes the n-th *Hermite polynomial*, which can be defined via the *Rodrigues formula*:

$$H_n(y) = (-1)^n\, e^{y^2}\, \frac{d^n}{dy^n} e^{-y^2} \tag{4.38}$$

A detailed derivation of the Hermite polynomials can be found in Appendix 10.2. By inserting $h(y) = H_n(y)$ in Ansatz (4.19), we finally obtain the (dimensionless) solution:

$$\overline{\psi}_n(y) = N e^{-\frac{1}{2}y^2}\, H_n(y) \tag{4.39}$$

With the back substitution from (4.13), we get the final result for the position-dependent wave function $\psi_n(x)$. The normalization N is determined by the Hermite polynomials:

$$\psi_n(x) = \frac{1}{\pi^{1/4}} \frac{1}{\sqrt{x_0 2^n n!}}\, e^{-\frac{1}{2}(x/x_0)^2} H_n\left(\frac{x}{x_0}\right) \quad \text{with} \quad x_0 = \sqrt{\frac{\hbar}{m\omega}} \tag{4.40}$$

4.1.3 Limits of the Harmonic Oscillator

In the framework of the classical calculation of the harmonic oscillator, we encountered in (4.8) a correlation of the probability proportional to $P(x) \propto (x_u - x)^{-1/2}$. A particle in a classical oscillator potential will therefore be least likely found in the potential minimum, but mainly near the two turning points of its motion. The question arises as to how probability behaves in quantum theory.

The state with the most distinct "quantum-character" is the ground state with $n = 0$, where the probability is maximal at the potential minimum: This is in complete contradiction to classical intuition. Likewise, when energy increases and the principal quantum number n becomes larger, the likelihood of finding the particle in the potential minimum decreases, and we realize we approach the classical limit. A schematic representation of these states is shown in Figure 19. Here, although we still see a significantly higher probability around $x \approx 0$, the behavior of $P(x)$ increasingly approaches the classical expression from (4.8). This behavior is found in other quantum systems: With increasingly large quantum numbers, the system exhibits more "classical-like" behavior (another example is the Rydberg state for very high principal quantum numbers n of the hydrogen atom, which can be approximately described classically).

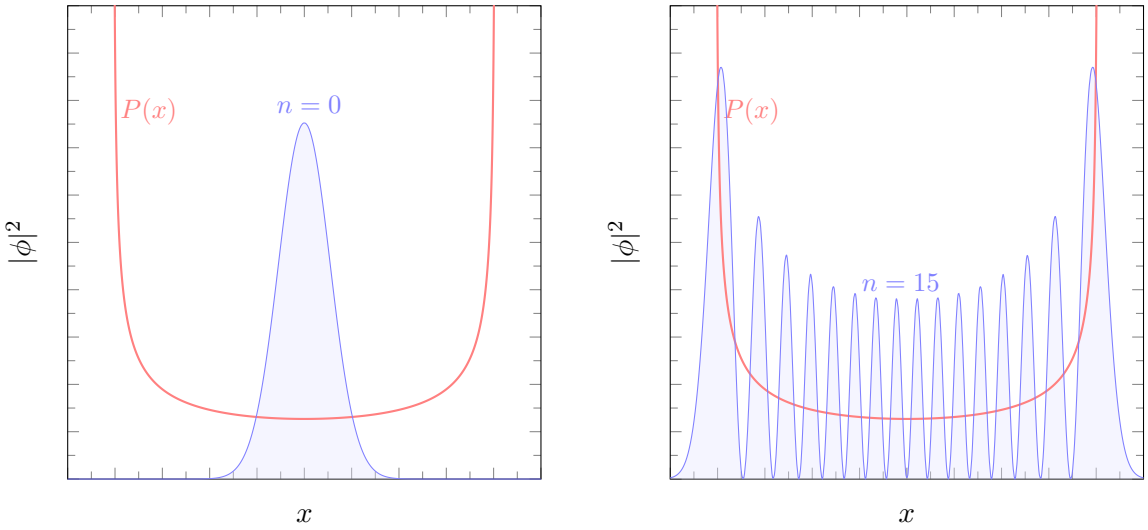


Fig. 19: Limits of the Harmonic Quantum Oscillator: In addition to the classical probability $P(x)$ (red), the wave functions for $n = 0$ and $n = 15$ are shown. The discrepancy between classical and quantum physics becomes particularly evident in the case $n = 0$.

4.2 Algebraic Solution

Motivation: Algebraic Solution Path of the Harmonic Oscillator

In this section, we describe a very elegant solution method for the quantum mechanical harmonic oscillator, developed by Paul Dirac. For this solution path, two so-called ladder operators $\hat{a}$ and $\hat{a}^\dagger$ are defined that respectively extract or add a quantum of energy $\hbar\omega$ to an oscillator.

Although we initially only use the algebraic method to describe the harmonic oscillator within this text, its significance extends far beyond. We will subsequently also use ladder operators to solve problems with quantum mechanical angular momentum more easily. Moreover, ladder operators are essential in the quantization (that is, the quantum mechanical representation) of fields like, for example, the electromagnetic field, which is the foundation of the „second quantization“.

4.2.1 Ladder Operators

Let's start from the one-dimensional, reduced Schrödinger equation of a particle in (4.18) and extract the (reduced) Hamiltonian operator $\bar{H}$:

$$\bar{H} = \frac{1}{2}\left[y^2 - \frac{d^2}{dy^2}\right] \tag{4.41}$$

Foresightedly, we can rewrite this Hamiltonian operator $\bar{H}$:

$$\begin{aligned}
\bar{H} &= \frac{1}{2}\left[\frac{1}{2}\left(y^2 - \frac{d^2}{dy^2}\right) + \frac{1}{2}\left(y^2 - \frac{d^2}{dy^2}\right)\right] = \\
&= \frac{1}{2}\left[\frac{1}{2}\left(y^2 - \frac{d^2}{dy^2} - y\frac{d}{dy} + y\frac{d}{dy}\right) + \frac{1}{2}\left(y^2 - \frac{d^2}{dy^2} + y\frac{d}{dy} - y\frac{d}{dy}\right)\right] = \\
&= \frac{1}{2}\left[\frac{1}{2}\left(y^2 - \frac{d^2}{dy^2} - y\frac{d}{dy} + 1 + y\frac{d}{dy}\right) + \frac{1}{2}\left(y^2 - \frac{d^2}{dy^2} + y\frac{d}{dy} - 1 - y\frac{d}{dy}\right)\right] = \\
&= \frac{1}{2}\left[\frac{1}{2}\left(y^2 - \frac{d^2}{dy^2} - y\frac{d}{dy} + \left\{1 + y\frac{d}{dy}\right\}\right) + \frac{1}{2}\left(y^2 - \frac{d^2}{dy^2} + y\frac{d}{dy} - \left\{1 + y\frac{d}{dy}\right\}\right)\right]
\end{aligned} \tag{4.42}$$

To simplify the curly braces, it helps to investigate the effect of a general operator $\hat{A} = \frac{d}{dy}y$ on a function $\overline{\psi}(y)$:

$$\begin{aligned}
\hat{A}\,\overline{\psi}(y) = \frac{d}{dy}\left(y\overline{\psi}(y)\right) = \overline{\psi}(y) + y\frac{d}{dy}\overline{\psi}(y) = \left\{1 + y\frac{d}{dy}\right\}\overline{\psi}(y) \implies \\
\hat{A} = \left\{1 + y\frac{d}{dy}\right\} \implies \frac{d}{dy}y = \left\{1 + y\frac{d}{dy}\right\}
\end{aligned} \tag{4.43}$$

Thus, we can substitute the two curly braces in (4.42) according to (4.43):

$$\begin{aligned}
\bar{H} &= \frac{1}{2}\left[\frac{1}{2}\left(y^2 - \frac{d^2}{dy^2} - y\frac{d}{dy} + \frac{d}{dy}y\right) + \frac{1}{2}\left(y^2 - \frac{d^2}{dy^2} + y\frac{d}{dy} - \frac{d}{dy}y\right)\right] = \\
&= \frac{1}{2}\Big[\underbrace{\frac{1}{\sqrt{2}}\left(y + \frac{d}{dy}\right)}_{\hat{a}}\underbrace{\frac{1}{\sqrt{2}}\left(y - \frac{d}{dy}\right)}_{\hat{a}^\dagger} + \underbrace{\frac{1}{\sqrt{2}}\left(y - \frac{d}{dy}\right)}_{\hat{a}^\dagger}\underbrace{\frac{1}{\sqrt{2}}\left(y + \frac{d}{dy}\right)}_{\hat{a}}\Big]
\end{aligned} \tag{4.44}$$

We replace the operations in parentheses by the new operators $\hat{a}$ and $\hat{a}^\dagger$. This allows us a new, more compact notation for the Hamiltonian $\bar{H}$:

$$\bar{H} = \frac{1}{2}\left(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}\right) = \frac{1}{2}\{\hat{a}, \hat{a}^\dagger\} \tag{4.45}$$

The dimensionally based Hamiltonian $\hat{H}$ arises from $\bar{H}$ via $\hat{H} = \hbar\omega\bar{H}$. We have now reduced the Hamiltonian to the ladder operators $\hat{a}$ and $\hat{a}^\dagger$. The operator $\hat{a}^\dagger$ is called the creation operator or raising operator:

$$\hat{a}^\dagger = \frac{1}{\sqrt{2}}\left(y - \frac{d}{dy}\right) = \frac{1}{\sqrt{2}}\left(\frac{x}{x_0} - x_0\frac{d}{dx}\right) = \frac{1}{\sqrt{2}}\left(\sqrt{\frac{m\omega}{\hbar}}x - \sqrt{\frac{\hbar}{m\omega}}\frac{d}{dx}\right) \tag{4.46}$$

The operator $\hat{a}$ is the annihilation operator or lowering operator:

$$\hat{a} = \frac{1}{\sqrt{2}}\left(y + \frac{d}{dy}\right) = \frac{1}{\sqrt{2}}\left(\frac{x}{x_0} + x_0\frac{d}{dx}\right) = \frac{1}{\sqrt{2}}\left(\sqrt{\frac{m\omega}{\hbar}}x + \sqrt{\frac{\hbar}{m\omega}}\frac{d}{dx}\right) \tag{4.47}$$

Both operators are generally termed the ladder operators; the reason for this naming will soon become clear. With the definitions (4.46) and (4.47), we express the one-dimensional position operator $\hat{x}$ and momentum operator $\hat{p}$ as a linear combination of the raising and lowering operators:

$$\hat{x} = \frac{x_0}{\sqrt{2}}\left(\hat{a} + \hat{a}^\dagger\right) \tag{4.48}$$

$$\hat{p} = -\frac{ip_0}{\sqrt{2}}\left(\hat{a} - \hat{a}^\dagger\right) \tag{4.49}$$

Example: Position and Momentum Operator as Ladder Operators

To show that the representations of $\hat{x}$ and $\hat{p}$ from (4.48) and (4.49) actually correspond to the position and momentum operator, we substitute the definitions of creation and annihilation from (4.46) and (4.47):

$$\hat{x} = \frac{x_0}{\sqrt{2}}\left(\hat{a} + \hat{a}^\dagger\right) = \frac{x_0}{2}\left(\frac{x}{x_0} + x_0\frac{d}{dx} + \frac{x}{x_0} - x_0\frac{d}{dx}\right) = x \quad \square$$

$$\hat{p} = -\frac{i\hbar}{\sqrt{2}x_0}\left(\hat{a} - \hat{a}^\dagger\right) = -\frac{i\hbar}{2x_0}\left(\frac{x}{x_0} + x_0\frac{d}{dx} - \frac{x}{x_0} + x_0\frac{d}{dx}\right) = -i\hbar\frac{d}{dx} \quad \square$$

The momentum operator is traditionally represented in the position space – therefore the derivative with respect to x can be reversed by the action of *hatp*:

$$\frac{d}{dx} = -\frac{1}{i\hbar}\hat{p} = \frac{i}{\hbar}\hat{p}$$

If we substitute this into the definitions (4.46) and (4.47), due to the complex term, it becomes immediately apparent that $\hat{a}$ and $\hat{a}^\dagger$ are not Hermitian:

$$\hat{a}^\dagger = \frac{1}{\sqrt{2}}\left(\frac{\hat{x}}{x_0} - i\frac{x_0}{\hbar}\hat{p}\right) \tag{4.50}$$

$$\hat{a} = \frac{1}{\sqrt{2}}\left(\frac{\hat{x}}{x_0} + i\frac{x_0}{\hbar}\hat{p}\right) \tag{4.51}$$

It holds $\hat{a} \neq \hat{a}^\dagger$ because they differ in the sign before the complex number in (4.50) and (4.51). Using the provided definitions (4.13), and (4.15), we express the position operator (4.48) and momentum operator (4.49) in terms of the dimensionless position operator $\hat{y}$ and momentum operator $\hat{p}_y$:

$$\hat{y} = \frac{1}{\sqrt{2}}(\hat{a} + \hat{a}^\dagger) \tag{4.52}$$

$$\hat{p}_y = -\frac{i}{\sqrt{2}}(\hat{a} - \hat{a}^\dagger) \tag{4.53}$$

4.2.2 Commutators of the Ladder Operators and Number State Operator

The two trivial commutators between the faller $\hat{a}$ and the raiser $\hat{a}^\dagger$ follow:

$$[\hat{a}, \hat{a}] = [\hat{a}^\dagger, \hat{a}^\dagger] = 0 \tag{4.54}$$

However, the two operators $\hat{a}$ and $\hat{a}^\dagger$ do not commute with each other. The following important commutator relationship hence follows:

$$[\hat{a}, \hat{a}^\dagger] = \hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a} = 1 \tag{4.55}$$

Example: Commutator between $\hat{a}$ and $\hat{a}^\dagger$

To prove the relation from (4.55), we substitute the definitions (4.46) and (4.47) for $\hat{a}^\dagger$ and $\hat{a}$. As these are operators with specific effects, carefully consider their operation on a preceded wave function $\overline{\psi}(y)$:

$$
\begin{aligned}
[\hat{a}, \hat{a}^\dagger]\,\overline{\psi}(y) &= \left(\hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a}\right)\overline{\psi} = \\
&= \frac{1}{2}\left[\left(y + \frac{d}{dy}\right)\left(y - \frac{d}{dy}\right)\overline{\psi} - \left(y - \frac{d}{dy}\right)\left(y + \frac{d}{dy}\right)\overline{\psi}\right] = \\
&= \frac{1}{2}\left[\left(y + \frac{d}{dy}\right)\left(y\overline{\psi} - \frac{d\overline{\psi}}{dy}\right) - \left(y - \frac{d}{dy}\right)\left(y\overline{\psi} + \frac{d\overline{\psi}}{dy}\right)\right] = \\
&= \frac{1}{2}\left[\left(y^2\overline{\psi} - y\frac{d\overline{\psi}}{dy} + \frac{d}{dy}\left(y\overline{\psi}\right) - \frac{d^2\overline{\psi}}{dy^2}\right) - \left(y^2\overline{\psi} + y\frac{d\overline{\psi}}{dy} - \frac{d}{dy}\left(y\overline{\psi}\right) - \frac{d^2\overline{\psi}}{dy^2}\right)\right] = \\
&= \frac{1}{2}\left[y^2\overline{\psi} - y\frac{d\overline{\psi}}{dy} + \frac{d}{dy}\left(y\overline{\psi}\right) - \frac{d^2\overline{\psi}}{dy^2} - y^2\overline{\psi} - y\frac{d\overline{\psi}}{dy} + \frac{d}{dy}\left(y\overline{\psi}\right) + \frac{d^2\overline{\psi}}{dy^2}\right] = \\
&= \frac{d}{dy}\left(y\overline{\psi}\right) - y\frac{d\overline{\psi}}{dy} = \overline{\psi} + y\frac{d\overline{\psi}}{dy} - y\frac{d\overline{\psi}}{dy} = 1 \cdot \overline{\psi} \quad \square
\end{aligned}
$$

Equivalently, a more concise approach avoids using position space representation but uses the canonical commutation relation $[\hat{x}, \hat{p}] = i\hbar$ from (3.133). We omit the wave function for this proof and imply the operation of the operators:

$$
\begin{aligned}
[\hat{a}, \hat{a}^\dagger] &= \frac{1}{2}\left[\left(\frac{\hat{x}}{x_0} + i\frac{x_0}{\hbar}\hat{p}\right), \left(\frac{\hat{x}}{x_0} - i\frac{x_0}{\hbar}\hat{p}\right)\right] = \\
&= \frac{1}{2}\left(\frac{1}{x_0^2}[\hat{x}, \hat{x}] - \frac{i}{\hbar}[\hat{x}, \hat{p}] + \frac{i}{\hbar}[\hat{p}, \hat{x}] + \frac{x_0^2}{\hbar^2}[\hat{p}, \hat{p}]\right) = \\
&= -\frac{1}{2}\left(\frac{i}{\hbar}[\hat{x}, \hat{p}] + \frac{i}{\hbar}[\hat{x}, \hat{p}]\right) = -\frac{i}{\hbar}[\hat{x}, \hat{p}] \overset{(3.133)}{=} \\
&= -\frac{i}{\hbar} i\hbar = 1 \quad \square
\end{aligned}
$$

With a simple transformation of (4.55), the following useful relation can be derived:

$$\hat{a}\hat{a}^\dagger = 1 + \hat{a}^\dagger\hat{a} \tag{4.56}$$

Using this relation, we can simplify the dimensionless Hamiltonian (4.45):

$$\bar{H} = \frac{1}{2}\left(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}\right) \overset{(4.56)}{=} \frac{1}{2}\left(1 + \hat{a}^\dagger\hat{a} + \hat{a}^\dagger\hat{a}\right) = \hat{a}^\dagger\hat{a} + \frac{1}{2}$$

Typically, instead of $\hat{a}^\dagger\hat{a}$, a new operator called the *number state operator* $\hat{N}$ is defined:

$$\hat{N} = \hat{a}^\dagger\hat{a} \tag{4.57}$$

Hence, we find a new, compact notation for the dimensioned Hamiltonian $\hat{H}$ using the number state operator $\hat{N}$:

$$\hat{H} = \hbar\omega\left(\hat{N} + \frac{1}{2}\right) \quad \text{with} \quad \hat{N} = \hat{a}^\dagger\hat{a} \tag{4.58}$$

Properties of the Number State Operator How does the number state operator act on an eigenstate of the Hamiltonian operator $\hat{H}$? From the analytical solution of the Schrödinger equation in the oscillator potential, we know the solutions (4.36) of the eigenenergies; based on the eigenvalue equation $\hat{H}\,|n\rangle = E_n\,|n\rangle$, where $|n\rangle$ are the wave functions of the n-th eigenenergy state of the harmonic oscillator:

$$0 = \left(\hat{H} - E_n\right)|n\rangle \overset{(4.36)}{=} \hbar\omega\left[\left(\hat{N} + \frac{1}{2}\right) - \left(n + \frac{1}{2}\right)\right]|n\rangle = \hbar\omega\left(\hat{N} - n\right)|n\rangle$$

Reorganizing this equation leads to another crucial expression, namely the effect of the number state operator $\hat{N}$ on an eigenenergy state $|n\rangle$ of the harmonic oscillator. Specifically, by its eigenvalue n, $\hat{N}$ conveys the information on which excitation state $|n\rangle$ resides.

$$\hat{N}\,|n\rangle = n\,|n\rangle \tag{4.59}$$

An eigenvalue n of the number state operator $\hat{N}$ can only take values greater than or equal to zero:

$$n = \langle n|\hat{N}|n\rangle = \langle n|\hat{a}^\dagger\hat{a}|n\rangle = \langle \hat{a}n|\hat{a}n\rangle = ||\,|\hat{a}n\rangle\,||^2 \geq 0 \tag{4.60}$$

For $n = 0$, this denotes the ground state of the system, and for $n > 0$, an excited state. Also, for arbitrary functions, the expectation value formation ensures the relationship $\langle\psi|\hat{N}|\psi\rangle \geq 0$ (positivity). The number state operator is *Hermitian* and *positive*; thus, we get positive, real eigenvalues in the eigen-system. Hermiticity follows directly from $(\hat{N})^\dagger = (\hat{a}^\dagger\hat{a})^\dagger = \hat{a}^\dagger\hat{a} = \hat{N}$.

Commutators of the Number State Operator Additionally, commutator relations exist between $\hat{N}$ and the ladder operators, utilisng the already known commutator relation (4.55):

$$\begin{aligned}[\hat{N}, \hat{a}] = [\hat{a}^\dagger\hat{a}, \hat{a}] = \hat{a}^\dagger\hat{a}\hat{a} - \hat{a}\hat{a}^\dagger\hat{a} = (\hat{a}^\dagger\hat{a} - \hat{a}\hat{a}^\dagger)\hat{a} = [\hat{a}^\dagger, \hat{a}]\hat{a} = -[\hat{a}, \hat{a}^\dagger]\hat{a} \overset{(4.55)}{=} \\ = -\hat{a}\end{aligned} \tag{4.61}$$

From a straightforward transformation of (4.61), the following useful relation can be derived:

$$\hat{N}\hat{a} - \hat{a}\hat{N} = -\hat{a} \implies \hat{N}\hat{a} = \hat{a}\hat{N} - \hat{a} \tag{4.62}$$

Exploiting the identity $[AB, C] = A[B, C] + [A, C]B$, the commutator $[\hat{N}, \hat{a}^\dagger]$ is simply determined:

$$\begin{aligned}[\hat{N}, \hat{a}^\dagger] = [\hat{a}^\dagger\hat{a}, \hat{a}^\dagger] = \hat{a}^\dagger[\hat{a}, \hat{a}^\dagger] - [\hat{a}^\dagger, \hat{a}^\dagger]\hat{a} = \hat{a}^\dagger[\hat{a}, \hat{a}^\dagger] \overset{(4.55)}{=} \\ = \hat{a}^\dagger\end{aligned} \tag{4.63}$$

This yields another beneficial relation via rearrangement of (4.63):

$$\hat{N}\hat{a}^\dagger - \hat{a}^\dagger\hat{N} = \hat{a}^\dagger \implies \hat{N}\hat{a}^\dagger = \hat{a}^\dagger + \hat{a}^\dagger\hat{N} \tag{4.64}$$

4.2.3 Action of Raising and Lowering Operators

Let $\hat{N}$ act on a state $|\hat{a}^\dagger n\rangle$, we can explicitly determine the action of $\hat{a}^\dagger$:

$$\begin{aligned}\hat{N}\,|\hat{a}^\dagger n\rangle = \hat{N}\hat{a}^\dagger\,|n\rangle \overset{(4.64)}{=} (\hat{a}^\dagger + \hat{a}^\dagger\hat{N})\,|n\rangle = \hat{a}^\dagger(1 + \hat{N})\,|n\rangle = \hat{a}^\dagger(1 + n)\,|n\rangle = \\ = (n + 1)\hat{a}^\dagger\,|n\rangle = (n + 1)\,|\hat{a}^\dagger n\rangle\end{aligned} \tag{4.65}$$

The state $|\hat{a}^\dagger n\rangle$ is thus an eigenfunction of $\hat{N}$ with the eigenvalue $n + 1$. Or, in other words: the number state or the excitation level of $|n\rangle$ is increased by one by the application of $\hat{a}^\dagger$; this is why $\hat{a}^\dagger$ is also called the raising operator! We get the same eigenvalue $n + 1$ when we let the

number state operator $\hat{N}$ act on a state $|n+1\rangle$ – up to a scaling factor c, this will correspond to $|\hat{a}^\dagger n\rangle \equiv \hat{a}^\dagger |n\rangle = c\,|n+1\rangle$. The specific form of c can be determined from the modulus squared construction of $|\alpha^\dagger n\rangle$:

$$\begin{aligned} || \, |\hat{a}^\dagger n\rangle \, ||^2 &= \langle n+1|c^* c|n+1\rangle = |c|^2 \, \langle n+1|n+1\rangle = |c|^2 \\ || \, |\hat{a}^\dagger n\rangle \, ||^2 &= \langle n|\hat{a}\,\hat{a}^\dagger|n\rangle \overset{(4.56)}{=} \langle n|\hat{a}^\dagger\hat{a}+1|n\rangle = \langle n|\hat{N}+1|n\rangle = \langle n|n+1|n\rangle = \\ &= (n+1)\langle n|n\rangle = n+1 \end{aligned} \tag{4.66}$$

We find two solutions for $c = \pm\sqrt{n+1}$ (where we disregard a global phase $e^{i\varphi}$). The negative solution is neglected to guarantee a positive eigenvalue of $\hat{N}$ according to (4.60). From (4.65) and (4.66) it follows thus for the action of the raising operator $\hat{a}^\dagger$:

$$\hat{a}^\dagger \, |n\rangle = \sqrt{n+1} \, |n+1\rangle \tag{4.67}$$

In a similar way, we can also determine the action of the lowering operator. First, we consider the action of $\hat{N}$ on a state $|\hat{a}n\rangle$:

$$\begin{aligned} \hat{N} \, |\hat{a}n\rangle = \hat{N}\hat{a} \, |n\rangle &\overset{(4.62)}{=} (\hat{a}\hat{N} - \hat{a}) \, |n\rangle = \hat{a}(\hat{N}-1) \, |n\rangle = \hat{a}(n-1) \, |n\rangle = \\ &= (n-1)\hat{a} \, |n\rangle = (n-1) \, |\hat{a}n\rangle \end{aligned} \tag{4.68}$$

The state $|\hat{a}n\rangle$ is thus an eigenfunction of $\hat{N}$ with the eigenvalue $n-1$, just like the state $|n-1\rangle$. Or, in other words: the number state or the excitation level of $|n\rangle$ is decreased by one by the application of $\hat{a}$; this is why $\hat{a}$ is also called the lowering operator! The eigenstate $|\hat{a}n\rangle$ again corresponds, up to a scaling factor c, to the state $|n-1\rangle$ – it follows: $|\hat{a}n\rangle = c\,|n-1\rangle$. From the construction of the modulus squared, we can determine the form of c:

$$\begin{aligned} || \, |\hat{a}n\rangle \, ||^2 &= \langle n-1|c^* c|n-1\rangle = |c|^2 \, \langle n-1|n-1\rangle = |c|^2 \\ || \, |\hat{a}n\rangle \, ||^2 &= \langle n|\hat{a}^\dagger\hat{a}|n\rangle \overset{(4.57)}{=} \langle n|\hat{N}|n\rangle = \langle n|n|n\rangle = n\langle n|n\rangle = \\ &= n \end{aligned} \tag{4.69}$$

Again, we find two solutions $c = \sqrt{n}$, whereby due to the positivity of the number state operator $\hat{N}$, the negative scaling factor can be excluded (and a global phase $e^{i\varphi}$ ignored). Thus, from (4.68) and (4.69) the complete action of the lowering operator results:

$$\hat{a} \, |n\rangle = \sqrt{n} \, |n-1\rangle \tag{4.70}$$

Applying the lowering operator to the ground state $|0\rangle$, one obtains zero:

$$\hat{a} \, |0\rangle = 0 \tag{4.71}$$

Intuitively, this is also logical: if a state is already in the ground state, it cannot be lowered further. Starting from the ground state $|0\rangle$, we can generate a state $|n\rangle$ by applying the raising operator $\hat{a}^\dagger$ n times:

$$(\hat{a}^\dagger)^n \, |0\rangle = \sqrt{n!} \, |n\rangle \tag{4.72}$$

Rearranging (4.72) for $|n\rangle$ gives us an expression algebraically equivalent to Hermite polynomials from (4.40):

$$|n\rangle = \frac{1}{\sqrt{n!}} (\hat{a}^\dagger)^n \, |0\rangle \tag{4.73}$$

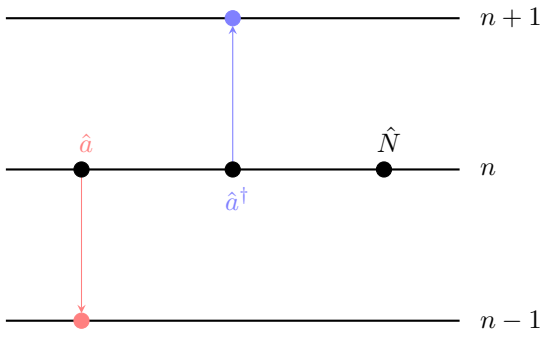


Fig. 20: Schematic action of the raising operator $\hat{a}^\dagger$, the lowering operator $\hat{a}$, and the number state operator $\hat{N}$.

Example: Action of $\hat{N}$, $\hat{a}^\dagger$ and $\hat{a}$ on the Ground State in Position Space

We have now encountered two equivalent formalisms for describing the quantum harmonic oscillator: the analytical approach, which led us to the Hermite polynomials (4.40) as functions of position x, and the algebraic approach, through which new operators $\hat{N}$, $\hat{a}^\dagger$, and $\hat{a}$ were introduced. How do the number operator, raising operator, and lowering operator act when the oscillator eigenstates $|n\rangle$ are projected into the "dimensionless position space" $|y\rangle$? We will demonstrate this using the simplest case, the ground state $|0\rangle$:

$$\overline{\psi}_0(y) = \langle y|0\rangle \overset{(4.21)}{=} Ne^{-\frac{1}{2}y^2}$$

From (4.46) and (4.47), the dimensionless number operator can be derived:

$$\hat{N} = \frac{1}{2}\left(y - \frac{d}{dy}\right)\left(y + \frac{d}{dy}\right) = \frac{1}{2}\left(y^2 - \frac{d^2}{dy^2} - \frac{d}{dy}y + y\frac{d}{dy}\right)$$

The action of the number operator $\hat{N}$ in position space is thus:

$$\begin{aligned}
\hat{N}\,|0\rangle \to \frac{1}{2}\left(y^2 - \frac{d^2}{dy^2} - \frac{d}{dy}y + y\frac{d}{dy}\right)Ne^{-\frac{1}{2}y^2} = \\
= \frac{N}{2}\left(y^2e^{-\frac{1}{2}y^2} - \underbrace{\frac{d^2}{dy^2}e^{-\frac{1}{2}y^2}}_{A} - \underbrace{\frac{d}{dy}\left(ye^{-\frac{1}{2}y^2}\right)}_{B} + \underbrace{y\frac{d}{dy}e^{-\frac{1}{2}y^2}}_{C}\right) = \\
= \frac{N}{2}\left(y^2e^{-\frac{1}{2}y^2} - A - B + C\right) \qquad (4.74)
\end{aligned}$$

In a short side calculation, we determine the expressions for A, B, and C:

$$\begin{aligned}
A = \frac{d^2}{dy^2}e^{-\frac{1}{2}y^2} = \frac{d}{dy}\left(-e^{-\frac{1}{2}y^2}\frac{1}{2}2\,y\right) = -\frac{d}{dy}\left(ye^{-\frac{1}{2}y^2}\right) = \\
= -\left(\frac{d}{dy}y\right)e^{-\frac{1}{2}y^2} - y\frac{d}{dy}e^{-\frac{1}{2}y^2} = -1e^{-\frac{1}{2}y^2} + ye^{-\frac{1}{2}y^2}\frac{1}{2}2\,y \\
= y^2e^{-\frac{1}{2}y^2} - e^{-\frac{1}{2}y^2} \qquad (4.75)
\end{aligned}$$

$$B = \frac{d}{dy}\left(ye^{-\frac{1}{2}y^2}\right) = \left(\frac{d}{dy}y\right)e^{-\frac{1}{2}y^2} + y\frac{d}{dy}e^{-\frac{1}{2}y^2} = 1e^{-\frac{1}{2}y^2} - ye^{-\frac{1}{2}y^2}\frac{1}{2}2\,y$$
$$= e^{-\frac{1}{2}y^2} - y^2e^{-\frac{1}{2}y^2} \tag{4.76}$$

$$C = y\frac{d}{dy}e^{-\frac{1}{2}y^2} = -ye^{-\frac{1}{2}y^2}\frac{1}{2}2\,y = -y^2e^{-\frac{1}{2}y^2} \tag{4.77}$$

By inserting (4.75), (4.76), and (4.77) into (4.74), we finally obtain the following result:

$$\hat{N}\,|0\rangle \to \frac{N}{2}\Big(y^2e^{-\frac{1}{2}y^2}\underbrace{-y^2e^{-\frac{1}{2}y^2} + e^{-\frac{1}{2}y^2}}_{-A}\underbrace{-e^{-\frac{1}{2}y^2} + y^2e^{-\frac{1}{2}y^2}}_{-B}\underbrace{-y^2e^{-\frac{1}{2}y^2}}_{+C}\Big) =$$
$$= \frac{N}{2}\left(e^{-\frac{1}{2}y^2} - e^{-\frac{1}{2}y^2}\right) = 0$$

We obtain the expected result: the ground state corresponds to an occupation number of $n = 0$ (since $\langle y|\hat{N}|0\rangle = 0 \cdot \langle y|n\rangle = 0$). For the action of the raising operator, we find

$$\langle y|\hat{a}^\dagger|0\rangle = \frac{N}{\sqrt{2}}\left(y - \frac{d}{dy}\right)e^{-\frac{1}{2}y^2} = \frac{N}{\sqrt{2}}\left(ye^{-\frac{1}{2}y^2} + ye^{-\frac{1}{2}y^2}\right) = \frac{\tilde{N}}{\sqrt{2}}\,2y\,e^{-\frac{1}{2}y^2}$$

The new term $2y$ corresponds to the first Hermite polynomial $H_1(y) = 2y$ (without explicitly recalculating the normalization $\tilde{N}$). For the lowering operator, we have:

$$\langle y|\hat{a}|0\rangle = \frac{N}{\sqrt{2}}\left(y + \frac{d}{dy}\right)e^{-\frac{1}{2}y^2} = \frac{N}{\sqrt{2}}\left(ye^{-\frac{1}{2}y^2} - ye^{-\frac{1}{2}y^2}\right) = 0$$

The ground state can therefore no longer be reduced by the annihilation operator $\hat{a}$.

Example: Ground state in position space from $\hat{a}\,|0\rangle = 0$

According to (4.71): $\hat{a}\,|0\rangle = 0$; in fact, from this we can calculate the position representation of the ground state $\overline{\psi}_0(y) \equiv \overline{\psi}_0$. If we project the entire equation onto the reduced position space $|y\rangle$, the annihilation operator changes accordingly (4.47) and we can write down the following differential equation:

$$\langle y|\hat{a}|0\rangle = \frac{1}{\sqrt{2}}\left(y - \frac{d}{dy}\right)\overline{\psi}_0 = 0 \Longrightarrow \frac{d\overline{\psi}_0}{\overline{\psi}_0} = -ydy \Longrightarrow \ln(\overline{\psi}_0) = -\frac{1}{2}y^2 + C$$

In the last step, both sides were integrated independently; exponentiating both sides gives the already known result (4.21) for the ground state:

$$\overline{\psi}_0 = Ne^{-\frac{1}{2}y^2}$$

In-Depth: Algebraic and Analytical Solutions

How can the algebraic and the analytical solution be transformed into one another? Project a general eigenstate $|n\rangle$ into the reduced position space $\langle y|$ and use (4.40), (10.17), and (4.38), we obtain:

$$\begin{aligned}
\langle y|n\rangle = \psi_n(y) &= \frac{1}{\pi^{1/4}}\frac{1}{\sqrt{2^n n!}} e^{-\frac{1}{2}y^2} H_n(y) \overset{(4.38)}{=} \\
&= \frac{1}{\pi^{1/4}}\frac{1}{\sqrt{2^0 n!}} e^{-\frac{1}{2}y^2} \cdot \left\{ (-1)^n e^{y^2} \frac{d^n}{dy^n} e^{-y^2} \right\} \overset{(10.17)}{=} \\
&= \frac{1}{(2\pi)^{1/4}}\frac{1}{\sqrt{n!}} e^{-\frac{1}{2}y^2} \cdot \left\{ e^{\frac{1}{2}y^2} \left[y - \frac{d}{dy} \right]^n e^{-\frac{1}{2}y^2} \right\} = \\
&= \frac{1}{\sqrt{n!}} \left[y - \frac{d}{dy} \right]^n \left\{ \frac{1}{(2\pi)^{1/4}} e^{-\frac{1}{2}y^2} \right\} \overset{(4.46)}{=} \\
&= \frac{1}{\sqrt{n!}} \langle y|(\hat{a}^\dagger)^n|0\rangle
\end{aligned} \tag{4.78}$$

We have managed to derive an expression from the wave function in position space and the Rodrigues formula of Hermite polynomials, which matches the representation of the raising operator in position space (4.46). However, we must note that the normalization changes by a factor of $2^{-1/4}$ in the transition from the Rodrigues formula (4.38) to (10.17), since the Gaussian function now decreases at half the rate.

4.2.4 Heisenberg Uncertainty Relation

To construct the Heisenberg uncertainty relation concerning the eigenstates of the harmonic oscillator, we need to calculate the position variance $\langle(\Delta\hat{x})^2)\rangle$ and momentum variance $\langle(\Delta\hat{p})^2)\rangle$, as described in (3.140). For reminder: for an arbitrary operator $\hat{A}$, it is generally true with the fluctuation operator $\Delta\hat{A}$:

$$\langle(\Delta\hat{A})^2\rangle = \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2$$

To calculate the expectation value of the position $\langle\hat{x}\rangle$ for a particular excitation state $|n\rangle$, we use the representation of the position operator with ladder operators from (4.48), and the action of the ladder operators on an eigenstate $|n\rangle$ according to (4.67) and (4.70):

$$\begin{aligned}
\langle\hat{x}\rangle &= \langle n|\hat{x}|n\rangle \overset{(4.48)}{=} \\
&= \frac{x_0}{\sqrt{2}} \langle n|\hat{a} + \hat{a}^\dagger|n\rangle = \\
&= \frac{x_0}{\sqrt{2}} \left(\langle n|\hat{a}|n\rangle + \langle n|\hat{a}^\dagger|n\rangle \right) \overset{(4.67,4.70)}{=} \\
&= \frac{x_0}{\sqrt{2}} \left(\sqrt{n}\langle n|n-1\rangle + \sqrt{n+1}\langle n|n+1\rangle \right) = 0
\end{aligned} \tag{4.79}$$

The eigenstates $|n+1\rangle$ and $|n-1\rangle$ are both orthogonal to $|n\rangle$, as the energy eigenfunctions form a complete orthonormal system $\{|n\rangle\}$. We thus expect an average localization at $x = 0$ in the middle of the potential. At this point, we also recognize the power of the ladder operators: instead of working with the complicated representation in position space, we find here only abstract eigenstates, whose behavior we can exactly describe under certain operations ($\hat{a}^\dagger$,$\hat{a}$ and $\hat{N}$)! Complicated computational operations can thus be elegantly and compactly expressed in algebraic notation.

For the expectation value of the momentum $\langle\hat{p}\rangle$, we can analogously use the representation of the momentum operator with raising operators (4.49):

$$\begin{aligned}
\langle\hat{p}\rangle &= \langle n|\hat{p}|n\rangle \overset{(4.49)}{=} \\
&= -\frac{i}{\sqrt{2}}\frac{\hbar}{x_0}\langle n|\hat{a}-\hat{a}^\dagger|n\rangle = \\
&= -\frac{i}{\sqrt{2}}\frac{\hbar}{x_0}\left(\langle n|\hat{a}|n\rangle - \langle n|\hat{a}^\dagger|n\rangle\right) \overset{(4.67,4.70)}{=} \\
&= -\frac{i}{\sqrt{2}}\frac{\hbar}{x_0}\left(\sqrt{n}\langle n|n-1\rangle - \sqrt{n+1}\langle n|n+1\rangle\right) = 0
\end{aligned} \tag{4.80}$$

The expectation value of the momentum thus also vanishes! For the second moment of $\hat{x}$, we find an expression different from zero:

$$\begin{aligned}
\langle\hat{x}^2\rangle &= \langle n|\hat{x}^2|n\rangle \overset{(4.48)}{=} \\
&= \frac{x_0^2}{2}\langle n|(\hat{a}+\hat{a}^\dagger)(\hat{a}+\hat{a}^\dagger)|n\rangle = \\
&= \frac{x_0^2}{2}\langle n|\hat{a}\hat{a}+\hat{a}^\dagger\hat{a}^\dagger+\hat{a}\hat{a}^\dagger+\hat{a}^\dagger\hat{a}|n\rangle \overset{(4.56)}{=} \\
&= \frac{x_0^2}{2}\langle n|\hat{a}\hat{a}+\hat{a}^\dagger\hat{a}^\dagger+1+\hat{a}^\dagger\hat{a}+\hat{a}^\dagger\hat{a}|n\rangle = \\
&= \frac{x_0^2}{2}\left(\langle n|\hat{a}\hat{a}|n\rangle + \langle n|\hat{a}^\dagger\hat{a}^\dagger|n\rangle + 2\langle n|\hat{a}^\dagger\hat{a}|n\rangle + \langle n|n\rangle\right) \overset{(4.57)}{=} \\
&= \frac{x_0^2}{2}\left(\langle n|\hat{a}\hat{a}|n\rangle + \langle n|\hat{a}^\dagger\hat{a}^\dagger|n\rangle + 2\langle n|\hat{N}|n\rangle + 1\right) \overset{(4.67,4.70)}{=} \\
&= \frac{x_0^2}{2}\left(\sqrt{n}\sqrt{n-1}\langle n|n-2\rangle + \sqrt{n+1}\sqrt{n+2}\langle n|n+2\rangle + 2n\langle n|n\rangle + 1\right) = \\
&= x_0^2\left(n+\frac{1}{2}\right) \overset{(4.13)}{=} \frac{\hbar}{m\omega}\left(n+\frac{1}{2}\right)
\end{aligned} \tag{4.81}$$

While the expectation value seems to be centered around the midpoint of the oscillator potential $\langle\hat{x}\rangle = 0$, the variance of the position does not vanish; there are therefore fluctuations around the equilibrium position. The second moment of the momentum operator can be calculated analogously:

$$\begin{aligned}
\langle\hat{p}^2\rangle &= \langle n|\hat{p}^2|n\rangle \overset{(4.49)}{=} \\
&= -\frac{p_0^2}{2}\langle n|(\hat{a}-\hat{a}^\dagger)(\hat{a}-\hat{a}^\dagger)|n\rangle = \\
&= -\frac{p_0^2}{2}\langle n|\hat{a}\hat{a}+\hat{a}^\dagger\hat{a}^\dagger-\hat{a}\hat{a}^\dagger-\hat{a}^\dagger\hat{a}|n\rangle = \\
&= \frac{p_0^2}{2}\langle n|-\hat{a}\hat{a}-\hat{a}^\dagger\hat{a}^\dagger+\hat{a}\hat{a}^\dagger+\hat{a}^\dagger\hat{a}|n\rangle \overset{(4.56)}{=} \\
&= \frac{p_0^2}{2}\langle n|-\hat{a}\hat{a}-\hat{a}^\dagger\hat{a}^\dagger+1+\hat{a}^\dagger\hat{a}+\hat{a}^\dagger\hat{a}|n\rangle = \\
&= \frac{p_0^2}{2}\left(-\langle n|\hat{a}\hat{a}|n\rangle - \langle n|\hat{a}^\dagger\hat{a}^\dagger|n\rangle + 2\langle n|\hat{a}^\dagger\hat{a}|n\rangle + \langle n|n\rangle\right) \overset{(4.57)}{=} \\
&= \frac{p_0^2}{2}\left(-\langle n|\hat{a}\hat{a}|n\rangle - \langle n|\hat{a}^\dagger\hat{a}^\dagger|n\rangle + 2\langle n|\hat{N}|n\rangle + 1\right) \overset{(4.67,4.70)}{=} \\
&= \frac{p_0^2}{2}\left(-\sqrt{n}\sqrt{n-1}\langle n|n-2\rangle - \sqrt{n+1}\sqrt{n+2}\langle n|n+2\rangle + 2n\langle n|n\rangle + 1\right) \\
&= p_0^2\left(n+\frac{1}{2}\right) \overset{(4.15)}{=} m\hbar\omega\left(n+\frac{1}{2}\right)
\end{aligned} \tag{4.82}$$

Thus, for the variance of position and momentum, we obtain:

$$\langle(\Delta\hat{x})^2\rangle = \langle\hat{x}^2\rangle - \langle\hat{x}\rangle^2 \overset{(4.81,4.79)}{=} \frac{\hbar}{m\omega}\left(n+\frac{1}{2}\right)$$
$$\langle(\Delta\hat{p})^2\rangle = \langle\hat{p}^2\rangle - \langle\hat{p}\rangle^2 \overset{(4.82,4.80)}{=} m\hbar\omega\left(n+\frac{1}{2}\right)$$

The square root of each variance represents the standard deviation; forming the product Δx and Δp we finally obtain the following expression for the Heisenberg uncertainty relation:

$$\Delta x \Delta p = \sqrt{\langle(\Delta\hat{x})^2\rangle}\sqrt{\langle(\Delta\hat{p})^2\rangle} = \hbar\left(n+\frac{1}{2}\right) \tag{4.83}$$

If the harmonic quantum oscillator is in the ground state $|0\rangle$ (with $n = 0$), it possesses the minimum uncertainty according to (4.83):

$$\Delta x \Delta p = \frac{\hbar}{2} \tag{4.84}$$

4.3 Time-Evolved Oscillator States

4.3.1 Time-Dependent Oscillator Wave Functions

Let's now investigate the time-dependent, one-dimensional Schrödinger equation for a particle in the harmonic oscillator potential $V(x)$:

$$\hat{H}(x)\psi_n(x,t) = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2}m\omega^2x^2\right)\psi_n(x,t) = i\hbar\frac{\partial}{\partial t}\psi_n(x,t) \tag{4.85}$$

The Hamiltonian operator $\hat{H}$ is not explicitly time-dependent. Therefore, we can apply the separation Ansatz $\psi_n(x,t) = \phi_n(x)\varphi_n(t)$ as described in Chapter 2.1.4. For the time-dependent part, we have the following differential equation according to (2.35) and (4.36):

$$i\hbar\frac{d}{dt}\varphi_n(t) = E_n\varphi_n(t) \overset{(4.36)}{=} \hbar\omega\left(n+\frac{1}{2}\right)\varphi_n(t) \tag{4.86}$$

This differential equation can, as shown in (2.36), be easily solved, and we obtain the time evolution of the eigenstates of the harmonic oscillator:

$$\psi_n(x,t) \overset{(2.36)}{=} \phi_n(x)\, e^{-iE_nt/\hbar} \overset{(4.36)}{=} \phi_n(x)\, e^{-i\omega\left(n+\frac{1}{2}\right)t} = \phi_n(x)\, e^{-i\omega nt}\, e^{-\frac{1}{2}i\omega t} \tag{4.87}$$

The time-dependent part leads, with increasing excitation of the system, to an ever faster oscillation of the phase of the stationary eigenstate. A general state in space (at a fixed time t_0) $\psi(x,t_0)$ can be constructed as a linear combination of the time-independent eigenstates of the quantum oscillator $\phi_n(x)$. The further time evolution of $\psi(x,t)$ for times $t > t_0$ then simply follows from each of these base states $\phi_n(x)$ evolving with its associated time development:

$$\psi(x,t) = \sum_{n=0}^{\infty} a_n\psi_n(x,t) = \sum_{n=0}^{\infty} a_n\phi_n(x)e^{-i\omega nt}\, e^{-\frac{1}{2}i\omega t} = e^{-\frac{1}{2}i\omega t}\sum_{n=0}^{\infty} a_n\phi_n(x)e^{-i\omega nt} \tag{4.88}$$

The expansion coefficients a_n can, as usual, be calculated according to the following relation:

$$a_n = \int_{-\infty}^{+\infty} dx'\, \phi_n^*(x')\psi(x',t) \tag{4.89}$$

We will use the above-found relationships to construct states that can behave like classical particles.

4.3.2 Coherent Glauber States

Motivation: Classical Oscillator Motion with Quantum Physics

In this chapter, we ask the question, which time-dependent wave functions in the harmonic oscillator potential come closest to the harmonic motion of a particle known from classical physics. Erwin Schrödinger already succeeded in constructing such states in 1926, which are today called "coherent states" or "Glauber states" (after Roy Glauber). As we will see, these states correspond to Gaussian wave packets that run back and forth in the harmonic potential and maintain the minimal position and momentum uncertainty at all times.

The desired *coherent* or *Glauber state* is referred to as $|\alpha\rangle$ and corresponds to a superposition of time-dependent eigenstates of the harmonic oscillator. The wave packets resulting from this superposition should have the following properties:

- they follow the classical equations of motion, so that $\langle \hat{x} \rangle = C \cos(\omega t - \delta)$;
- they have minimal uncertainty for all times $t > 0$;
- the uncertainty is equally distributed over y and p_y.

As a starting point, we consider how $x(t)$ and $p(t)$ of a particle behave in a classical harmonic oscillator (assuming, without loss of generality, a phase $\varphi = 0$ for $x(t)$ from (4.4)):

$$\begin{aligned} x(t) &= A\cos(\omega t) = A\,\mathrm{Re}(e^{i\omega t}) \\ p(t) &= -B\sin(\omega t) = B\cos(\frac{\pi}{2} + \omega t) = B\,\mathrm{Re}(\underbrace{e^{i\pi/2}}_{i}\, e^{i\omega t}) = B\,\mathrm{Re}(ie^{i\omega t}) \end{aligned} \tag{4.90}$$

In the second line, we use the identity $-\sin(x) = \cos(x + \pi/2)$. After this transformation, we recognize that the momentum $p(t)$ is shifted by the phase $e^{i\pi/2} = i$ compared to the position $x(t)$. Let us now formulate the condition that the sought quantum states $|\phi_\alpha\rangle$ must satisfy for the quantum mechanical uncertainty to be equally distributed over the (dimensionless) quantities y and p_y:

$$(\hat{y} - \langle \hat{y} \rangle)\,|\phi_\alpha\rangle = -i\,(\hat{p}_y - \langle \hat{p}_y \rangle)\,|\phi_\alpha\rangle \tag{4.91}$$

The factor $-i = e^{-i\pi/2}$ on the right-hand side of equation (4.91) serves to compensate for the phase shift of $e^{i\pi/2} = i$ between p_y and y. As p_y and y are dimensionless, no further pre-factors are necessary. By transforming (4.91), we arrive at an eigenvalue equation that the desired state $|\alpha\rangle$ must satisfy:

$$(\hat{y} + i\hat{p}_y)\,|\phi_\alpha\rangle = (\langle \hat{y} \rangle + i\,\langle \hat{p}_y \rangle)\,|\phi_\alpha\rangle \tag{4.92}$$

The right-hand side of (4.92) corresponds to a complex eigenvalue. Substituting the definitions (4.52) and (4.53) into (4.92), we obtain:

$$\begin{aligned} (\langle \hat{y} \rangle + i\,\langle \hat{p}_y \rangle)\,|\phi_\alpha\rangle &= \left[\frac{1}{\sqrt{2}}\left(\hat{a} + \hat{a}^\dagger\right) + \frac{1}{\sqrt{2}}\left(\hat{a} - \hat{a}^\dagger\right)\right]|\phi_\alpha\rangle = \\ &= \frac{1}{\sqrt{2}}\left(\hat{a} + \hat{a}^\dagger + \hat{a} - \hat{a}^\dagger\right)|\phi_\alpha\rangle = \\ &= \sqrt{2}\hat{a}\,|\phi_\alpha\rangle \end{aligned}$$

We only need to divide the above equation by $\sqrt{2}$. Then we see: when the lowering operator $\hat{a}$ acts on the Glauber state $|\phi_\alpha\rangle$, it can be expressed in terms of the corresponding expectation values of the position and momentum operators, which allows us to define an eigenvalue α:

$$\hat{a}\,|\phi_\alpha\rangle = \frac{1}{\sqrt{2}}\,(\langle \hat{y} \rangle + i\,\langle \hat{p}_y \rangle)\,|\phi_\alpha\rangle = \alpha\,|\phi_\alpha\rangle$$

We are looking for states $|\phi_\alpha\rangle$ that are eigenstates of the lowering operator $\hat{a}$ and thus satisfy the following eigenvalue equation:

$$\hat{a}\,|\phi_\alpha\rangle = \alpha\,|\phi_\alpha\rangle \tag{4.93}$$

The complex-conjugate version of the eigenvalue equation also holds analogously:

$$\langle\phi_\alpha|\,\hat{a}^\dagger = \langle\phi_\alpha|\,\alpha^* \tag{4.94}$$

The solution for the above eigenvalue problem (4.92) at a fixed time $t = 0$ is:

$$|\phi_\alpha\rangle_{t=0} = N\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\,|n\rangle \overset{(4.73)}{=} N\sum_{n=0}^{\infty}\frac{(\alpha\hat{a}^\dagger)^n}{n!}\,|0\rangle \tag{4.95}$$

For the latter expression, we use that $|n\rangle$ can always be expressed by the ground state through the raising operator. For the normalization factor N, we calculate:

$$\begin{aligned}
\frac{1}{N^2} &\overset{!}{=} \langle\phi_\alpha|\phi_\alpha\rangle_{t=0} \overset{(4.95)}{=} \\
&= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{\alpha^n(\alpha^*)^m}{\sqrt{m!n!}}\,\langle m|n\rangle = \\
&= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{\alpha^n(\alpha^*)^m}{\sqrt{m!n!}}\delta_{nm} = \\
&= \sum_{n=0}^{\infty}\frac{(\alpha^*\alpha)^n}{n!} = \\
&= \sum_{n=0}^{\infty}\frac{|\alpha|^{2n}}{n!} = e^{|\alpha|^2}
\end{aligned}$$

Rearranging and taking the square root yields the finished normalization, depending on the magnitude of the eigenvalue:

$$N = e^{-\frac{1}{2}|\alpha|^2} \tag{4.96}$$

We can insert this normalization factor into (4.95), and thus express the normalized Glauber state $|\phi_\alpha\rangle_{t=0}$ at time $t = 0$ as follows:

$$|\phi_\alpha\rangle_{t=0} = e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\,|n\rangle = e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{(\alpha\hat{a}^\dagger)^n}{n!}\,|0\rangle \tag{4.97}$$

Letting $\hat{a}$ act on $|\phi_\alpha\rangle_{t=0}$, we see that it is indeed an eigenstate of the lowering operator:

$$\begin{aligned}
\hat{a}\,|\phi_\alpha\rangle_{t=0} &= \hat{a}\left(e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\,|n\rangle\right) = \\
&= e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\hat{a}\,|n\rangle \overset{(4.70)}{=} \\
&= e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\sqrt{n}\,|n-1\rangle = \\
&= e^{-\frac{1}{2}|\alpha|^2}\sum_{n=1}^{\infty}\frac{\alpha\alpha^{n-1}}{\sqrt{(n-1)!}}\,|n-1\rangle = \\
&= \alpha\,e^{-\frac{1}{2}|\alpha|^2}\sum_{m=1}^{\infty}\frac{\alpha^{m-1}}{\sqrt{(m-1)!}}\,|m-1\rangle = \quad |\,m = n+1 \\
&= \alpha e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\,|n\rangle = \\
&= \alpha\,|\phi_\alpha\rangle_{t=0}
\end{aligned} \tag{4.98}$$

All eigenstates of the lowering operator are normalized by construction, so it always holds:

$$\langle\phi_\alpha|\phi_\alpha\rangle = 1 \tag{4.99}$$

However, two arbitrary, different eigenstates $|\alpha\rangle$ and $|\beta\rangle$ of the lowering operator are *not* orthogonal to each other:

$$\langle\phi_\alpha|\phi_\beta\rangle \neq \delta_{\phi_\alpha\beta} \tag{4.100}$$

This was expected because $\hat{a}$ is not a Hermitian operator. Let's verify now whether the requirement of minimal uncertainty in position and momentum is met. For this, we need to calculate the first and second moments of x and p. Let's start with $\langle\hat{x}\rangle$:

$$\begin{aligned}
\langle\hat{x}\rangle = \langle\phi_\alpha|\hat{x}|\phi_\alpha\rangle &\overset{(4.48)}{=} \frac{x_0}{\sqrt{2}}\langle\phi_\alpha|\hat{a}+\hat{a}^\dagger|\phi_\alpha\rangle = \\
&= \frac{x_0}{\sqrt{2}}\Big(\langle\phi_\alpha|\hat{a}|\phi_\alpha\rangle + \langle\phi_\alpha|\hat{a}^\dagger|\phi_\alpha\rangle\Big) = \\
&= \frac{x_0}{\sqrt{2}}\Big(\alpha\langle\phi_\alpha|\phi_\alpha\rangle + \alpha^*\langle\phi_\alpha|\phi_\alpha\rangle\Big) = \frac{x_0}{\sqrt{2}}(\alpha+\alpha^*) = \\
&= \frac{x_0}{\sqrt{2}}\,2\,\mathrm{Re}(\alpha) = \sqrt{2}x_0\,\mathrm{Re}(\alpha)
\end{aligned} \tag{4.101}$$

For the expectation value $\langle\hat{p}\rangle$, a completely equivalent calculation can be performed:

$$\begin{aligned}
\langle\hat{p}\rangle = \langle\phi_\alpha|\hat{p}|\phi_\alpha\rangle &\overset{(4.49)}{=} -\frac{i}{\sqrt{2}}\frac{\hbar}{x_0}\langle\phi_\alpha|\hat{a}-\hat{a}^\dagger|\phi_\alpha\rangle = \\
&= -\frac{i}{\sqrt{2}}\frac{\hbar}{x_0}\Big(\langle\phi_\alpha|\hat{a}|\phi_\alpha\rangle - \langle\phi_\alpha|\hat{a}^\dagger|\phi_\alpha\rangle\Big) = \\
&= -\frac{i}{\sqrt{2}}\frac{\hbar}{x_0}\Big(\alpha\langle\phi_\alpha|\phi_\alpha\rangle - \alpha^*\langle\phi_\alpha|\phi_\alpha\rangle\Big) = -\frac{i}{\sqrt{2}}\frac{\hbar}{x_0}(\alpha-\alpha^*) = \\
&= \frac{1}{\sqrt{2}}\frac{\hbar}{x_0}\,2\,\mathrm{Im}(\alpha) = \sqrt{2}p_0\,\mathrm{Im}(\alpha)
\end{aligned} \tag{4.102}$$

To calculate the variance of position and momentum, we also need the second moments $\langle\hat{x}^2\rangle$ and $\langle\hat{p}^2\rangle$. Let's consider the second moment of position:

$$\begin{aligned}
\langle\hat{x}^2\rangle = \langle\alpha|\hat{x}^2|\alpha\rangle &\overset{(4.48)}{=} \\
&= \frac{x_0^2}{2}\langle\alpha|(\hat{a}+\hat{a}^\dagger)(\hat{a}+\hat{a}^\dagger)|\alpha\rangle = \\
&= \frac{x_0^2}{2}\langle\alpha|\hat{a}\hat{a}+\hat{a}^\dagger\hat{a}^\dagger+\hat{a}\hat{a}^\dagger+\hat{a}^\dagger\hat{a}|\alpha\rangle \overset{(4.56)}{=} \\
&= \frac{x_0^2}{2}\langle\alpha|\hat{a}\hat{a}+\hat{a}^\dagger\hat{a}^\dagger+1+\hat{a}^\dagger\hat{a}+\hat{a}^\dagger\hat{a}|\alpha\rangle = \\
&= \frac{x_0^2}{2}\Big(\langle\alpha|\hat{a}\hat{a}|\alpha\rangle + \langle\alpha|\hat{a}^\dagger\hat{a}^\dagger|\alpha\rangle + 2\langle\alpha|\hat{a}^\dagger\hat{a}|\alpha\rangle + \langle\alpha|\alpha\rangle\Big) = \\
&= \frac{x_0^2}{2}\Big(\alpha^2 + (\alpha^*)^2 + 2\alpha\alpha^* + 1\Big) = \\
&= \frac{x_0^2}{2}\Big((\alpha+\alpha^*)^2 + 1\Big)
\end{aligned} \tag{4.103}$$

For the second moment of momentum, a similar calculation applies:

$$\begin{aligned}
\langle\hat{p}^2\rangle = \langle\phi_\alpha|\hat{p}^2|\phi_\alpha\rangle &\overset{(4.49)}{=} \\
&= -\frac{1}{2}\frac{\hbar^2}{x_0^2}\langle\phi_\alpha|(\hat{a}-\hat{a}^\dagger)(\hat{a}-\hat{a}^\dagger)|\phi_\alpha\rangle =
\end{aligned}$$

$$
\begin{aligned}
&= -\frac{1}{2}\frac{\hbar^2}{x_0^2}\left\langle\phi_\alpha|\hat{a}\hat{a} + \hat{a}^\dagger\hat{a}^\dagger - \hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a}|\phi_\alpha\right\rangle \overset{(4.56)}{=} \\
&= -\frac{1}{2}\frac{\hbar^2}{x_0^2}\left\langle\phi_\alpha|\hat{a}\hat{a} + \hat{a}^\dagger\hat{a}^\dagger - 1 - \hat{a}^\dagger\hat{a} - \hat{a}^\dagger\hat{a}|\phi_\alpha\right\rangle = \\
&= -\frac{1}{2}\frac{\hbar^2}{x_0^2}\Big(\left\langle\phi_\alpha|\hat{a}\hat{a}|\phi_\alpha\right\rangle + \left\langle\phi_\alpha|\hat{a}^\dagger\hat{a}^\dagger|\phi_\alpha\right\rangle - 2\left\langle\phi_\alpha|\hat{a}^\dagger\hat{a}|\phi_\alpha\right\rangle - \left\langle\phi_\alpha|\phi_\alpha\right\rangle\Big) = \\
&= -\frac{1}{2}\frac{\hbar^2}{x_0^2}\Big(\alpha^2 + (\alpha^*)^2 - 2\alpha\alpha^* - 1\Big) = \\
&= -\frac{1}{2}\frac{\hbar^2}{x_0^2}\Big((\alpha - \alpha^*)^2 - 1\Big)
\end{aligned}
\tag{4.104}
$$

Thus, we can finally calculate the variance for position and momentum:

$$
\begin{aligned}
\left\langle(\Delta\hat{x})^2\right\rangle &= \left\langle\hat{x}^2\right\rangle - \left\langle\hat{x}\right\rangle^2 \overset{(4.103,4.101)}{=} \\
&= \frac{x_0^2}{2}\Big((\alpha + \alpha^*)^2 + 1 - (\alpha + \alpha^*)^2\Big) = \\
&= \frac{x_0^2}{2} \overset{(4.14)}{=} \frac{1}{2}\frac{\hbar}{m\omega}
\end{aligned}
\tag{4.105}
$$

$$
\begin{aligned}
\left\langle(\Delta\hat{p})^2\right\rangle &= \left\langle\hat{p}^2\right\rangle - \left\langle\hat{p}\right\rangle^2 \overset{(4.104,4.102)}{=} \\
&= \frac{1}{2}\frac{\hbar^2}{x_0^2}\Big(1 - (\alpha - \alpha^*)^2 + (\alpha - \alpha^*)^2\Big) = \\
&= \frac{1}{2}\frac{\hbar^2}{x_0^2} \overset{(4.14)}{=} \frac{1}{2}\hbar m\omega
\end{aligned}
\tag{4.106}
$$

Finally, we obtain the following expression for the Heisenberg uncertainty relation:

$$
\Delta x\Delta p = \sqrt{\left\langle(\Delta\hat{x})^2\right\rangle}\sqrt{\left\langle(\Delta\hat{p})^2\right\rangle} = \sqrt{\frac{\hbar}{2m\omega}}\sqrt{\frac{\hbar m\omega}{2}} = \frac{\hbar}{2} \tag{4.107}
$$

We recognize that this corresponds to the same uncertainty as in the ground state of the harmonic oscillator (4.84). It corresponds – as required – to the *minimal* uncertainty.

Time Evolution of Glauber States So far, we have only considered coherent states at a fixed time $t = 0$. If we drop this assumption, we must take into account the time evolution $e^{-i\omega nt}\, e^{-\frac{1}{2}i\omega t}$ of the eigenstates $|n\rangle$ derived in (4.87). To do this, we rewrite the expression for the time-independent state from (4.97) and replace $|n\rangle$ with $e^{-i\omega nt}\, e^{-\frac{1}{2}i\omega t}\,|n\rangle$. In addition, we rename the eigenvalue α to α_0 to indicate that it is the eigenvalue at time $t = 0$.

$$
\begin{aligned}
|\alpha(t)\rangle &= e^{-\frac{1}{2}|\alpha_0|^2}\sum_{n=0}^{\infty}\frac{\alpha_0^n}{\sqrt{n!}}\, e^{-i\omega nt}e^{-\frac{1}{2}i\omega t}\,|n\rangle = \\
&= e^{-\frac{1}{2}i\omega t}e^{-\frac{1}{2}|\alpha_0|^2}\sum_{n=0}^{\infty}\frac{\alpha_0^n}{\sqrt{n!}}\left(e^{-i\omega t}\right)^n|n\rangle = \\
&= e^{-\frac{1}{2}i\omega t}e^{-\frac{1}{2}|\alpha_0|^2}\sum_{n=0}^{\infty}\frac{1}{\sqrt{n!}}\left(\alpha_0\, e^{-i\omega t}\right)^n|n\rangle = \quad \big|\ \alpha(t) = \alpha_0\, e^{-i\omega t} \\
&= e^{-\frac{1}{2}i\omega t}e^{-\frac{1}{2}|\alpha_0|^2}\sum_{n=0}^{\infty}\frac{\alpha(t)^n}{\sqrt{n!}}\,|n\rangle
\end{aligned}
$$

In the final step, we have defined $\alpha(t) = \alpha_0\, e^{-i\omega t}$; in a sense, this is the time evolution of the eigenvalue α_0. However, since $|\alpha_0|^2$ is equal to $|\alpha_0\, e^{-i\omega t}|^2$, we can also replace $|\alpha_0|^2$ with $|\alpha(t)|^2$

in the equation above:

$$|\phi_\alpha(t)\rangle = e^{-\frac{1}{2}i\omega t} e^{-\frac{1}{2}|\alpha(t)|^2} \sum_{n=0}^{\infty} \frac{\alpha(t)^n}{\sqrt{n!}} |n\rangle$$

This expression is, apart from the phase factor $e^{-\frac{1}{2}i\omega t}$, identical to the expression for $|\phi_\alpha\rangle_{t=0}$ in (4.97), if we replace α with $\alpha(t)$. We can therefore write:

$$|\phi_\alpha(t)\rangle = e^{-\frac{1}{2}i\omega t} |\phi_\alpha\rangle_{t=0} \tag{4.108}$$

We have just shown that the time-dependent Glauber state $|\phi_\alpha(t)\rangle$ can be expressed at any time t in terms of the expression for the time-independent state $|\phi_\alpha\rangle_{t=0}$, by multiplying it with $e^{-\frac{1}{2}i\omega t}$ and substituting α with $\alpha(t) = \alpha_0\, e^{-i\omega t}$. This means that all previous proofs remain valid since we have always calculated with a general (arbitrary) value for α, and the phase pre-factor $e^{-\frac{1}{2}i\omega t}$ does not play a role. Therefore, for instance, the minimal uncertainty (4.107) remains preserved for all times t. The wave packet does not "disperse".

Analogy to the Classical Oscillator We can now calculate how the expectation value of the position $\langle\hat{x}(t)\rangle$ changes over time. For this, we use expression (4.101), replacing α with $\alpha(t)$:

$$\langle\hat{x}\rangle = \frac{x_0}{\sqrt{2}}\, 2\,\mathrm{Re}(\alpha(t)) \overset{(4.108)}{=} \frac{x_0}{\sqrt{2}}\, 2\,\mathrm{Re}(\alpha_0\, e^{-i\omega t}) \tag{4.109}$$

The complex number α_0 can also be represented by its magnitude $|\alpha_0|$ and its phase φ, so $\alpha_0 = |\alpha_0| e^{i\varphi}$. Substituting this into (4.109), we get:

$$\langle\hat{x}\rangle = \frac{x_0}{\sqrt{2}}\, 2\,\mathrm{Re}\left(|\alpha_0|\, e^{i\varphi}\, e^{-i\omega t}\right) = \sqrt{2}\, x_0\, |\alpha_0| \cos\left(\omega t - \varphi\right) \tag{4.110}$$

The coherent wave packet thus swings in the harmonic oscillator potential like a classical particle; just as we derived it for the classical harmonic oscillator in (4.4).

5 Angular Momentum

Motivation: The Orbital Angular Momentum (and Other Angular Momenta)

In this chapter, we will discuss the important quantum mechanical observable of *angular momentum* $\mathbf{L}$. Using the correspondence principle, we will first derive a quantum mechanical angular momentum operator $\hat{\mathbf{L}}$ from the classical definition of $\mathbf{L}$. The observable thus defined is called the *orbital angular momentum* because it is the quantum mechanical equivalent of the classical angular momentum exhibited, for example, by electrons orbiting the atomic nucleus.

Since angular momentum is a vector quantity, the angular momentum operator $\hat{\mathbf{L}}$ is consequently a vector operator. Subsequently, we will determine a set of commutation relations that the angular momentum operator $\hat{\mathbf{L}}$ satisfies. It turns out that these relations are so general that we can call *any* operator that satisfies them an angular momentum operator.

Indeed, in quantum mechanics, there are other angular momentum operators besides the orbital angular momentum operator $\hat{\mathbf{L}}$, such as the spin operator $\hat{\mathbf{S}}$, which we will examine in more detail later.

5.1 Angular Momentum Operator

The classical angular momentum is defined as $\mathbf{L} = \mathbf{r} \times \mathbf{p}$; from the correspondence principle, it follows for the angular momentum operator that $\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$. In position representation, for the position and momentum operators it holds that $\hat{\mathbf{r}} \to \mathbf{r}$ and $\hat{\mathbf{p}} \to -i\hbar\boldsymbol{\nabla}$. Consequently, we write in position space $\{|\mathbf{r}\rangle\}$:

$$\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}} \xrightarrow{\{|\mathbf{r}\rangle\}} -i\hbar\, \mathbf{r} \times \boldsymbol{\nabla} = -i\hbar \begin{pmatrix} x \\ y \\ z \end{pmatrix} \times \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} \tag{5.1}$$

By explicitly carrying out the cross product, we can find a position representation for each individual component of the angular momentum:

$$\hat{\mathbf{L}} = \begin{pmatrix} \hat{L}_x \\ \hat{L}_y \\ \hat{L}_z \end{pmatrix} = \begin{pmatrix} \hat{y}\hat{p}_z - \hat{z}\hat{p}_y \\ \hat{z}\hat{p}_x - \hat{x}\hat{p}_z \\ \hat{x}\hat{p}_y - \hat{y}\hat{p}_x \end{pmatrix} \xrightarrow{\{|\mathbf{r}\rangle\}} -i\hbar \begin{pmatrix} y\partial_z - z\partial_y \\ z\partial_x - x\partial_z \\ x\partial_y - y\partial_x \end{pmatrix} \tag{5.2}$$

In the course of this chapter, no strict distinction is made between the abstract operator notation and the expression in the position basis; instead of "$\to$", "$=$" is written directly.

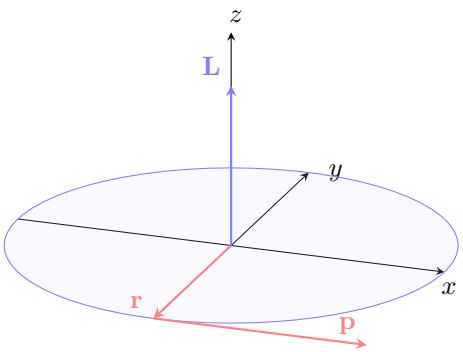


Fig. 21: Representation of the classical angular momentum.

$\hat{\mathbf{L}}$ is also referred to as a vector operator. In index notation, the cross product can be expressed using the Levi-Civita tensor ε_{ijk}:

$$\hat{L}_i = \varepsilon_{ijk}\hat{r}_j\hat{p}_k = -i\hbar\varepsilon_{ijk}r_j\partial_k \tag{5.3}$$

In addition to the individual components, a magnitude operator of the angular momentum $\hat{\mathbf{L}}^2$ can also be defined:

$$\hat{\mathbf{L}}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 \tag{5.4}$$

In-Depth: Index Notation

General: In index notation, instead of writing a vector $\mathbf{r} = (r_x, r_y, r_z)^T$, one simply writes r_i, where the index i stands for the i-th element of a vector $\mathbf{r}$, so: $r_1 = r_x, r_2 = r_y, r_3 = r_z$. Of course, any other letter can be used as a placeholder instead of i.

$$(\mathbf{r})_i \hat{=} r_i$$

(Einstein's) Summation Convention: If the same index appears exactly twice in a product term (*"is saturated"*), it is summed over, for example:

$$a_i r_i = \sum_{i=1}^{3} a_i r_i = a_1 r_1 + a_2 r_2 + a_3 r_3$$

The above example corresponds to the inner product of vectors $\mathbf{a}$ and $\mathbf{r}$: $a_i r_i = \mathbf{a} \cdot \mathbf{r}$. Two variables with the same index in a product term represent the inner product, and thus a scalar.

Free and Saturated Indices: If a product term consists only of "double" (*"saturated"*) indices, then this term represents a scalar expression. For example: $a_i b_i c_j d_j$. However, in all product terms that are additively linked in an equation, only saturated indices must appear (adding a vector to a scalar is not possible). Conversely: If a product term contains a "free" index (i.e., an index that appears only once), then this product term represents a vector. The left and right sides of the equation, along with all the terms involved, must have the same free index, for example: $a_j = b_i c_i d_j + b_i c_i a_j$.

Kronecker Delta: The Kronecker delta δ_{ij} is defined as follows:

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \tag{5.5}$$

Levi-Civita Symbol: The Levi-Civita symbol is defined as follows:

$$\varepsilon_{ijk} = \begin{cases} 1, \text{ if } ijk \text{ is an even permutation of } 123 \\ -1, \text{ if } ijk \text{ is an odd permutation of } 123 \\ 0, \text{ otherwise (if at least two indices are the same)} \end{cases} \tag{5.6}$$

Derivative: The derivative operator in index notation is defined as follows:

$$\partial_i = (\nabla)_i = \frac{\partial}{\partial r_i} \iff \partial_1 = \frac{\partial}{\partial x}, \quad \partial_2 = \frac{\partial}{\partial y}, \quad \partial_3 = \frac{\partial}{\partial z} \tag{5.7}$$

Calculation Rules: Here's a list of calculation rules, where the occasionally used letter n stands for the number of dimensions (in our case $n = 3$). We mainly consider important rules for δ_{ij}, ε_{ijk}, and ∂_i:

$$\begin{aligned}
\delta_{ij} &= \delta_{ji} \\
\delta_{ii} &= n \\
\delta_{ij}\delta_{jk} &= \delta_{ik} \\
\delta_{ij}x_j &= x_i \\
\varepsilon_{ijk}\varepsilon_{klm} &= \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl} \\
\varepsilon_{ijk}\varepsilon_{ijl} &= 2\delta_{kl} \\
\varepsilon_{ijk}\varepsilon_{ijk} &= 3! = 6 \\
\varepsilon_{ijk}\delta_{ij} &= \varepsilon_{ijk}\delta_{ik} = \varepsilon_{ijk}\delta_{jk} = 0 \qquad (5.8) \\
\varepsilon_{ijk} &= \varepsilon_{jki} = \varepsilon_{kij} \qquad (5.9) \\
\varepsilon_{ijk} &= -\varepsilon_{ikj} = -\varepsilon_{jik} \\
\varepsilon_{iik} &= \varepsilon_{iki} = \varepsilon_{ikk} = \varepsilon_{iii} = 0 \\
\partial_i x_j &= \delta_{ij} \\
\partial_i x_i &= \delta_{ii} = n
\end{aligned}$$

Connection between Vector and Index Notation:

Operation	Vector Notation	Index Notation
Dot Product	$\mathbf{x} \cdot \mathbf{y}$	$x_i y_i$
Magnitude Squared	$\|\|\mathbf{x}\|\|^2 = \mathbf{x} \cdot \mathbf{x}$	$x_i x_i$
Gradient (scalar)	$\nabla f(\mathbf{r})$	$\partial_i f(\mathbf{r})$
Divergence	$\nabla \cdot \mathbf{x}$	$\partial_i x_i$
Cross Product	$\mathbf{x} \times \mathbf{y}$	$\varepsilon_{ijk} x_j y_k$
Curl	$\nabla \times \mathbf{x}$	$\varepsilon_{ijk} \partial_j x_k$
Laplacian (scalar)	$\nabla^2 f(\mathbf{r}) = \nabla \cdot (\nabla f(\mathbf{r}))$	$\partial_i \partial_i f(\mathbf{r})$
Vector Gradient	$\nabla \otimes \mathbf{x}$	$\partial_j x_i$
Laplacian (vectorial)	$\nabla^2 \mathbf{x} = \begin{pmatrix} \nabla^2 x_1 \\ \nabla^2 x_2 \\ \nabla^2 x_3 \end{pmatrix}$	$\partial_j \partial_j x_i$

We now turn to the question of whether the angular momentum operator $\hat{\mathbf{L}}$ satisfies the Hermiticity condition. To do this, we need to prove that each component $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ is individually Hermitian. As a starting point for the proof, we take equation (5.3) and consider in the last step that $\hat{p}_k$ and $\hat{r}_j$ are Hermitian.

$$\hat{L}_i^\dagger = (\varepsilon_{ijk}\hat{r}_j\hat{p}_k)^\dagger = \varepsilon_{ijk}(\hat{r}_j\hat{p}_k)^\dagger = \varepsilon_{ijk}\hat{p}_k^\dagger\hat{r}_j^\dagger = \varepsilon_{ijk}\hat{p}_k\hat{r}_j \qquad (5.10)$$

To complete the proof of the Hermiticity of $\hat{L}_i$, we would need to reverse the order of $\hat{p}_k$ and $\hat{r}_j$. However, position and momentum operators do *not* commute per se, but satisfy the canonical commutation relation $[\hat{r}_i, \hat{p}_j] = i\hbar\delta_{ij}$ according to (3.133). This implies: $\hat{r}_j\hat{p}_k - \hat{p}_k\hat{r}_j = i\hbar\delta_{jk}$, and thus $\hat{p}_k\hat{r}_j = \hat{r}_j\hat{p}_k - i\hbar\delta_{jk}$. Substituting this into (5.10), we obtain:

$$\hat{L}_i^\dagger = \varepsilon_{ijk}\left(\hat{r}_j\hat{p}_k - i\hbar\delta_{jk}\right) = \varepsilon_{ijk}\hat{r}_j\hat{p}_k - i\hbar\varepsilon_{ijk}\delta_{jk} \overset{(5.3)}{=} \hat{L}_i - i\hbar\varepsilon_{ijk}\delta_{jk} = L_i \ \square \qquad (5.11)$$

Considering the calculation rule (5.8), the last term vanishes, and we have successfully proven that all components of the operator, and thus $\hat{\mathbf{L}}$ as a whole, are Hermitian:

$$\hat{\mathbf{L}}^\dagger = \hat{\mathbf{L}} \quad \text{and} \quad \hat{L}_i^\dagger = \hat{L}_i \qquad (5.12)$$

For a Hermitian operator $\hat{A}$, it holds that any powers of it are also Hermitian: $(\hat{A}^2)^\dagger = \hat{A}^\dagger\hat{A}^\dagger = \hat{A}\hat{A} = \hat{A}^2$; it follows immediately that the magnitude operator of the angular momentum is also Hermitian:

$$(\hat{\mathbf{L}}^2)^\dagger = \hat{\mathbf{L}}^2 \quad \text{and} \quad (\hat{L}_i^2)^\dagger = \hat{L}_i^2 \tag{5.13}$$

5.1.1 Commutator Relations

The angular momentum operator satisfies a special commutator relation with both the position and the momentum operators. Let's first consider the position operator $\hat{\mathbf{r}}$:

$$[\hat{L}_i, \hat{r}_j] = i\hbar\varepsilon_{ijk}\hat{r}_k \tag{5.14}$$

Example: Commutator between $\hat{L}_x$ and $\hat{\mathbf{r}}$

We first use the relation (5.14). This allows the commutator relationships between $\hat{L}_x$ and the components of the position operator $\hat{x}$, $\hat{y}$, and $\hat{z}$ to be determined very easily (the same calculation steps naturally also apply to $\hat{L}_y$ and $\hat{L}_z$):

$$[\hat{L}_x, \hat{x}] \equiv [\hat{L}_1, \hat{r}_1] = i\hbar\varepsilon_{11k}\hat{r}_k = i\hbar\varepsilon_{111}\hat{r}_1 + i\hbar\varepsilon_{112}\hat{r}_2 + i\hbar\varepsilon_{113}\hat{r}_3 = 0 \tag{5.15}$$

$$[\hat{L}_x, \hat{y}] \equiv [\hat{L}_1, \hat{r}_2] = i\hbar\varepsilon_{12k}\hat{r}_k = i\hbar\varepsilon_{121}\hat{r}_1 + i\hbar\varepsilon_{122}\hat{r}_2 + i\hbar\varepsilon_{123}\hat{r}_3 = i\hbar\hat{r}_3 \equiv i\hbar\hat{z} \tag{5.16}$$

$$[\hat{L}_x, \hat{z}] \equiv [\hat{L}_1, \hat{r}_3] = i\hbar\varepsilon_{13k}\hat{r}_k = i\hbar\varepsilon_{131}\hat{r}_1 + i\hbar\varepsilon_{132}\hat{r}_2 + i\hbar\varepsilon_{133}\hat{r}_3 = -i\hbar\hat{r}_2 \equiv -i\hbar\hat{y} \tag{5.17}$$

To demonstrate that these relationships are indeed correct, we can also use the explicit representation (5.3) for $\hat{L}_i$. Note that the operators are only defined by their effect on a following state $\psi(x) \equiv \psi$, which must be implicitly included in the calculation. We explicitly carry out the calculation for $[\hat{L}_x, \hat{x}]$ and $[\hat{L}_x, \hat{y}]$ (where $[\hat{L}_x, \hat{z}]$ again follows through equivalent calculation steps):

$$\begin{aligned}
[\hat{L}_x, \hat{x}]\psi \equiv [\hat{L}_1, \hat{r}_1]\psi &= \hat{L}_1\hat{r}_1\psi - \hat{r}_1\hat{L}_1\psi = -i\hbar\left[\varepsilon_{1jk}r_j\partial_k\left(r_1\psi\right) - r_1\varepsilon_{1jk}r_j\partial_k\psi\right] = \\
&= -i\hbar\left[\varepsilon_{1jk}r_j\psi\partial_k r_1 + \varepsilon_{1jk}r_j r_1\partial_k\psi - r_1\varepsilon_{1jk}r_j\partial_k\psi\right] = \quad \left|\, \varepsilon_{1jk} = \varepsilon_{123} + \varepsilon_{132}\right. \\
&= -i\hbar\left[\varepsilon_{123}r_2\psi\partial_3 r_1 + \varepsilon_{132}r_3\psi\partial_2 r_1\right] = -i\hbar\left[r_2\psi\delta_{31} - r_3\psi\delta_{21}\right] = 0
\end{aligned}$$

As expected, we obtain the same result (5.15) in the position representation! For $[\hat{L}_x, \hat{y}]$, we proceed in an entirely analogous manner:

$$\begin{aligned}
[\hat{L}_x, \hat{y}]\psi \equiv [\hat{L}_1, \hat{r}_2]\psi &= \hat{L}_1\hat{r}_2\psi - \hat{r}_2\hat{L}_1\psi = -i\hbar\left[\varepsilon_{1jk}r_j\partial_k\left(r_2\psi\right) - r_2\varepsilon_{1jk}r_j\partial_k\psi\right] = \\
&= -i\hbar\left[\varepsilon_{1jk}r_j\psi\partial_k r_2 + \varepsilon_{1jk}r_j r_2\partial_k\psi - r_2\varepsilon_{1jk}r_j\partial_k\psi\right] = \quad \left|\, \varepsilon_{1jk} = \varepsilon_{123} + \varepsilon_{132}\right. \\
&= -i\hbar\left[\varepsilon_{123}r_2\psi\partial_3 r_2 + \varepsilon_{132}r_3\psi\partial_2 r_2\right] = -i\hbar\left[r_2\psi\delta_{32} - r_3\psi\delta_{22}\right] = \\
&= i\hbar r_3\psi = i\hbar\hat{r}_3\psi \equiv i\hbar\hat{z}\psi
\end{aligned}$$

The commutator between the angular momentum operator and the momentum operator looks quite identical:

$$[\hat{L}_i, \hat{p}_j] = i\hbar\varepsilon_{ijk}\hat{p}_k \tag{5.18}$$

Example: Commutator between $\hat{L}_z$ and $\hat{\mathbf{p}}$

We first use the relation (5.18). This allows the commutator relationships between $\hat{L}_z$ and the components of the momentum operator $\hat{p}_x$, $\hat{p}_y$, and $\hat{p}_z$ to again be determined very easily. Since we continue to remain in the position representation, at the end we substitute the respective momentum operator components $\hat{p}_i$ with their explicit position representation $\hat{p}_i = -i\hbar\partial_i$. Analogous to the position operator components, it follows:

$$[\hat{L}_z, \hat{p}_x] \equiv [\hat{L}_3, \hat{p}_1] = i\hbar\varepsilon_{31k}\hat{p}_k = i\hbar\varepsilon_{312}\hat{p}_2 = i\hbar\hat{p}_2 = i\hbar\left(-i\hbar\partial_2\right) = \hbar^2\partial_2 \tag{5.19}$$

$$[\hat{L}_z, \hat{p}_y] \equiv [\hat{L}_3, \hat{p}_2] = i\hbar\varepsilon_{32k}\hat{p}_k = i\hbar\varepsilon_{321}\hat{p}_1 = -i\hbar\hat{p}_1 = -i\hbar\left(-i\hbar\partial_1\right) = -\hbar^2\partial_1 \tag{5.20}$$

$$[\hat{L}_z, \hat{p}_z] \equiv [\hat{L}_3, \hat{p}_3] = i\hbar\varepsilon_{33k}\hat{p}_k = 0 \tag{5.21}$$

To demonstrate that these relationships hold, we can again use the explicit representation in the position basis (5.3) for $\hat{L}_i$ and $-i\hbar\partial_i$ for $\hat{p}_i$. The state ψ, on which the operators act, must be included in the calculation. We demonstrate the calculation explicitly for $[\hat{L}_z, \hat{p}_x]$ and $[\hat{L}_z, \hat{p}_y]$ (again the proof for $[\hat{L}_z, \hat{p}_y]$ works analogously). It holds:

$$\begin{aligned}
[\hat{L}_z, \hat{p}_x]\psi &\equiv [\hat{L}_3, \hat{p}_1]\psi = \hat{L}_3\hat{p}_1\psi - \hat{p}_1\hat{L}_3\psi = \\
&= \left(-i\hbar\varepsilon_{3jk}r_j\partial_k\right)\left(-i\hbar\partial_1\right)\psi - \left(-i\hbar\partial_1\right)\left(-i\hbar\varepsilon_{3jk}r_j\partial_k\right)\psi = \\
&= -\hbar^2\varepsilon_{3jk}r_j\partial_k\partial_1\psi + \hbar^2\varepsilon_{3jk}\partial_1\left(r_j\partial_k\psi\right) = \\
&= -\hbar^2\varepsilon_{3jk}r_j\partial_k\partial_1\psi + \hbar^2\varepsilon_{3jk}\left(\partial_1 r_j\right)\left(\partial_k\psi\right) + \hbar^2\varepsilon_{3jk}r_j\partial_k\partial_1\psi = \quad \left|\, \varepsilon_{3jk} = \varepsilon_{312} + \varepsilon_{321}\right. \\
&= \hbar^2\varepsilon_{312}\left(\partial_1 r_1\right)\left(\partial_2\psi\right) + \hbar^2\varepsilon_{321}\left(\partial_1 r_2\right)\left(\partial_1\psi\right) = \\
&= \hbar^2\varepsilon_{312}\delta_{11}\left(\partial_2\psi\right) + \hbar^2\varepsilon_{321}\delta_{12}\left(\partial_1\psi\right) = \hbar^2\partial_2\psi \equiv \hbar^2\frac{\partial\psi}{\partial y}
\end{aligned}$$

All calculation steps can be analogously transferred to the commutator $[\hat{L}_z, \hat{p}_y]$:

$$\begin{aligned}
[\hat{L}_z, \hat{p}_y]\psi &\equiv [\hat{L}_3, \hat{p}_2]\psi = \hat{L}_3\hat{p}_2\psi - \hat{p}_2\hat{L}_3\psi = \\
&= \left(-i\hbar\varepsilon_{3jk}r_j\partial_k\right)\left(-i\hbar\partial_2\right)\psi - \left(-i\hbar\partial_2\right)\left(-i\hbar\varepsilon_{3jk}r_j\partial_k\right)\psi = \\
&= -\hbar^2\varepsilon_{3jk}r_j\partial_k\partial_2\psi + \hbar^2\varepsilon_{3jk}\partial_2\left(r_j\partial_k\psi\right) = \\
&= -\hbar^2\varepsilon_{3jk}r_j\partial_k\partial_2\psi + \hbar^2\varepsilon_{3jk}\left(\partial_2 r_j\right)\left(\partial_k\psi\right) + \hbar^2\varepsilon_{3jk}r_j\partial_k\partial_2\psi = \quad \left|\, \varepsilon_{3jk} = \varepsilon_{312} + \varepsilon_{321}\right. \\
&= \hbar^2\varepsilon_{312}\left(\partial_2 r_1\right)\left(\partial_2\psi\right) + \hbar^2\varepsilon_{321}\left(\partial_2 r_2\right)\left(\partial_1\psi\right) = \\
&= \hbar^2\varepsilon_{312}\delta_{21}\left(\partial_2\psi\right) + \hbar^2\varepsilon_{321}\delta_{22}\left(\partial_1\psi\right) = -\hbar^2\partial_1\psi \equiv -\hbar^2\frac{\partial\psi}{\partial x}
\end{aligned}$$

In general, the following holds: If the components of an operator $\hat{\mathbf{V}}$ satisfy the following commutator relationship, then $\hat{\mathbf{V}}$ is a *vector operator*:

$$[\hat{L}_i, \hat{V}_j] = i\hbar\varepsilon_{ijk}\hat{V}_k \tag{5.22}$$

This relation expresses that the angular momentum operator $\hat{\mathbf{L}}$ is the generator of rotations. A vector operator is characterized by transforming like a vector under rotations. Vector operators are special cases of tensor operators whose matrix elements can be described by the Wigner-Eckart theorem. The commutator of $\hat{\mathbf{L}}$ with a scalar operator (5.30) vanishes (since scalar quantities are invariant under rotations). This topic and all the above claims will be discussed in more detail in the lecture `Quantum Theory II`.

Since the angular momentum operator $\hat{\mathbf{L}}$ is itself a vector operator, the following transformation

behavior must be valid according to (5.22):

$$[\hat{L}_i, \hat{L}_j] = i\hbar\varepsilon_{ijk}\hat{L}_k \tag{5.23}$$

The relationship (5.23) leads to an important conclusion: Two different components of the angular momentum operator $\hat{\mathbf{L}}$ never commute and therefore exclude simultaneous measurement. Later, we will recognize that this is the cause of angular momentum uncertainty!

Example: Commutator between $\hat{L}_x$ and $\hat{L}_y$

We first use the relationship (5.23), to calculate the commutator $[\hat{L}_x, \hat{L}_y]$:

$$[\hat{L}_x, \hat{L}_y] \equiv [\hat{L}_1, \hat{L}_2] = i\hbar\varepsilon_{12k}\hat{L}_k = i\hbar\varepsilon_{123}\hat{L}_3 = i\hbar\hat{L}_3 \equiv i\hbar\hat{L}_z$$

To show that this relation is valid, we now perform the calculation explicitly with the representation (5.3):

$$\begin{aligned}
[\hat{L}_x, \hat{L}_y]\psi \equiv [\hat{L}_1, \hat{L}_2]\psi &= \hat{L}_1\hat{L}_2\psi - \hat{L}_2\hat{L}_1\psi = \\
&= (-i\hbar\varepsilon_{1jk}r_j\partial_k)\,(-i\hbar\varepsilon_{2lm}r_l\partial_m)\,\psi - (-i\hbar\varepsilon_{2lm}r_l\partial_m)\,(-i\hbar\varepsilon_{1jk}r_j\partial_k)\,\psi = \\
&= \hbar^2\varepsilon_{1jk}\varepsilon_{2lm}\left[-r_j\partial_k\,(r_l\partial_m\psi) + r_l\partial_m\,(r_j\partial_k\psi)\right] = \\
&= \hbar^2\varepsilon_{1jk}\varepsilon_{2lm}\left[-r_j\,(\partial_k r_l)\,(\partial_m\psi) - r_j r_l\partial_m\partial_k\psi + r_l\,(\partial_m r_j)\,(\partial_k\psi) + r_l r_j\partial_m\partial_k\psi\right] = \\
&= -\hbar^2\varepsilon_{1jk}\varepsilon_{2lm}\delta_{kl}r_j\partial_m\psi + \hbar^2\varepsilon_{1jk}\varepsilon_{2lm}\delta_{mj}r_l\partial_k\psi = \\
&= -\hbar^2\varepsilon_{1jk}\varepsilon_{km2}r_j\partial_m\psi + \hbar^2\varepsilon_{k1j}\varepsilon_{j2l}r_l\partial_k\psi = \\
&= -\hbar^2(\delta_{1m}\delta_{j2} - \delta_{12}\delta_{jm})r_j\partial_m\psi + \hbar^2(\delta_{k2}\delta_{1l} - \delta_{kl}\delta_{12})r_l\partial_k\psi = \\
&= -\hbar^2\delta_{j2}r_j\delta_{1m}\partial_m\psi - \hbar^2\delta_{1l}r_l\delta_{k2}\partial_k\psi = \\
&= \hbar^2\,(-r_2\partial_1 + r_1\partial_2)\,\psi = \quad |\,\varepsilon_{321} = -1;\ \varepsilon_{312} = +1 \\
&= \hbar^2\,(\varepsilon_{321}r_2\partial_1 + \varepsilon_{312}r_1\partial_2)\,\psi = \\
&= \hbar^2\,(\varepsilon_{3jk}r_j\partial_k)\,\psi = \\
&= i\hbar\,(-i\hbar\varepsilon_{3jk}r_j\partial_k)\,\psi = i\hbar\hat{L}_z\psi
\end{aligned}$$

5.1.2 Polar and Axial Vector Operators

In this section, we consider how the angular momentum, by construction, behaves differently under symmetry transformations than position vectors $\mathbf{r}$ or momentum vectors $\mathbf{p}$. To do so, we first introduce the parity operator $\hat{\Pi}$, which will be discussed more thoroughly in the `Quantum Theory II` lecture. The parity operator $\hat{\Pi}$ performs a point reflection about the origin of the coordinate system. It acts on a general wavefunction in position representation as follows:

$$\hat{\Pi}\psi(\mathbf{r}) = \psi(-\mathbf{r}) \tag{5.24}$$

There are wavefunctions $\varphi_\pi(\mathbf{r})$ with positive parity, so that $\hat{\Pi}\varphi_\pi(\mathbf{r}) = \varphi_\pi(-\mathbf{r}) = \varphi_\pi(\mathbf{r})$, and those with negative parity, so that $\hat{\Pi}\varphi_\pi(\mathbf{r}) = \varphi_\pi(-\mathbf{r}) = -\varphi_\pi(\mathbf{r})$. Wavefunctions that behave this way are evidently eigenstates of the parity operator with an eigenvalue of plus or minus one. Therefore, we can write:

$$\hat{\Pi}\,|\varphi_\pi\rangle = p_\pi\,|\varphi_\pi\rangle \quad \text{with} \quad p_\pi = \pm 1 \tag{5.25}$$

Any arbitrary wavefunction can be decomposed into a part with positive parity and a part with negative parity (i.e., into the eigenstates $|\varphi_\pi\rangle$ of the parity operator $\hat{\Pi}$). The parity operator is both Hermitian and unitary; therefore, the following relationship holds:

$$\hat{\Pi} = \hat{\Pi}^\dagger = \hat{\Pi}^{-1} \tag{5.26}$$

Depending on the symmetry of a state $|\varphi_\pi\rangle$, the eigenvalue is either $p_\pi = 1$ (then we speak of even states) or $p_\pi = -1$ (then we speak of odd states). This principle can also be applied to operators. We will briefly discuss the symmetry behavior of the operators $\hat{\mathbf{r}}$, $\hat{\mathbf{p}}$, and $\hat{\mathbf{L}}$ here.

A *polar* vector (or, due to the correspondence principle, a vector operator) indicates a *direction* in space. Accordingly, when a space reflection occurs, the direction in space should also logically change. We are already familiar with two polar operators:

$$\begin{aligned} \hat{\Pi}\hat{\mathbf{r}}\hat{\Pi}^{-1} &= -\hat{\mathbf{r}} \\ \hat{\Pi}\hat{\mathbf{p}}\hat{\Pi}^{-1} &= -\hat{\mathbf{p}} \end{aligned} \tag{5.27}$$

What applies to the operators holds also for the associated polar wavefunctions:

$$\begin{aligned} \hat{\Pi}\,|\mathbf{r}\rangle &= -\,|\mathbf{r}\rangle \\ \hat{\Pi}\,|\mathbf{p}\rangle &= -\,|\mathbf{p}\rangle \end{aligned} \tag{5.28}$$

On the other hand, an *axial* vector (or an axial vector operator) indicates the *rotation sense* in space and does not change even when a space reflection occurs! For example, the angular momentum operator is subject to the following:

$$\hat{\Pi}\hat{\mathbf{L}}\hat{\Pi}^{-1} = +\hat{\mathbf{L}} \tag{5.29}$$

This can be easily verified by inserting $\mathbb{1} = \hat{\Pi}^{-1}\hat{\Pi}$:

$$\hat{\Pi}\hat{\mathbf{L}}\hat{\Pi}^{-1} = \hat{\Pi}\,(\hat{\mathbf{r}} \times \hat{\mathbf{p}})\,\hat{\Pi}^{-1} = \hat{\Pi}\hat{\mathbf{r}}\hat{\Pi}^{-1} \times \hat{\Pi}\hat{\mathbf{p}}\hat{\Pi}^{-1} = (-\hat{\mathbf{r}}) \times (-\hat{\mathbf{p}}) = \hat{\mathbf{L}} \quad \square$$

It should be noted here: The parity operator leads to a *point reflection* and is thus not to be confused with a *plane reflection*, which would lead to an entirely different symmetry behavior.

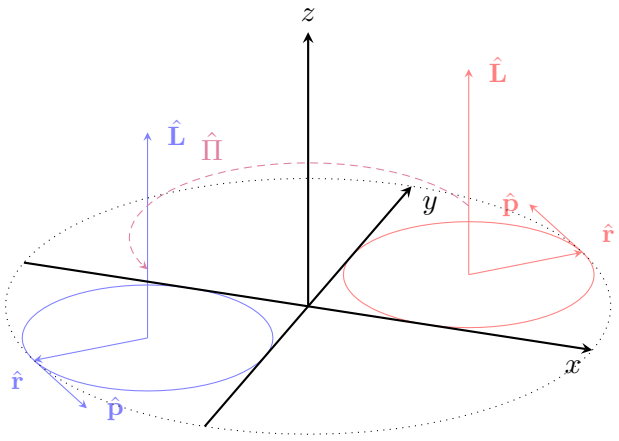


Fig. 22: Behavior of polar and axial vectors under a point reflection.

5.1.3 Commutator Relations with Scalar Operators

Similar to the commutator relationship between the angular momentum and vector operator $\hat{\mathbf{L}}$ and $\hat{\mathbf{V}}$ from (5.22), a relation can also be defined between the angular momentum operator and a *scalar operator* $\hat{S}$:

$$[\hat{L}_i, \hat{S}] = 0 \tag{5.30}$$

Such scalar operators are, for example, the squared forms of vector operators, such as the operator $\hat{\mathbf{L}}^2$ defined in (5.4), or the squared position operator $\hat{\mathbf{r}}^2 = \hat{x}^2 + \hat{y}^2 + \hat{z}^2$. In general, scalars behave the same under rotations as (5.30).

Example: Commutator between $\hat{L}_x$ and $\hat{\mathbf{r}}^2$

Here, we examine the commutator relationship between the angular momentum component in the x-direction and the magnitude of the position operator $\hat{\mathbf{r}}^2$. For the individual terms of $\hat{\mathbf{r}}^2$, the following holds:

$$[\hat{L}_x, \hat{x}^2] = \hat{x}[\hat{L}_x, \hat{x}] + [\hat{L}_x, \hat{x}]\hat{x} \overset{(5.15)}{=} 0 + 0 = 0$$
$$[\hat{L}_x, \hat{y}^2] = \hat{y}[\hat{L}_x, \hat{y}] + [\hat{L}_x, \hat{y}]\hat{y} \overset{(5.16)}{=} \hat{y}(i\hbar\hat{z}) + (i\hbar\hat{z})\hat{y} = 2i\hbar\hat{y}\hat{z}$$
$$[\hat{L}_x, \hat{z}^2] = \hat{z}[\hat{L}_x, \hat{z}] + [\hat{L}_x, \hat{z}]\hat{z} \overset{(5.17)}{=} \hat{z}(-i\hbar\hat{y}) + (-i\hbar\hat{y})\hat{z} = -2i\hbar\hat{y}\hat{z}$$

Thus, for the sum of the three terms, we derive the desired expression; the scalar operator $\hat{\mathbf{r}}^2$ commutes, as expected, with the angular momentum operator:

$$[\hat{L}_x, \hat{\mathbf{r}}^2] = [\hat{L}_x, \hat{x}^2 + \hat{y}^2 + \hat{z}^2] = [\hat{L}_x, \hat{x}^2] + [\hat{L}_x, \hat{y}^2] + [\hat{L}_x, \hat{z}^2] = 2i\hbar\hat{y}\hat{z} - 2i\hbar\hat{y}\hat{z} = 0 \ \square$$

An important special case of a commutator between a vector and a scalar operator is the commutator between a component of the angular momentum operator and the magnitude operator $\hat{\mathbf{L}}^2$:

$$[\hat{L}_i, \hat{\mathbf{L}}^2] = 0 \tag{5.31}$$

This relation can be verified through a simple calculation using index exchange and renaming:

$$\begin{aligned}
[\hat{L}_i, \hat{\mathbf{L}}^2] = [\hat{L}_i, \hat{L}_j\hat{L}_j] &= \hat{L}_j[\hat{L}_i, \hat{L}_j] + [\hat{L}_i, \hat{L}_j]\hat{L}_j \overset{(5.23)}{=} \\
&= \hat{L}_j\left(i\hbar\varepsilon_{ijk}\hat{L}_k\right) + \left(i\hbar\varepsilon_{ijk}\hat{L}_k\right)\hat{L}_j = \\
&= i\hbar\left(\varepsilon_{ijk}\hat{L}_j\hat{L}_k + \varepsilon_{ijk}\hat{L}_k\hat{L}_j\right) = \ | \ \varepsilon_{ijk} = -\varepsilon_{ikj} \\
&= i\hbar\left(\varepsilon_{ijk}\hat{L}_j\hat{L}_k - \varepsilon_{ikj}\hat{L}_k\hat{L}_j\right) = \ | \ \text{second term: } k \leftrightarrow j \\
&= i\hbar\left(\varepsilon_{ijk}\hat{L}_j\hat{L}_k - \varepsilon_{ijk}\hat{L}_j\hat{L}_k\right) = 0 \ \square
\end{aligned}$$

$\hat{L}_i$ and $\hat{\mathbf{L}}^2$ are thus compatible operators and thereby lead to a common, complete orthonormal system between the angular momentum operator and the magnitude of the angular momentum operator. Therefore, a simultaneous, accurate measurement of the eigenvalues of $\hat{L}_i$ and $\hat{\mathbf{L}}^2$ is possible. We will engage in constructing such an eigen system in the following chapters.

5.1.4 Ladder Operators of the Angular Momentum

Similar to the raising and lowering operators of the harmonic oscillator, *raising and lowering operators* $\hat{L}_\pm$ for angular momentum $\hat{\mathbf{L}}$ can also be constructed. Their specific effect will be discussed in the next chapter, as only there the eigenstates of the angular momentum operator are introduced. Here, we primarily focus on defining $\hat{L}_\pm$ and familiarizing ourselves with useful commutator relations. The raising and lowering operator for angular momentum can be expressed as follows:

$$\hat{L}_\pm = \hat{L}_x \pm i\hat{L}_y \tag{5.32}$$

In this definition, the z-component of the angular momentum is chosen as the preferred direction according to convention. How exactly the sign is interpreted will be discussed in the next chapter. In anticipation, we denote $\hat{L}_+$ as the *raising operator* and $\hat{L}_-$ as the *lowering operator*. Transposition and complex conjugation lead to the switch between the raising and lowering operators:

$$\hat{L}_\pm^\dagger = \left(\hat{L}_x \pm i\hat{L}_y\right)^\dagger = \hat{L}_x^\dagger \mp i\hat{L}_y^\dagger \overset{(5.12)}{=} \hat{L}_x \mp i\hat{L}_y = \hat{L}_\mp \tag{5.33}$$

Relevant Relationships A useful relation is obtained by forming the product between the raising and lowering operators (and vice versa). In the last step, the relationship (5.23) is used:

$$\begin{aligned}\hat{L}_+\hat{L}_- &= \left(\hat{L}_x + i\hat{L}_y\right)\left(\hat{L}_x - i\hat{L}_y\right) = \hat{L}_x^2 + \hat{L}_y^2 + i\left(\hat{L}_y\hat{L}_x - \hat{L}_x\hat{L}_y\right) = \\ &= \hat{L}_x^2 + \hat{L}_y^2 - i[\hat{L}_x, \hat{L}_y] \overset{(5.23)}{=} \hat{L}_x^2 + \hat{L}_y^2 + \hbar\hat{L}_z \qquad (5.34)\\ \hat{L}_-\hat{L}_+ &= \left(\hat{L}_x - i\hat{L}_y\right)\left(\hat{L}_x + i\hat{L}_y\right) = \hat{L}_x^2 + \hat{L}_y^2 - i\left(\hat{L}_y\hat{L}_x - \hat{L}_x\hat{L}_y\right) = \\ &= \hat{L}_x^2 + \hat{L}_y^2 + i[\hat{L}_x, \hat{L}_y] \overset{(5.23)}{=} \hat{L}_x^2 + \hat{L}_y^2 - \hbar\hat{L}_z \qquad (5.35)\end{aligned}$$

From (5.4), it follows that $\hat{L}_x^2 + \hat{L}_y^2 = \hat{\mathbf{L}}^2 - \hat{L}_z^2$. Substituting this into (5.34) and (5.35), we get:

$$\hat{L}_\pm\hat{L}_\mp = \hat{L}_x^2 + \hat{L}_y^2 \pm \hbar\hat{L}_z = \hat{\mathbf{L}}^2 - \hat{L}_z^2 \pm \hbar\hat{L}_z \qquad (5.36)$$

The sum of the expressions from (5.34) and (5.35) (i.e., the anticommutator between the raising and lowering operators) leads us to:

$$\hat{L}_+\hat{L}_- + \hat{L}_-\hat{L}_+ = 2\left(\hat{L}_x^2 + \hat{L}_y^2\right) = 2\left(\hat{\mathbf{L}}^2 - \hat{L}_z^2\right) \qquad (5.37)$$

We can rewrite this expression by again using the relationship $\hat{L}_x^2 + \hat{L}_y^2 = \hat{\mathbf{L}}^2 - \hat{L}_z^2$ derived from (5.4) and substituting this into (5.37):

$$\hat{\mathbf{L}}^2 = \frac{1}{2}\left(\hat{L}_+\hat{L}_- + \hat{L}_-\hat{L}_+\right) + \hat{L}_z^2 \qquad (5.38)$$

In the next chapter, we want to derive the effect of $\hat{\mathbf{L}}^2$, $\hat{L}_z$, and $\hat{L}_\pm$. Both the relationship (5.38) and the following commutator relations are very helpful. We start with the commutator between the raising and lowering operators $\hat{L}_+$ and $\hat{L}_-$, which is easily calculated (in the last step, the relationship (5.23) using $[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z$ is used):

$$\begin{aligned}[\hat{L}_\pm, \hat{L}_\mp] &= [\hat{L}_x \pm i\hat{L}_y, \hat{L}_x \mp i\hat{L}_y] = \\ &= [\hat{L}_x, \hat{L}_x] + [\hat{L}_y, \hat{L}_y] \pm i([\hat{L}_y, \hat{L}_x] - [\hat{L}_x, \hat{L}_y]) = \\ &= \mp 2i[\hat{L}_x, \hat{L}_y] \overset{(5.23)}{=} \pm 2\hbar\hat{L}_z \qquad (5.39)\end{aligned}$$

Another important relationship is the commutator between $\hat{L}_z$ and $\hat{L}_\pm$ (we use (5.23); more precisely, the variants $[\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y$ and $[\hat{L}_z, \hat{L}_y] = -i\hbar\hat{L}_x$):

$$\begin{aligned}[\hat{L}_z, \hat{L}_\pm] &= [\hat{L}_z, \hat{L}_x \pm i\hat{L}_y] = [\hat{L}_z, \hat{L}_x] \pm i[\hat{L}_z, \hat{L}_y] \overset{(5.23)}{=} \\ &= i\hbar\hat{L}_y \pm \hbar\hat{L}_x = \pm\hbar\left(\hat{L}_x \pm i\hat{L}_y\right) \overset{(5.32)}{=} \\ &= \pm\hbar\hat{L}_\pm \qquad (5.40)\end{aligned}$$

Writing out the commutator in (5.40) explicitly, it becomes $\hat{L}_z\hat{L}_\pm - \hat{L}_\pm\hat{L}_z = \pm\hbar\hat{L}_\pm$, which can be rewritten as:

$$\hat{L}_z\hat{L}_\pm = \hat{L}_\pm\hat{L}_z \pm \hbar\hat{L}_\pm \qquad (5.41)$$

Due to (5.31), the relationship between the commutator of $\hat{\mathbf{L}}^2$ and $\hat{L}_\pm$ is trivially zero:

$$[\hat{\mathbf{L}}^2, \hat{L}_\pm] \overset{(5.32)}{=} [\hat{\mathbf{L}}^2, \hat{L}_x \pm i\hat{L}_y] \overset{(5.31)}{=} [\hat{\mathbf{L}}^2, \hat{L}_x] \pm i[\hat{\mathbf{L}}^2, \hat{L}_y] = 0 \qquad (5.42)$$

Now that we are equipped with a *"toolbox"* of useful relationships, we want to investigate the eigenstates of $\hat{\mathbf{L}}^2$ and $\hat{L}_z$ in the next chapter.

5.2 Eigen System of Angular Momentum

Due to (5.23), the components $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ *do not* commute among themselves and thus *cannot* be measured simultaneously. This is a typical quantum phenomenon. While in classical physics, angular momentum can be represented by a vector whose components obviously exist and can be measured independently, this is not the case in quantum physics. As soon as we measure one vector component of angular momentum, we automatically affect the other two components of angular momentum! Because compatible operators can be found to have a common eigen system, at least the simultaneous measurement of the *magnitude* of angular momentum via $\hat{\mathbf{L}}^2$ and a *single* angular momentum component is possible as already shown in (5.31). We conventionally settle on the $\hat{L}_z$ operator, i.e., the z-component of angular momentum. Without initially imposing restrictions, we define $\{|a,b\rangle\}$ as the common, complete eigen system of $\hat{L}_z$ and $\hat{\mathbf{L}}^2$. The eigenvalue equations are:

$$\hat{\mathbf{L}}^2 |a,b\rangle = a |a,b\rangle \tag{5.43}$$

$$\hat{L}_z |a,b\rangle = b |a,b\rangle \tag{5.44}$$

While the eigenvalue a of $\hat{\mathbf{L}}^2$ represents the length of the angular momentum vector, the eigenvalue b of $\hat{L}_z$ represents the z-component of the angular momentum. However, the eigenvalues a and b must satisfy the following relationship:

$$0 \leq |b|^2 \leq a \quad \text{with} \quad a \geq 0 \tag{5.45}$$

This makes sense: A single component of angular momentum cannot be longer than the entire length of the angular momentum (in the classical case: $\mathbf{L}^2 = x^2 + y^2 + z^2$ with $L^2 = a$ and $z^2 = |b|^2$). We focus on the inequality $|b|^2 \leq a$, which will be discussed in more detail later in the context of the uncertainty relation.

5.2.1 Construction of an Angular Momentum Multiplet

We want to construct our angular momentum eigenstate $|a,b\rangle$ as a linear *multiplet*: This means that starting from a fixed quantum number a, we can reach every state of the multiplet $\{|a,b\rangle\}$ through ladder operators. $\hat{L}_\pm$ should therefore only lead to a change in the quantum number b while leaving a unchanged. But how do the ladder operators act on an eigenstate of angular momentum? To find this out, we let $\hat{L}_z$ act on $\left(\hat{L}_\pm |a,b\rangle\right)$:

$$\begin{aligned}\hat{L}_z \left(\hat{L}_\pm |a,b\rangle\right) = \hat{L}_z \hat{L}_\pm |a,b\rangle &\overset{(5.41)}{=} (\hat{L}_\pm \hat{L}_z \pm \hbar \hat{L}_\pm) |a,b\rangle = \hat{L}_\pm (\hat{L}_z \pm \hbar) |a,b\rangle \overset{(5.44)}{=} \\ &= \hat{L}_\pm (b \pm \hbar) |a,b\rangle = (b \pm \hbar) \hat{L}_\pm |a,b\rangle = \\ &= (b \pm \hbar) \left(\hat{L}_\pm |a,b\rangle\right)\end{aligned} \tag{5.46}$$

As required, the eigenvalue $(b \pm \hbar)$ of the $\hat{L}_z$ operator shows that the ladder operators only act on b. We now perform the same calculation for $\hat{\mathbf{L}}^2$, using the fact that according to (5.42), the operators $\hat{\mathbf{L}}^2$ and $\hat{L}_\pm$ commute:

$$\begin{aligned}\hat{\mathbf{L}}^2 \left(\hat{L}_\pm |a,b\rangle\right) = \hat{\mathbf{L}}^2 \hat{L}_\pm |a,b\rangle &\overset{(5.42)}{=} \hat{L}_\pm \hat{\mathbf{L}}^2 |a,b\rangle \overset{(5.43)}{=} \hat{L}_\pm a |a,b\rangle = a \hat{L}_\pm |a,b\rangle = \\ &= a \left(\hat{L}_\pm |a,b\rangle\right)\end{aligned} \tag{5.47}$$

In summary, we have shown: The raising operator increases the eigenvalue of $\hat{L}_z$ by $\hbar$, while the lowering operator decreases it by $\hbar$. However, the raising and lowering operators do not influence the eigenvalue of $\hat{\mathbf{L}}^2$ and thus do not change the magnitude of the angular momentum. We have thus constructed a multiplet state! By repeatedly applying $\hat{L}_\pm$, we can switch between

all states corresponding to b for a fixed a. By associating the eigenvalue $b \pm \hbar$ in (5.47) with the corresponding part of the eigenvector, we can express the action of $\hat{L}_\pm$ as follows (with N as a normalization factor):

$$\hat{L}_\pm |a, b\rangle = N |a, b \pm \hbar\rangle$$

Since the state $|a, b\rangle$ was changed by the action of $\hat{L}_\pm$, the relation above is no longer an eigenvalue equation! We can now determine the norm N, and thus the complete action of the ladder operator. To do this, we rewrite the ladder operators in terms of $\hat{\mathbf{L}}^2$ and $\hat{L}_z$, using (5.36), as we already know their action from (5.43) and (5.44):

$$\begin{aligned}
|N|^2 \langle a, b \pm \hbar | a, b \pm \hbar\rangle &= \langle \hat{L}_\pm(a,b) | \hat{L}_\pm(a,b)\rangle = \\
&= \langle a, b | \hat{L}_\pm^\dagger \hat{L}_\pm | a, b\rangle \overset{(5.33)}{=} \\
&= \langle a, b | \hat{L}_\mp \hat{L}_\pm | a, b\rangle \overset{(5.36)}{=} \\
&= \langle a, b | \hat{\mathbf{L}}^2 - \hat{L}_z^2 \mp \hbar \hat{L}_z | a, b\rangle = \\
&= \langle a, b | \hat{\mathbf{L}}^2 | a, b\rangle - \langle a, b | \hat{L}_z^2 | a, b\rangle \mp \hbar \langle a, b | \hat{L}_z | a, b\rangle \overset{(5.43,5.44)}{=} \\
&= \langle a, b | a | a, b\rangle - \langle a, b | b^2 | a, b\rangle \mp \hbar \langle a, b | b | a, b\rangle = \\
&= \left(a - b^2 \mp \hbar b\right) \langle a, b | a, b\rangle = \\
&= a - b(b \pm \hbar)
\end{aligned} \tag{5.48}$$

We thus obtain the following relation for the action of the ladder operator on a multiplet state $|a, b\rangle$:

$$\hat{L}_\pm |a, b\rangle = \sqrt{a - b(b \pm \hbar)}\, |a, b \pm \hbar\rangle \tag{5.49}$$

Let us now consider the states with the maximum and minimum eigenvalues of $\hat{L}_z$, calling them $b_{\max}$ and $b_{\min}$. Assuming such states exist is justified because we assume a fixed a, and according to (5.45), b is bounded. For example, if we let the raising operator act on the state with maximum angular momentum in the z-direction, we have:

$$\hat{L}_+ |a, b_{\max}\rangle = 0 \tag{5.50}$$

This must be so because the maximum value of $\hat{L}_z$ is already maximal and cannot be increased further. The same applies for the action of the lowering operator on the minimum value of $\hat{L}_z$:

$$\hat{L}_- |a, b_{\min}\rangle = 0 \tag{5.51}$$

We also know that the ladder operator $\hat{L}_+$ (except for the state $|a, b_{\max}\rangle$) increases the eigenvalue b by $\hbar$, and correspondingly the ladder operator $\hat{L}_-$ (except for the state $|a, b_{\min}\rangle$) decreases the eigenvalue b by $\hbar$. Consider a state with an arbitrary eigenvalue b. Since a maximum value $b_{\max}$ exists, there must be a corresponding number $n_{\max}$ times the ladder operator $\hat{L}_+$ needs to be applied to reach b_{max} starting from b. Correspondingly, $n_{\min}$ is the necessary number of applications of $\hat{L}_-$ to reach the minimum eigenvalue b_{min}. We can express this as follows:

$$\begin{aligned}
b_{\max} &= b + n_{\max}\hbar \\
b_{\min} &= b - n_{\min}\hbar
\end{aligned} \tag{5.52}$$

Therefore, the difference between the boundary eigenvalues is as follows:

$$b_{\max} - b_{\min} = \hbar(n_{\max} + n_{\min}) = \hbar n \tag{5.53}$$

Here, n is the maximum number of consecutive identical ladder operations that can be performed. Thus, starting from a boundary eigenvalue of the multiplet state, we can reach *any* other state

of the multiplet with consistent application of a ladder operator. To show the size of n, we apply $\hat{L}_-\hat{L}_+$ to $|a, b_{\max}\rangle$ and consider that the result must be zero due to relation (5.50):

$$\begin{aligned} 0 &= \hat{L}_-\hat{L}_+ |a, b_{\max}\rangle \overset{(5.36)}{=} \\ &= \left(\hat{\mathbf{L}}^2 - \hat{L}_z^2 - \hbar\hat{L}_z\right) |a, b_{\max}\rangle \overset{(5.46,5.47)}{=} \\ &= \left(a - b_{\max}^2 - \hbar b_{\max}\right) |a, b_{\max}\rangle = \\ &= b_{\max}^2 + \hbar b_{\max} - a \end{aligned} \tag{5.54}$$

We can solve a quadratic equation for $b_{\max}$ and obtain two solutions. Here we choose only the positive solution, as we are searching for the maximum value $b_{\max}$:

$$b_{\max} = -\frac{\hbar}{2} \pm \sqrt{\frac{\hbar^2}{4} + a} = -\frac{\hbar}{2} + \sqrt{\frac{\hbar^2}{4} + a}$$

We perform an equivalent calculation for the action of the lowering operator on the minimum state, using (5.51):

$$\begin{aligned} 0 &= \hat{L}_+\hat{L}_- |a, b_{\min}\rangle \overset{(5.36)}{=} \\ &= \left(\hat{\mathbf{L}}^2 - \hat{L}_z^2 + \hbar\hat{L}_z\right) |a, b_{\min}\rangle \overset{(5.46),(5.47)}{=} \\ &= \left(a - b_{\min}^2 + \hbar b_{\min}\right) |a, b_{\min}\rangle = \\ &= b_{\min}^2 - \hbar b_{\min} - a \end{aligned} \tag{5.55}$$

Again, we solve a quadratic equation, but this time we choose the negative solution to obtain the minimum state:

$$b_{\min} = \frac{\hbar}{2} \pm \sqrt{\frac{\hbar^2}{4} + a} = \frac{\hbar}{2} - \sqrt{\frac{\hbar^2}{4} + a}$$

We see that $b_{\max} = -b_{\min}$ must hold. This allows us, using the definition $n = 2j$, to rewrite the relation (5.53) as follows:

$$b_{\max} - b_{\min} = 2b_{\max} = n\hbar = 2j\hbar \implies b_{\max} = j\hbar \tag{5.56}$$

While n must always be an integer, j can, due to our definition, indeed be a half-integer (and thus the eigenvalue b). At the moment, however, we want to restrict j also to *integer* values. In this case, by convention, j is renamed to $j = l$ and this quantum number is called the *(orbital) angular momentum quantum number*. We can now express the eigenvalue a as a function of j (or rather l). First, we rewrite relation (5.54) for a, then insert (5.56). Finally, we rename j to l as explained earlier:

$$a = b_{\max}^2 + \hbar b_{\max} \overset{(5.56)}{=} \hbar^2(j^2 + j) = \hbar^2 j(j+1) = \hbar^2 l(l+1) \tag{5.57}$$

Since we now express the eigenvalue a as a function of l, it is more logical to describe the quantum state not as $|a, b\rangle$ but as $|l, b\rangle$. We also perform another substitution: $b \to m$:

$$b = m\hbar \tag{5.58}$$

Here, m is referred to as the *magnetic quantum number*. Consequently, the quantum state $|a, b\rangle$ is now represented as $|l, m\rangle$. Thus, for the angular momentum magnitude operator $\hat{\mathbf{L}}^2$, we obtain, using (5.43) and (5.57), the following eigenvalue equation:

$$\hat{\mathbf{L}}^2 \,|l,m\rangle = \hbar^2 l(l+1)\,|l,m\rangle \tag{5.59}$$

For the z-component of the angular momentum operator $\hat{L}_z$, we obtain, taking into account (5.44) and (5.58), an equally simple expression:

$$\hat{L}_z \,|l,m\rangle = \hbar m\,|l,m\rangle \tag{5.60}$$

The action of the raising and lowering operators can also be expressed using the quantum numbers l and m. To do this, we substitute (5.57) and (5.58) into (5.49). This yields the following relation:

$$\hat{L}_\pm \,|l,m\rangle = \hbar\sqrt{l(l+1) - m(m\pm 1)}\,|l,m\pm 1\rangle \tag{5.61}$$

Here, $|a, b \pm \hbar\rangle$ was transformed into $|l, m \pm 1\rangle$ because, according to (5.58), $m = b/\hbar$. The range of values for the magnetic quantum number m can be determined using inequality (5.45). By inserting the substitutions (5.57) and (5.58), we obtain:

$$m^2 \leq l(l+1) \tag{5.62}$$

Assuming the extremal values of the magnetic quantum number are $m = \pm l$, then the inequality (5.62) is evidently satisfied, because $l^2 < l(l+1)$. On the other hand, if we try the extremal value $m = \pm(l+1)$, the inequality (5.62) would not hold, because $(l+1)^2 > l(l+1)$. Therefore, we can write the following relationship between the orbital angular momentum quantum number l and the magnetic quantum number m:

$$|m| \leq l \tag{5.63}$$

All m fulfilling (5.63) belong to a multiplet with quantum number l and can be reached via ladder operators. Since ladder operators only affect the eigenvalue m, we cannot leave the multiplet of l using $\hat{L}_\pm$.

We will also focus on integer angular momentum quantum numbers in the following chapter and explore the significance of half-integer angular momentum quantum numbers (still denoted by j) later. How many possible settings of m are there for a fixed quantum number l? From (5.56), we already know that the maximum and minimum values of the magnetic quantum number m are fixed at $m_{\max} = l$ and $m_{\min} = -l$. We sum over all possible m, using the Gaussian summation formula and the substitution $m = n - l$:

$$p(l) = \sum_{m=-l}^{l} 1 = \sum_{n=0}^{2l} 1 = 2l + 1$$

Thus, according to the orbital angular momentum quantum number l, we find a number $p(l)$ of independent magnetic states:

$$p(l) = 2l + 1 \tag{5.64}$$

Figure 23 illustrates a two-dimensional angular momentum system with $l = 3$: While the length of the angular momentum clearly remains constant for each realization of the multiplet, in this case, $p(3) = 2 \cdot 3 + 1 = 7$ settings for m are allowed. Additionally, the action of the raising and lowering operators is schematically depicted.

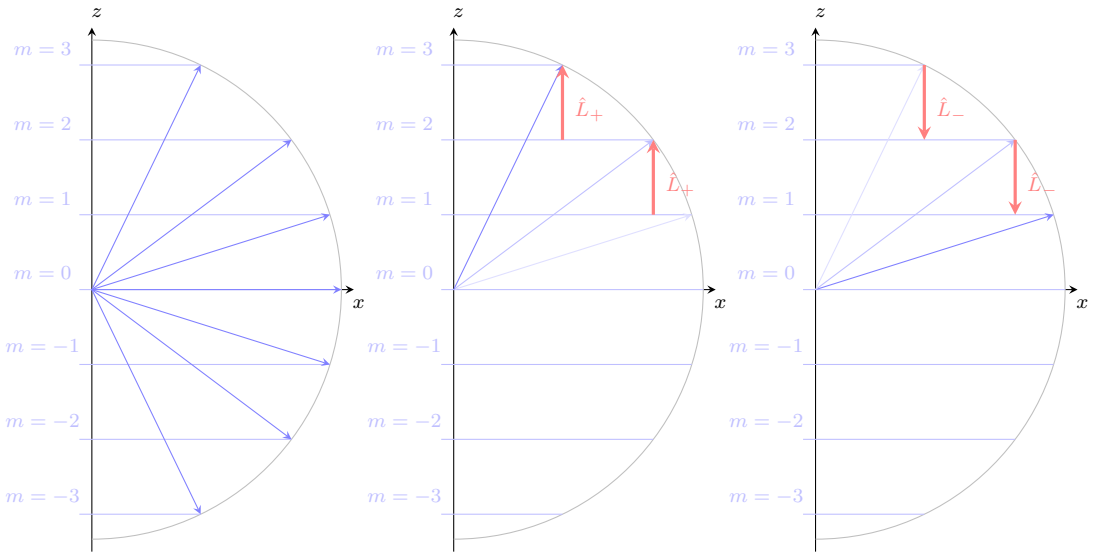


Fig. 23: (left) Schematic, two-dimensional representation of an angular momentum system with $l = 3$. (center) Action of the raising operator $\hat{L}_+$ and (right) the lowering operator $\hat{L}_-$.

5.2.2 Angular Momentum Uncertainty

The commutator relation (5.23) shows that the commutator between two different angular momentum components does not vanish, and the $\hat{L}_x$-, $\hat{L}_y$-, and $\hat{L}_z$-components cannot be measured sharply at the same time. Typically, the measurable quantity by convention (arbitrarily) is the z-component of the angular momentum. This also means that the poles of the angular momentum sphere (see Figure 24) can never be reached by $\hat{\mathbf{L}}$. This must naturally be so because if the angular momentum vector pointed exactly in the (positive or negative) z-direction, both the x- and y-components would be precisely set to zero. However, this is excluded by the commutator relation (5.23)! According to convention, only $\hat{L}_z$ is supposed to be a sharply measurable quantity, so it holds:

$$\langle (\Delta \hat{L}_z)^2 \rangle = \langle \hat{L}_z^2 \rangle - \langle \hat{L}_z \rangle^2 = 0 \tag{5.65}$$

For $\langle (\Delta \hat{L}_x)^2 \rangle$ and $\langle (\Delta \hat{L}_y)^2 \rangle$, this relationship no longer holds. They are smeared in the xy-plane. That means: If one measures the x- or y-component in a system with a fixed z-component, you will (in many repetitions of the experiment) get a different value each time! To calculate $\langle (\Delta \hat{L}_x)^2 \rangle$ and $\langle (\Delta \hat{L}_y)^2 \rangle$ concretely, it helps to express $\hat{L}_x$ and $\hat{L}_y$ through ladder operators. To this end, we consider the two relations from (5.32):

$$\hat{L}_+ = \hat{L}_x + i\hat{L}_y \quad \text{and} \quad \hat{L}_- = \hat{L}_x - i\hat{L}_y \tag{5.66}$$

By adding $\hat{L}_+$ and $\hat{L}_-$ from (5.32), one obtains the relation $\hat{L}_+ + \hat{L}_- = 2\hat{L}_x$, which can be rewritten as:

$$\hat{L}_x = \frac{1}{2}\left(\hat{L}_+ + \hat{L}_-\right) \tag{5.67}$$

If $\hat{L}_-$ is subtracted from $\hat{L}_+$ from (5.32), one obtains $\hat{L}_+ - \hat{L}_- = 2i\hat{L}_y$, which can again be rewritten as:

$$\hat{L}_y = \frac{1}{2i}\left(\hat{L}_+ - \hat{L}_-\right) \tag{5.68}$$

Let us first consider the uncertainty of the x-component of the angular momentum. For the first moment $\langle \hat{L}_x \rangle$, we obtain:

$$\begin{aligned}
\langle \hat{L}_x \rangle &= \langle l,m|\hat{L}_x|l,m\rangle \overset{(5.67)}{=} \\
&= \frac{1}{2}\langle l,m|\hat{L}_+ + \hat{L}_-|l,m\rangle = \\
&= \frac{1}{2}\left(\langle l,m|\hat{L}_+|l,m\rangle + \langle l,m|\hat{L}_-|l,m\rangle\right) \overset{(5.61)}{=} \\
&= \frac{1}{2}\left(N_+\langle l,m|l,m+1\rangle + N_-\langle l,m|l,m-1\rangle\right) = 0
\end{aligned} \tag{5.69}$$

On average, $\hat{L}_x$ fluctuates around zero, so the first moment $\langle \hat{L}_x \rangle = 0$ vanishes. However, because each individual measurement yields a non-zero value, we expect the second moment $\langle \hat{L}_x^2 \rangle$ to be non-zero. This will be checked in the following:

$$\begin{aligned}
\langle \hat{L}_x^2 \rangle &= \langle l,m|\hat{L}_x^2|l,m\rangle \overset{(5.67)}{=} \\
&= \frac{1}{4}\langle l,m|(\hat{L}_+ + \hat{L}_-)^2|l,m\rangle = \\
&= \frac{1}{4}\left(\langle l,m|\hat{L}_+^2|l,m\rangle + \langle l,m|\hat{L}_-^2|l,m\rangle + \langle l,m|\hat{L}_+\hat{L}_- + \hat{L}_-\hat{L}_+|l,m\rangle\right) \overset{(5.37)}{=} \\
&= \frac{1}{4}\left(N_+^2\langle l,m|l,m+2\rangle + N_-^2\langle l,m|l,m-2\rangle + 2\langle l,m|\hat{\mathbf{L}}^2 - \hat{L}_z^2|l,m\rangle\right) = \\
&= \frac{1}{2}\left(\langle l,m|\hat{\mathbf{L}}^2|l,m\rangle - \langle l,m|\hat{L}_z^2|l,m\rangle\right) \overset{(5.59,5.60)}{=} \\
&= \frac{1}{2}\left[\hbar^2 l(l+1)\langle l,m|l,m\rangle - m^2\hbar^2\langle l,m|l,m\rangle\right] = \\
&= \frac{\hbar^2}{2}\left[l(l+1) - m^2\right]
\end{aligned} \tag{5.70}$$

The variance $\langle (\Delta \hat{L}_x)^2 \rangle$ as a measure of the uncertainty of the x angular momentum component can be expressed concretely as:

$$\langle (\Delta \hat{L}_x)^2 \rangle = \langle \hat{L}_x^2 \rangle - \langle \hat{L}_x \rangle^2 \overset{(5.69)(5.70)}{=} \frac{\hbar^2}{2}\left[l(l+1) - m^2\right] \tag{5.71}$$

If we want to know the *minimal* uncertainty, we must insert the *maximum* value for m in (5.71). According to (5.63), this is $m = l$, so we obtain for the minimum uncertainty:

$$\langle (\Delta \hat{L}_x)^2 \rangle \geq \frac{\hbar^2}{2}\left(l(l+1) - l^2\right) = \left(l^2 + 1 - l^2\right) = \frac{\hbar^2}{2} l \tag{5.72}$$

Finally, to obtain the value for the smallest possible uncertainty, we only need to replace l with the smallest possible value of the quantum number l in (5.72). Since l is the quantum number for angular momentum (and can only take integer values), the minimum angular momentum setting is $l = l_{\min} = 1$. Inserting this into (5.72), we find:

$$\langle (\Delta \hat{L}_x)^2 \rangle \geq \frac{\hbar^2}{2} \quad \text{for} \quad l = l_{\min} = 1 \tag{5.73}$$

While $l = 0$ would also be a valid orbital angular momentum quantum number, the angular momentum would vanish and thus would not be meaningful for a discussion of uncertainty. If half-integer angular momentum values are possible, we replace l with j. The minimum value is $j = j_{\min} = \frac{1}{2}$, giving:

$$\langle (\Delta \hat{L}_x)^2 \rangle \geq \frac{\hbar^2}{4} \quad \text{for} \quad j = j_{\min} = \frac{1}{2} \tag{5.74}$$

The same results are obtained following an analogous calculation for the y angular momentum component.

In-Depth: Robertson-Schrödinger Uncertainty Relations

The starting point is the *Robertson Uncertainty Relation*, which can be expressed in the form of the following inequality for the product of two standard deviations of the operators $\hat{A}$ and $\hat{B}$:

$$\Delta A \cdot \Delta B \equiv \sigma_A \sigma_B \geq \frac{1}{2} \left| \langle [\hat{A}, \hat{B}] \rangle \right|$$

It is also supposed to hold that $[\hat{A}, \hat{B}] \neq 0$. To abbreviate the notation of the following scalar products, we introduce the definitions for the standard deviations σ_A and σ_B:

$$\sigma_A^2 = \langle \hat{A} - \langle \hat{A} \rangle \, | \hat{A} - \langle \hat{A} \rangle \rangle = \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2 \implies |\alpha\rangle = |\hat{A} - \langle A \rangle\rangle$$
$$\sigma_B^2 = \langle \hat{B} - \langle \hat{B} \rangle \, | \hat{B} - \langle \hat{B} \rangle \rangle = \langle \hat{B}^2 \rangle - \langle \hat{B} \rangle^2 \implies |\beta\rangle = |\hat{B} - \langle B \rangle\rangle$$

Note that, actually, $|\hat{A} - \langle A \rangle\rangle = |(\hat{A} - \langle A \rangle)\psi\rangle$ holds, since the operators always act on an underlying state. Over the Cauchy-Schwarz inequality from (3.18), we can estimate the product $\sigma_A \sigma_B$ (it holds $||\alpha|| = \sqrt{\langle \alpha | \alpha \rangle}$):

$$\sigma_A^2 \sigma_B^2 = \langle \alpha | \alpha \rangle \, \langle \beta | \beta \rangle \geq | \, \langle \alpha | \beta \rangle \, |^2 \tag{5.75}$$

The scalar product of α and β can be expressed using the appropriate commutator and anticommutator:

$$\langle \alpha | \beta \rangle = \langle \alpha \beta \rangle = \frac{1}{2} \langle \alpha\beta + \alpha\beta \rangle = \frac{1}{2} \langle \alpha\beta - \beta\alpha + \alpha\beta + \beta\alpha \rangle = \frac{1}{2} \langle [\alpha, \beta] + \{\alpha, \beta\} \rangle \tag{5.76}$$

In general, the scalar product corresponds to a complex number. We could already show that the commutator is always anti-hermitian, and therefore, its eigenvalue is purely complex, while the anticommutator is hermitian and has a purely real eigenvalue. Taking the magnitude of $\langle \alpha | \beta \rangle$ follows the same rules as taking the magnitude of a complex number $|z|^2 = |x + iy|^2 = |x|^2 + |y|^2$:

$$| \, \langle \alpha | \beta \rangle \, |^2 \overset{(5.76)}{=} \frac{1}{4} | \, \langle [\alpha, \beta] \rangle + \langle \{\alpha, \beta\} \rangle \, |^2 = \frac{1}{4} | \, \langle [\alpha, \beta] \rangle \, |^2 + \frac{1}{4} | \, \langle \{\alpha, \beta\} \rangle \, |^2 \tag{5.77}$$

We insert (5.77) into (5.75) and take the square root of the entire expression; the result is termed the *Robertson-Schrödinger Uncertainty Relation*:

$$\sigma_A \sigma_B \geq \frac{1}{2} \sqrt{| \, \langle [\alpha, \beta] \rangle \, |^2 + | \, \langle \{\alpha, \beta\} \rangle \, |^2} \geq \frac{1}{2} | \, \langle [\alpha, \beta] \rangle \, | \tag{5.78}$$

Only the last estimation leads us to the desired expression of the Robertson Uncertainty, which is trivial by neglecting the second term. We can verify the relation from (5.78) on the already well-known Heisenberg uncertainty relation $\sigma_x \sigma_p \geq \hbar/2$:

$$\sigma_x \sigma_p \overset{(5.78)}{\geq} \frac{1}{2} | \, \langle [\hat{\mathbf{x}}, \hat{\mathbf{p}}] \rangle \, | \overset{(3.132)}{=} \frac{1}{2} | \, \langle i\hbar \rangle \, | = \frac{\hbar}{2}$$

The relationship from (5.74) can also be easily confirmed with the *Robertson Uncertainty Relation* from (5.78); the following connection holds:

$$\sigma_{L_x} \sigma_{L_y} \geq \frac{1}{2} | \, \langle [\hat{L}_x, \hat{L}_y] \rangle \, | = \frac{\hbar}{2} | \, \langle \hat{L}_z \rangle \, | \overset{(5.60)}{=} \frac{\hbar^2}{2} m \tag{5.79}$$

In the case of an angular momentum setting at maximum m, it also holds that $m = l$. Since the uncertainty product $\sigma_{L_x} \sigma_{L_y}$ is non-zero, there will always be angular momentum components in the x and y directions for every quantum number m (if $l \neq 0$, this is also true for $m = 0$, as seen

from (5.73)). Only in the case $l = 0$ do the x and y components of angular momentum vanish; this does not contradict the uncertainty relation since no angular momentum exists when $l = 0$.

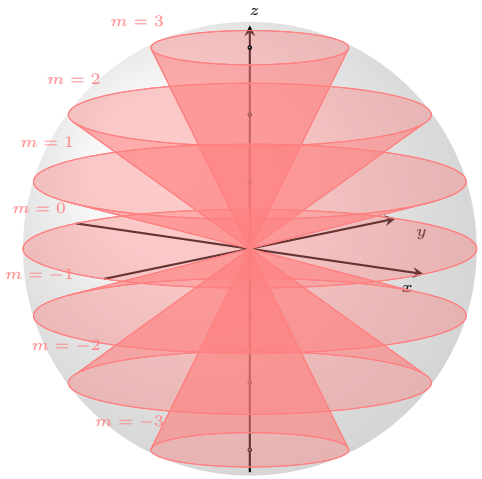


Fig. 24: Schematic, three-dimensional representation of an angular momentum system with $l = 3$: While the z-component is sharply defined, the x and y components are smeared and lie in this sketch on the intersections of the sphere and the corresponding cones.

Let's recall the start of the chapter, specifically (5.45). Here, the inequality $0 \leq |b|^2 \leq a$ or $0 \leq |m|^2 \leq l(l+1)$ becomes clear: If equality $|m|^2 = l(l+1)$ were to hold, the product $\sigma_{L_x}\sigma_{L_y}$ would vanish according to (5.74) and thus allow all angular momentum components to be determined *exactly*. We have, however, already recognized that this is non-physical and cannot correspond to reality. Only without angular momentum $l = 0$ may the equality relation be fulfilled!

5.3 Angular Momentum in Position Space

Until now, we have come to know the eigenstates of angular momentum as angular momentum multiplets as abstract states in the angular momentum eigenspace. However, to verify our predictions in experiments, we must continue our considerations in position space. This is achieved by projection onto $\{|\mathbf{r}\rangle\}$, where we will transition from Cartesian coordinates to a spherically symmetric coordinate system $\{|\mathbf{r}\rangle\} \equiv \{|\vartheta, \varphi\rangle\}$. It holds:

$$Y_l^m(\mathbf{r}) = \langle \mathbf{r}|l, m\rangle \implies Y_l^m(\vartheta, \varphi) = \langle \vartheta, \varphi|l, m\rangle \tag{5.80}$$

Later, we will associate $Y_l^m(\mathbf{r})$ with spherical harmonics, but for now, $Y_l^m(\mathbf{r})$ should just represent the position representation of an angular momentum eigenstate. The relevant eigenvalue equations from (5.59) and (5.60) thus result in:

$$\hat{\mathbf{L}}^2 Y_l^m(\mathbf{r}) = \hbar^2 l(l+1) Y_l^m(\mathbf{r}) \tag{5.81}$$

$$\hat{L}_z Y_l^m(\mathbf{r}) = m\hbar Y_l^m(\mathbf{r}) \tag{5.82}$$

But what do the eigenfunctions $Y_l^m(\mathbf{r})$ look like concretely? In the following chapter, we will perform the transformation of $\hat{\mathbf{L}}$ into spherical coordinates and will find an explicit representation of $|l, m\rangle$ in position space!

5.3.1 Angular Momentum in Spherical Coordinates

As a repetition: The angular momentum in operator notation is given by $\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$. We know from position and momentum operators the *cartesian* position representation, and have found

the following expression in (5.2):

$$\hat{\mathbf{L}} = -i\hbar \begin{pmatrix} y\partial_z - z\partial_y \\ z\partial_x - x\partial_z \\ x\partial_y - y\partial_x \end{pmatrix}$$

However, it is more natural to express the angular momentum in spherically-symmetric coordinates. We must therefore represent both $\{x, y, z\}$ and the partial derivatives $\{\partial_x, \partial_y, \partial_z\}$ in spherical coordinates. For a position vector $\mathbf{r}$, the transformation occurs via:

$$\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r\sin(\vartheta)\cos(\varphi) \\ r\sin(\vartheta)\sin(\varphi) \\ r\cos(\vartheta) \end{pmatrix} \tag{5.83}$$

For the present problem, however, the inversion of (5.83) is relevant, since we want to transform from the Cartesian coordinate system to spherical coordinates. Respectively, it holds:

$$r = \sqrt{x^2 + y^2 + z^2} \tag{5.84}$$

$$\vartheta = \arccos\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) \tag{5.85}$$

$$\varphi = \arctan\left(\frac{y}{x}\right) \tag{5.86}$$

From Figure 25, the relations in (5.84) become easily understandable. While there are no restrictions for the Cartesian coordinate system ($\{x, y, z\} \in \mathbb{R}$), for $\{r, \vartheta, \varphi\}$, we must note that the radius $r \in [0, +\infty)$ must be positive and the polar angle $\vartheta \in [0, \pi]$ is *not* defined over the entire angle range, while this is the case for the azimuthal angle $\varphi \in [0, 2\pi]$.

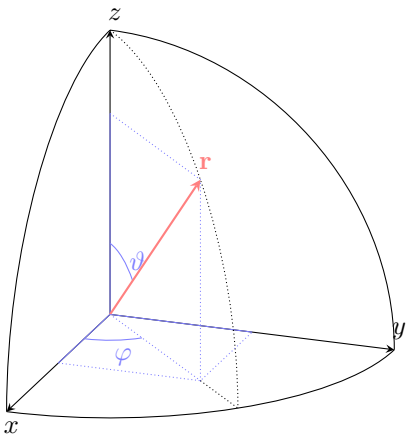


Fig. 25: Coordinates of a spherically symmetric system; here, only the first octant is shown for simplicity. We recognize that every point of the coordinate system can be uniquely identified by $\{r, \vartheta, \varphi\}$.

Let us turn to the partial derivatives. From (5.83), it becomes apparent that in the Cartesian coordinate system $x(r, \vartheta, \varphi)$, $y(r, \vartheta, \varphi)$, and $z(r, \vartheta, \varphi)$ all depend on the radius r, as well as the polar angle ϑ and the azimuthal angle φ. This allows for an extension of the respective differential operators ∂_x, ∂_y, and ∂_z using the chain rule – thus, it follows:

$$\begin{aligned} \partial_x \equiv \frac{\partial}{\partial x} &= \frac{\partial r}{\partial x}\frac{\partial}{\partial r} + \frac{\partial \vartheta}{\partial x}\frac{\partial}{\partial \vartheta} + \frac{\partial \varphi}{\partial x}\frac{\partial}{\partial \varphi} \\ \partial_y \equiv \frac{\partial}{\partial y} &= \frac{\partial r}{\partial y}\frac{\partial}{\partial r} + \frac{\partial \vartheta}{\partial y}\frac{\partial}{\partial \vartheta} + \frac{\partial \varphi}{\partial y}\frac{\partial}{\partial \varphi} \\ \partial_z \equiv \frac{\partial}{\partial z} &= \frac{\partial r}{\partial z}\frac{\partial}{\partial r} + \frac{\partial \vartheta}{\partial z}\frac{\partial}{\partial \vartheta} + \frac{\partial \varphi}{\partial z}\frac{\partial}{\partial \varphi} \end{aligned} \tag{5.87}$$

For the total differential of a function that depends on Cartesian coordinates (for example, $f = r(x,y,z)$, $f = \vartheta(x,y,z)$, or $f = \varphi(x,y,z)$), it can be written:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz \tag{5.88}$$

The actual forms of dr, $d\vartheta$, and $d\varphi$ can be found in (5.92), (5.96), and (5.100).

Example: Derivation of dr, $d\vartheta$, and $d\varphi$

Radius r: This part of the derivation is quite simple: We differentiate (5.84) with respect to x, y, and z, then substitute $\sqrt{x^2+y^2+z^2}$ in the denominator with r and convert the numerator using (5.83) into spherical coordinates. Thus, the following results:

$$\frac{\partial r}{\partial x} = \frac{1}{2}\frac{2x}{\sqrt{x^2+y^2+z^2}} \overset{(5.84)}{=} \frac{x}{r} \overset{(5.83)}{=} \frac{r\sin(\vartheta)\cos(\varphi)}{r} = \sin(\vartheta)\cos(\varphi) \tag{5.89}$$

$$\frac{\partial r}{\partial y} = \frac{1}{2}\frac{2y}{\sqrt{x^2+y^2+z^2}} \overset{(5.84)}{=} \frac{y}{r} \overset{(5.83)}{=} \frac{r\sin(\vartheta)\sin(\varphi)}{r} = \sin(\vartheta)\sin(\varphi) \tag{5.90}$$

$$\frac{\partial r}{\partial z} = \frac{1}{2}\frac{2z}{\sqrt{x^2+y^2+z^2}} \overset{(5.84)}{=} \frac{z}{r} \overset{(5.83)}{=} \frac{r\cos(\vartheta)}{r} = \cos(\vartheta) \tag{5.91}$$

For the total differential dr, we finally insert (5.89–5.91) into (5.88):

$$dr = \frac{\partial r}{\partial x}dx + \frac{\partial r}{\partial y}dy + \frac{\partial r}{\partial z}dz = \sin(\vartheta)\cos(\varphi)dx + \sin(\vartheta)\sin(\varphi)dy + \cos(\vartheta)dz \tag{5.92}$$

Polar Angle $d\vartheta$: This derivation is similar, but involves a little more calculation effort. We start by differentiating the expression (5.85) with respect to x, y, and z:

$$\begin{aligned}
\frac{\partial\vartheta}{\partial x} &= -\frac{1}{\sqrt{1-\frac{z^2}{x^2+y^2+z^2}}}\left(-\frac{1}{2}\right)\frac{z}{(x^2+y^2+z^2)^{\frac{3}{2}}}2x \overset{(5.84)}{=} \frac{1}{\sqrt{1-\frac{z^2}{r^2}}}\frac{xz}{r^3} = \\
&= \frac{1}{\frac{1}{r}\sqrt{r^2-z^2}}\frac{xz}{r^3} = \frac{xz}{r^2\sqrt{r^2-z^2}} \overset{(5.84)}{=} \frac{xz}{r^2\sqrt{x^2+y^2}} \overset{(5.83)}{=} \\
&= \frac{r^2\sin(\vartheta)\cos(\varphi)\cos(\vartheta)}{r^2\sqrt{r^2\sin^2(\vartheta)\cos^2(\varphi)+r^2\sin^2(\vartheta)\sin^2(\varphi)}} = \\
&= \frac{\sin(\vartheta)\cos(\varphi)\cos(\vartheta)}{\sqrt{r^2\sin^2(\vartheta)\left[\cos^2(\varphi)+\sin^2(\varphi)\right]}} = \frac{\cos(\varphi)\cos(\vartheta)}{r}
\end{aligned} \tag{5.93}$$

$$\begin{aligned}
\frac{\partial\vartheta}{\partial y} &= -\frac{1}{\sqrt{1-\frac{z^2}{x^2+y^2+z^2}}}\left(-\frac{1}{2}\right)\frac{z}{(x^2+y^2+z^2)^{\frac{3}{2}}}2y \overset{(5.84)}{=} \frac{1}{\sqrt{1-\frac{z^2}{r^2}}}\frac{yz}{r^3} = \\
&= \frac{1}{\frac{1}{r}\sqrt{r^2-z^2}}\frac{yz}{r^3} = \frac{yz}{r^2\sqrt{r^2-z^2}} \overset{(5.84)}{=} \frac{yz}{r^2\sqrt{x^2+y^2}} \overset{(5.83)}{=} \\
&= \frac{r^2\sin(\vartheta)\sin(\varphi)\cos(\vartheta)}{r^2\sqrt{r^2\sin^2(\vartheta)\cos^2(\varphi)+r^2\sin^2(\vartheta)\sin^2(\varphi)}} = \\
&= \frac{\sin(\vartheta)\sin(\varphi)\cos(\vartheta)}{\sqrt{r^2\sin^2(\vartheta)\left[\cos^2(\varphi)+\sin^2(\varphi)\right]}} = \frac{\sin(\varphi)\cos(\vartheta)}{r}
\end{aligned} \tag{5.94}$$

$$\frac{\partial\vartheta}{\partial z} = -\frac{1}{\sqrt{1-\frac{z^2}{x^2+y^2+z^2}}}\left(\frac{1}{\sqrt{x^2+y^2+z^2}} - \frac{1}{2}\frac{z}{(x^2+y^2+z^2)^{\frac{3}{2}}}2z\right) \overset{(5.84)}{=}$$

$$= -\frac{1}{\sqrt{1-\frac{z^2}{r^2}}}\left(\frac{1}{r}-\frac{z^2}{r^3}\right) = -\frac{1}{\frac{1}{r}\sqrt{r^2-z^2}}\left(\frac{1}{r}-\frac{z^2}{r^3}\right) = -\frac{1-\frac{z^2}{r^2}}{\sqrt{r^2-z^2}} \overset{(5.84)}{=}$$
$$= -\frac{1-\frac{z^2}{r^2}}{\sqrt{x^2+y^2}} \overset{(5.83)}{=} -\frac{1-\cos^2(\vartheta)}{\sqrt{r^2\sin^2(\vartheta)\cos^2(\varphi)+r^2\sin^2(\vartheta)\sin^2(\varphi)}} =$$
$$= -\frac{\sin^2(\vartheta)}{\sqrt{r^2\sin^2(\vartheta)\left[\cos^2(\varphi)+\sin^2(\varphi)\right]}} = -\frac{\sin(\vartheta)}{r} \tag{5.95}$$

Substituting (5.93–5.95) into the total derivative (5.88), we obtain the expression:

$$d\vartheta = \frac{\partial\vartheta}{\partial x}dx + \frac{\partial\vartheta}{\partial y}dy + \frac{\partial\vartheta}{\partial z}dz = \frac{\cos(\varphi)\cos(\vartheta)}{r}dx + \frac{\sin(\varphi)\cos(\vartheta)}{r}dy - \frac{\sin(\vartheta)}{r}dz \tag{5.96}$$

Azimuthal Angle $d\varphi$: This is finally the analogous computation of the total derivative $d\varphi$:

$$\begin{aligned}
\frac{\partial\varphi}{\partial x} &= \frac{1}{1+\frac{y^2}{x^2}}\left(-\frac{y}{x^2}\right) = -\frac{x^2}{x^2+y^2}\frac{y}{x^2} \overset{(5.83)}{=} \\
&= -\frac{r\sin(\vartheta)\sin(\varphi)}{r^2\sin^2(\vartheta)\cos^2(\varphi)+r^2\sin^2(\vartheta)\sin^2(\varphi)} = \\
&= -\frac{r\sin(\vartheta)\sin(\varphi)}{r^2\sin^2(\vartheta)\left[\cos^2(\varphi)+\sin^2(\varphi)\right]} = -\frac{\sin(\varphi)}{r\sin(\vartheta)} \qquad (5.97) \\
\frac{\partial\varphi}{\partial y} &= \frac{1}{1+\frac{y^2}{x^2}}\frac{1}{x} = \frac{1}{1+\frac{y^2}{x^2}}\frac{x}{x^2} = \frac{x}{x^2+y^2} \overset{(5.83)}{=} \\
&= \frac{r\sin(\vartheta)\cos(\varphi)}{r^2\sin^2(\vartheta)\cos^2(\varphi)+r^2\sin^2(\vartheta)\sin^2(\varphi)} = \\
&= \frac{r\sin(\vartheta)\cos(\varphi)}{r^2\sin^2(\vartheta)\left[\cos^2(\varphi)+\sin^2(\varphi)\right]} = \frac{\cos(\varphi)}{r\sin(\vartheta)} \qquad (5.98) \\
\frac{\partial\varphi}{\partial z} &= 0 \qquad (5.99)
\end{aligned}$$

Substituting results (5.97–5.99) again into (5.88), we obtain for $d\varphi$:

$$d\varphi = \frac{\partial\varphi}{\partial x}dx + \frac{\partial\varphi}{\partial y}dy + \frac{\partial\varphi}{\partial z}dz = -\frac{\sin(\varphi)}{r\sin(\vartheta)}dx + \frac{\cos(\varphi)}{r\sin(\vartheta)}dy \tag{5.100}$$

Transformation Matrix: We can summarize all partial derivatives in a transformation matrix T:

$$\mathrm{T} = \begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial\vartheta}{\partial x} & \frac{\partial\varphi}{\partial x} \\ \frac{\partial r}{\partial y} & \frac{\partial\vartheta}{\partial y} & \frac{\partial\varphi}{\partial y} \\ \frac{\partial r}{\partial z} & \frac{\partial\vartheta}{\partial z} & \frac{\partial\varphi}{\partial z} \end{pmatrix} = \begin{pmatrix} \sin(\vartheta)\cos(\varphi) & \frac{\cos(\vartheta)\cos(\varphi)}{r} & -\frac{\sin(\varphi)}{r\sin(\vartheta)} \\ \sin(\vartheta)\sin(\varphi) & \frac{\cos(\vartheta)\sin(\varphi)}{r} & \frac{\cos(\varphi)}{r\sin(\vartheta)} \\ \cos(\vartheta) & -\frac{\sin(\vartheta)}{r} & 0 \end{pmatrix} \tag{5.101}$$

Through matrix-vector multiplication, we can transform the vector $(\partial_x, \partial_y, \partial_z)^T$ into spherically symmetric coordinates in one step – the partial derivatives are accordingly written in the required form in (5.87). Using the inverse of the transformation matrix T, we can also calculate the partial differentials in spherical coordinates based on the Cartesian derivatives. For the matrix-vector multiplication, the following holds in our case in a compact

form:

$$\begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} = T \begin{pmatrix} \partial r \\ \partial_\vartheta \\ \partial_\varphi \end{pmatrix} \tag{5.102}$$

To write the Cartesian, partial derivatives explicitly over the spherical coordinates $\{r, \vartheta, \varphi\}$, we substitute the respective components from the transformation matrix (5.101) into (5.102). The following are three expressions for the Cartesian, partial derivatives, expressed in the desired coordinates:

$$\partial_x \equiv \frac{\partial}{\partial x} = \sin(\vartheta)\cos(\varphi)\frac{\partial}{\partial r} + \frac{\cos(\vartheta)\cos(\varphi)}{r}\frac{\partial}{\partial \vartheta} - \frac{\sin(\varphi)}{r\sin(\vartheta)}\frac{\partial}{\partial \varphi} \tag{5.103}$$

$$\partial_y \equiv \frac{\partial}{\partial y} = \sin(\vartheta)\sin(\varphi)\frac{\partial}{\partial r} + \frac{\cos(\vartheta)\sin(\varphi)}{r}\frac{\partial}{\partial \vartheta} + \frac{\cos(\varphi)}{r\sin(\vartheta)}\frac{\partial}{\partial \varphi} \tag{5.104}$$

$$\partial_z \equiv \frac{\partial}{\partial z} = \cos(\vartheta)\frac{\partial}{\partial r} - \frac{\sin(\vartheta)}{r}\frac{\partial}{\partial \vartheta} \tag{5.105}$$

Example: Derivation of $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$

To represent the individual components of the angular momentum operator $\hat{\mathbf{L}}$ in spherically-symmetric coordinates, we use the previously found relationships of the Cartesian partial derivatives from (5.103), (5.104), and (5.105). Substituting them in, considering the respective transformations in (5.2), we obtain the following relations.

***x*-Component**: According to (5.2), it holds for $\hat{L}_x = -i\hbar(y\partial_z - z\partial_y)$. Substituting the transformed relationships for y, z, ∂_y, and ∂_z, we obtain:

$$\begin{aligned}
\hat{L}_x &= -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) \overset{(5.83)}{=} \\
&= -i\hbar\left(r\sin(\vartheta)\sin(\varphi)\frac{\partial}{\partial z} - r\cos(\vartheta)\frac{\partial}{\partial y}\right) \overset{(5.104)(5.105)}{=} \\
&= -i\hbar\Bigg[r\sin(\vartheta)\sin(\varphi)\left(\cos(\vartheta)\frac{\partial}{\partial r} - \frac{\sin(\vartheta)}{r}\frac{\partial}{\partial \vartheta}\right) - \\
&\quad - r\cos(\vartheta)\left(\sin(\vartheta)\sin(\varphi)\frac{\partial}{\partial r} + \frac{\cos(\vartheta)\sin(\varphi)}{r}\frac{\partial}{\partial \vartheta} + \frac{\cos(\varphi)}{r\sin(\vartheta)}\frac{\partial}{\partial \varphi}\right)\Bigg] = \\
&= -i\hbar\left[-\sin(\varphi)\left(\cos^2(\vartheta) + \sin^2(\vartheta)\right)\frac{\partial}{\partial \vartheta} - \frac{\cos(\vartheta)}{\sin(\vartheta)}\cos(\varphi)\frac{\partial}{\partial \varphi}\right] = \\
&= -i\hbar\left[-\sin(\varphi)\frac{\partial}{\partial \vartheta} - \cot(\vartheta)\cos(\varphi)\frac{\partial}{\partial \varphi}\right]
\end{aligned}$$

***y*-Component**: According to (5.2), it holds for $\hat{L}_y = -i\hbar(z\partial_x - x\partial_z)$. Substituting the transformed relationships for x, z, ∂_x, and ∂_z, we can determine the partial differential

operator in spherically-symmetric coordinates and thus obtain the expression:

$$\begin{aligned}
\hat{L}_y &= -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) \overset{(5.83)}{=} \\
&= -i\hbar\left(r\cos(\vartheta)\frac{\partial}{\partial x} - r\sin(\vartheta)\cos(\varphi)\frac{\partial}{\partial z}\right) \overset{(5.103)(5.105)}{=} \\
&= -i\hbar\Big[r\cos(\vartheta)\left(\sin(\vartheta)\cos(\varphi)\frac{\partial}{\partial r} + \frac{\cos(\vartheta)\cos(\varphi)}{r}\frac{\partial}{\partial \vartheta} - \frac{\sin(\varphi)}{r\sin(\vartheta)}\frac{\partial}{\partial \varphi}\right) - \\
&\quad - r\sin(\vartheta)\cos(\varphi)\left(\cos(\vartheta)\frac{\partial}{\partial r} - \frac{\sin(\vartheta)}{r}\frac{\partial}{\partial \vartheta}\right)\Big] = \\
&= -i\hbar\left[\cos(\varphi)\left(\cos^2(\vartheta) + \sin^2(\vartheta)\right)\frac{\partial}{\partial \vartheta} - \frac{\cos(\vartheta)}{\sin(\vartheta)}\sin(\varphi)\frac{\partial}{\partial \varphi}\right] = \\
&= -i\hbar\left[\cos(\varphi)\frac{\partial}{\partial \vartheta} - \cot(\vartheta)\sin(\varphi)\frac{\partial}{\partial \varphi}\right]
\end{aligned}$$

z-Component: According to (5.2), it holds again for $\hat{L}_z = -i\hbar(x\partial_y - y\partial_x)$. Thus, substituting the transformed relationships for x, y, ∂_x, and ∂_y, it follows:

$$\begin{aligned}
\hat{L}_z &= -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) \overset{(5.83)}{=} \\
&= -i\hbar\left(r\sin(\vartheta)\cos(\varphi)\frac{\partial}{\partial y} - r\sin(\vartheta)\sin(\varphi)\frac{\partial}{\partial x}\right) \overset{(5.103)(5.104)}{=} \\
&= -i\hbar\Big[r\sin(\vartheta)\cos(\varphi)\left(\sin(\vartheta)\sin(\varphi)\frac{\partial}{\partial r} + \frac{\cos(\vartheta)\sin(\varphi)}{r}\frac{\partial}{\partial \vartheta} + \frac{\cos(\varphi)}{r\sin(\vartheta)}\frac{\partial}{\partial \varphi}\right) - \\
&\quad - r\sin(\vartheta)\sin(\varphi)\left(\sin(\vartheta)\cos(\varphi)\frac{\partial}{\partial r} + \frac{\cos(\vartheta)\cos(\varphi)}{r}\frac{\partial}{\partial \vartheta} - \frac{\sin(\varphi)}{r\sin(\vartheta)}\frac{\partial}{\partial \varphi}\right)\Big] = \\
&= -i\hbar\Big[r\sin(\vartheta)\cos(\varphi)\left(\frac{\cos(\varphi)}{r\sin(\vartheta)}\frac{\partial}{\partial \varphi}\right) + r\sin(\vartheta)\sin(\varphi)\left(\frac{\sin(\varphi)}{r\sin(\vartheta)}\frac{\partial}{\partial \varphi}\right)\Big] \\
&= -i\hbar\left(\cos^2(\varphi) + \sin^2(\varphi)\right)\frac{\partial}{\partial \varphi} = -i\hbar\frac{\partial}{\partial \varphi}
\end{aligned}$$

With more or less elaborate calculations, we could express the Cartesian angular momentum components in spherical coordinates (via the angle coordinates ϑ and φ). The radius r actually plays no role! Here are the results found for $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ summarized again. For the x-component of the angular momentum, it holds:

$$\hat{L}_x = i\hbar\left(\sin(\varphi)\frac{\partial}{\partial \vartheta} + \cot(\vartheta)\cos(\varphi)\frac{\partial}{\partial \varphi}\right) \tag{5.106}$$

In its form, the y-component of the angular momentum is very similar to $\hat{L}_x$. It follows:

$$\hat{L}_y = i\hbar\left(-\cos(\varphi)\frac{\partial}{\partial \vartheta} + \cot(\vartheta)\sin(\varphi)\frac{\partial}{\partial \varphi}\right) \tag{5.107}$$

For $\hat{L}_z$, we obtain the simplest differential expression – the z-component of the angular momentum only acts on the azimuthal angle. We can write $\hat{L}_z$ as:

$$\hat{L}_z = -i\hbar\frac{\partial}{\partial \varphi} \tag{5.108}$$

Example: Derivation of $\hat{\mathbf{L}}^2$

In the derivation of $\hat{\mathbf{L}}^2$, we use the previously found relationships from (5.106), (5.107), and (5.108). Compared to the previous calculations, the derivation of the magnitude operator of the angular momentum $\hat{\mathbf{L}}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$ in spherical coordinates is somewhat more laborious, since we must consider that through differential operators, the chain rule must be applied. However, we will recognize that a large portion of the terms obtained cancel out:

$$
\begin{aligned}
\hat{\mathbf{L}}^2 =& L_x^2 + L_y^2 + L_z^2 \overset{(5.106)(5.107)(5.108)}{=} \\
=& -\hbar^2 \Bigg[\left(\sin(\varphi)\frac{\partial}{\partial\vartheta} + \cot(\vartheta)\cos(\varphi)\frac{\partial}{\partial\varphi}\right)\left(\sin(\varphi)\frac{\partial}{\partial\vartheta} + \cot(\vartheta)\cos(\varphi)\frac{\partial}{\partial\varphi}\right) + \\
&+ \left(-\cos(\varphi)\frac{\partial}{\partial\vartheta} + \cot(\vartheta)\sin(\varphi)\frac{\partial}{\partial\varphi}\right)\left(-\cos(\varphi)\frac{\partial}{\partial\vartheta} + \cot(\vartheta)\sin(\varphi)\frac{\partial}{\partial\varphi}\right) + \frac{\partial^2}{\partial\varphi^2}\Bigg] = \\
=& -\hbar^2 \Bigg[\sin^2(\varphi)\frac{\partial^2}{\partial\vartheta^2} + \cot^2(\vartheta)\cos^2(\varphi)\frac{\partial^2}{\partial\varphi^2} + \sin(\varphi)\frac{\partial}{\partial\vartheta}\left(\cot(\vartheta)\cos(\varphi)\frac{\partial}{\partial\varphi}\right) + \\
&+ \cot(\vartheta)\cos(\varphi)\frac{\partial}{\partial\varphi}\left(\sin(\varphi)\frac{\partial}{\partial\vartheta}\right) + \cos^2(\varphi)\frac{\partial^2}{\partial\vartheta^2} + \cot^2(\vartheta)\sin^2(\varphi)\frac{\partial^2}{\partial\varphi^2} - \\
&- \cos(\varphi)\frac{\partial}{\partial\vartheta}\left(\cot(\vartheta)\sin(\varphi)\frac{\partial}{\partial\varphi}\right) - \cot(\vartheta)\sin(\varphi)\frac{\partial}{\partial\varphi}\left(\cos(\varphi)\frac{\partial}{\partial\vartheta}\right) + \frac{\partial^2}{\partial\varphi^2}\Bigg] = \\
=& -\hbar^2 \Bigg[\frac{\partial^2}{\partial\vartheta^2} + \left(1+\cot^2(\vartheta)\right)\frac{\partial^2}{\partial\varphi^2} - \frac{\sin(\varphi)\cos(\varphi)}{\sin^2(\vartheta)}\frac{\partial}{\partial\varphi} + \cot(\vartheta)\sin(\varphi)\cos(\varphi)\frac{\partial^2}{\partial\vartheta\partial\varphi} + \\
&+ \cot(\vartheta)\cos(\varphi)\left(\cos(\varphi)\frac{\partial}{\partial\vartheta} + \sin(\varphi)\frac{\partial^2}{\partial\vartheta\partial\varphi}\right) + \frac{\sin(\varphi)\cos(\varphi)}{\sin^2(\vartheta)}\frac{\partial}{\partial\varphi} - \\
&- \cot(\vartheta)\sin(\varphi)\cos(\varphi)\frac{\partial^2}{\partial\vartheta\partial\varphi} - \cot(\vartheta)\sin(\varphi)\left(-\sin(\varphi)\frac{\partial}{\partial\vartheta} + \cos(\varphi)\frac{\partial^2}{\partial\vartheta\partial\varphi}\right)\Bigg] = \\
=& -\hbar^2 \left[\frac{\partial^2}{\partial\vartheta^2} + \frac{\sin^2(\vartheta)+\cos^2(\vartheta)}{\sin^2(\vartheta)}\frac{\partial^2}{\partial\varphi^2} + \cot(\vartheta)\frac{\partial}{\partial\vartheta}\right] = \\
=& -\hbar^2 \left[\frac{\partial^2}{\partial\vartheta^2} + \cot(\vartheta)\frac{\partial}{\partial\vartheta} + \frac{1}{\sin^2(\vartheta)}\frac{\partial^2}{\partial\varphi^2}\right] = \\
=& -\hbar^2 \left[\frac{1}{\sin(\vartheta)}\frac{\partial}{\partial\vartheta}\left(\sin(\vartheta)\frac{\partial}{\partial\vartheta}\right) + \frac{1}{\sin^2(\vartheta)}\frac{\partial^2}{\partial\varphi^2}\right] = \\
=& -\frac{\hbar^2}{\sin^2(\vartheta)}\left[\sin(\vartheta)\frac{\partial}{\partial\vartheta}\left(\sin(\vartheta)\frac{\partial}{\partial\vartheta}\right) + \frac{\partial^2}{\partial\varphi^2}\right]
\end{aligned}
$$

The magnitude of the angular momentum operator $\hat{\mathbf{L}}^2$ can also be expressed as a function of ϑ and φ. From the above derivation, we obtain the following result:

$$
\hat{\mathbf{L}}^2 = -\frac{\hbar^2}{\sin^2(\vartheta)}\left[\sin(\vartheta)\frac{\partial}{\partial\vartheta}\left(\sin(\vartheta)\frac{\partial}{\partial\vartheta}\right) + \frac{\partial^2}{\partial\varphi^2}\right] \tag{5.109}
$$

Since the individual components $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ were independent of the radius, $\hat{\mathbf{L}}^2$ is naturally independent of r as well. As we will show in more detail later, the last term from (5.109) essentially corresponds to the square of $\hat{L}_z$; this is practical as it simplifies the calculation of the following differential equation. Moreover, the functional form of $\hat{\mathbf{L}}^2$ is very similar to the Laplace operator Δ in spherical coordinates, which we will utilize for deriving the wave functions of the hydrogen atom.

Example: Derivation of $\hat{L}_\pm$

The derivation of $\hat{L}_\pm = \hat{L}_x \pm i\hat{L}_y$ in spherical coordinates is again significantly simpler. We obtain just by inserting (5.106) and (5.107):

$$\begin{aligned}
\hat{L}_\pm &= i\hbar \left[\left(\sin(\varphi)\frac{\partial}{\partial\vartheta} + \cot(\vartheta)\cos(\varphi)\frac{\partial}{\partial\varphi} \right) \pm i \left(-\cos(\varphi)\frac{\partial}{\partial\vartheta} + \cot(\vartheta)\sin(\varphi)\frac{\partial}{\partial\varphi} \right) \right] = \\
&= i\hbar \left[\left(\sin(\varphi) \mp i\cos(\varphi) \right) \frac{\partial}{\partial\vartheta} + \cot(\vartheta) \left(\cos(\varphi) \pm i\sin(\varphi) \right) \frac{\partial}{\partial\varphi} \right] = \\
&= i\hbar \left[\mp i \left(\cos(\varphi) \pm \sin(\varphi) \right) \frac{\partial}{\partial\vartheta} + \cot(\vartheta) \left(\cos(\varphi) \pm i\sin(\varphi) \right) \frac{\partial}{\partial\varphi} \right] = \\
&= \hbar e^{\pm i\varphi} \left[\pm\frac{\partial}{\partial\vartheta} + i\cot(\vartheta)\frac{\partial}{\partial\varphi} \right]
\end{aligned}$$

For the raising and lowering operators $\hat{L}_+$ and $\hat{L}_-$, we therefore find the following expression collectively:

$$\hat{L}_\pm = \hbar e^{\pm i\varphi} \left[\pm\frac{\partial}{\partial\vartheta} + i\cot(\vartheta)\frac{\partial}{\partial\varphi} \right] \tag{5.110}$$

Canonical Commutation Relation For the position and momentum operators $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$, we learned the canonical commutation relation $[\hat{\mathbf{x}}, \hat{\mathbf{p}}] = i\hbar$ in Cartesian coordinates in (3.133). In spherically symmetric spherical coordinates, we can now observe an equivalent relation between the azimuthal angle $\hat{\varphi}$ and the z-component of the angular momentum operator $\hat{L}_z$:

$$[\hat{\varphi}, \hat{L}_z] = i\hbar \tag{5.111}$$

The relation (5.111) can be easily verified by transitioning into the position representation (keeping in mind that the operators act on a wave function, and therefore the chain rule must be considered):

$$[\hat{\varphi}, \hat{L}_z]\psi = -i\hbar \left[\varphi, \frac{\partial}{\partial\varphi} \right] \psi = -i\hbar \left(\varphi\frac{\partial}{\partial\varphi} - \frac{\partial}{\partial\varphi}\varphi \right) \psi = -i\hbar \left(\varphi\frac{\partial}{\partial\varphi}\psi - \frac{\partial}{\partial\varphi}(\varphi\psi) \right) \tag{5.112}$$

$$= -i\hbar \left(\varphi\frac{\partial}{\partial\varphi} - \frac{\partial\varphi}{\partial\varphi} - \varphi\frac{\partial}{\partial\varphi} \right) \psi = i\hbar\psi \quad \square \tag{5.113}$$

The azimuthal angle φ is thus the complementary quantity to the z-component of the angular momentum $\hat{L}_z$.

5.3.2 Legendre Polynomials and Spherical Harmonics

In the previous chapter, we got to know the position representation of the operators $\hat{\mathbf{L}}^2$ and $\hat{L}_z$, where we initially left out a specific representation of the eigenfunctions. First, let's recap our two fundamental eigenvalue equations (5.59) and (5.60) in the required representation space:

$$\begin{aligned}
\hat{\mathbf{L}}^2 Y_l^m(\vartheta,\varphi) &\overset{(5.109)}{=} -\frac{\hbar^2}{\sin^2(\vartheta)} \left[\sin(\vartheta)\frac{\partial}{\partial\vartheta} \left(\sin(\vartheta)\frac{\partial}{\partial\vartheta} \right) + \frac{\partial^2}{\partial^2\varphi} \right] Y_l^m(\vartheta,\varphi) \overset{(5.59)}{=} \\
&= \hbar^2 l(l+1) Y_l^m(\vartheta,\varphi)
\end{aligned} \tag{5.114}$$

$$\hat{L}_z Y_l^m(\vartheta,\varphi) \overset{(5.108)}{=} -i\hbar\frac{\partial}{\partial\varphi} Y_l^m(\vartheta,\varphi) \overset{(5.60)}{=} \hbar m Y_l^m(\vartheta,\varphi) \tag{5.115}$$

We recognize that the eigenvalue equation (5.115) of the z-component of the angular momentum operator depends solely on φ and there is also a clear separation between the ϑ- and φ-dependent

terms in (5.114). It is therefore likely that the eigenvalue equation of $\hat{\mathbf{L}}^2$ will be solvable by a separation approach. We define $Y_l^m = F(\vartheta)\Phi(\varphi)$ and directly substitute this approach into (5.115). $F(\vartheta)$ cancels out, and we obtain a simple differential equation for $\Phi(\varphi)$:

$$\begin{aligned} -i\hbar F(\vartheta)\frac{\partial}{\partial\varphi}\Phi(\varphi) &= \hbar m F(\vartheta)\Phi(\varphi) \mid \cdot \frac{i}{\hbar} \\ \frac{d\Phi(\varphi))}{d\varphi} &= im\Phi(\varphi) \end{aligned} \tag{5.116}$$

This differential equation can be solved quite simply with the following approach:

$$\Phi(\varphi) = C\, e^{+im\varphi} \tag{5.117}$$

The quantum number m in this simple form thus determines the phase of the solution functions. Let's now consider the eigenvalue equation for $\hat{\mathbf{L}}^2$ in (5.114) more closely. The derivative term $\frac{\partial^2}{\partial^2\varphi}$ appears in this equation. But we know from (5.108) that $\hat{L}_z = -i\hbar\partial_\varphi$ contains precisely this derivative. If we square $\hat{L}_z$ in this representation, we find a way to introduce the operator $\hat{L}_z$ into the eigenvalue equation of $\hat{\mathbf{L}}^2$ (5.114):

$$\hat{L}_z^2 = -\hbar^2\frac{\partial^2}{\partial\varphi^2} \implies \frac{\partial^2}{\partial\varphi^2} = -\frac{\hat{L}_z^2}{\hbar^2} \tag{5.118}$$

We can insert the expression obtained in this way into (5.114), and transform the ϑ-dependent differential equation accordingly:

$$\begin{aligned} 0 &= \left\{-\frac{\hbar^2}{\sin^2(\vartheta)}\left[\sin(\vartheta)\frac{\partial}{\partial\vartheta}\left(\sin(\vartheta)\frac{\partial}{\partial\vartheta}\right) + \frac{\partial^2}{\partial^2\varphi}\right] - \hbar^2 l(l+1)\right\} Y_l^m(\vartheta,\varphi) \overset{(5.118)}{=} \\ &= \left\{\frac{1}{\sin^2(\vartheta)}\left[\sin(\vartheta)\frac{\partial}{\partial\vartheta}\left(\sin(\vartheta)\frac{\partial}{\partial\vartheta}\right) - \frac{1}{\hbar^2}\hat{L}_z^2\right] + l(l+1)\right\} F(\vartheta)\Phi(\varphi) \overset{(5.115)}{=} \\ &= \left\{\frac{1}{\sin^2(\vartheta)}\left[\sin(\vartheta)\frac{\partial}{\partial\vartheta}\left(\sin(\vartheta)\frac{\partial}{\partial\vartheta}\right) - \frac{\hbar^2 m^2}{\hbar^2}\right] + l(l+1)\right\} F(\vartheta)\Phi(\varphi) = \\ &= \left\{\frac{1}{\sin^2(\vartheta)}\left[\sin(\vartheta)\frac{\partial}{\partial\vartheta}\left(\sin(\vartheta)\frac{\partial}{\partial\vartheta}\right)\right] + \left[l(l+1) - \frac{m^2}{\sin^2(\vartheta)}\right]\right\} F(\vartheta)\Phi(\varphi) \end{aligned}$$

Admittedly, this expression is somewhat unwieldy. Therefore, we shall introduce a substitution that will significantly simplify writing the differential equation and its solution in the further course:

$$u = \cos(\vartheta) \implies \sin(\vartheta) = \sqrt{1-u^2}$$

Whatever applies for u must also apply for the differentials du. Considering this, we obtain:

$$\frac{du}{d\vartheta} = -\sin(\vartheta) \implies du = -\sin(\vartheta)d\vartheta$$

The derivative with respect to ϑ can also be adapted to the new variable u:

$$\frac{d}{d\vartheta} = \frac{du}{d\vartheta}\frac{d}{du} = -\sin(\vartheta)\frac{d}{du}$$

If we now apply the substitution directly, the representation of the differential equation from (5.114) changes noticeably. It becomes a *Legendre's Differential Equation*:

$$\begin{aligned} 0 &= \left\{\frac{1}{1-u^2}\left[(1-u^2)\frac{\partial}{\partial u}\left((1-u^2)\frac{\partial}{\partial u}\right)\right] + \left[l(l+1) - \frac{m^2}{1-u^2}\right]\right\} f(u) = \\ &= \left\{\frac{\partial}{\partial u}\left[(1-u^2)\frac{\partial}{\partial u}\right] + \left[l(l+1) - \frac{m^2}{1-u^2}\right]\right\} f(u) = \\ &= \left\{(1-u^2)\frac{\partial^2}{\partial u^2} - 2u\frac{\partial}{\partial u} + \left[l(l+1) - \frac{m^2}{1-u^2}\right]\right\} f(u) \end{aligned} \tag{5.119}$$

How (5.119) can indeed be solved will be discussed in more detail in Appendix 10.3. Here we anticipate the solution in the form of the *Rodrigues Formula of Legendre's Differential Equation*:

$$Y_l^m(\vartheta,\varphi) = \frac{(-1)^{l+m}}{2^l l!}\sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}e^{+im\varphi}\sin^m(\vartheta)\frac{d^{m+l}}{d\cos(\vartheta)^{m+l}}\sin^{2l}(\vartheta) \tag{5.120}$$

Apart from the more or less complicated prefactor, we have a φ-dependent part that we call a phase factor and is already known from (5.117). The ϑ-dependent term corresponds to the *associated Legendre polynomials* and depends on both the orbital angular momentum and the magnetic quantum number l and m. If we are interested in spherical harmonics with specific quantum numbers l and m, we either evaluate (5.120) or look it up in tabulated works.

In-Depth: Construction of Spherical Harmonics through Ladder Operators

Similar to the eigenfunctions of the harmonic oscillator, we can also construct the eigenfunctions of $\hat{\mathbf{L}}^2$ and $\hat{L}_z$ by the application of ladder operators. If we carry this out directly in the position space, the resulting functions of the multiplet agree with (5.120).

Let's begin with the position representation of the ladder operators $\hat{L}_\pm$ of angular momentum:

$$\hat{L}_\pm = \hbar e^{\pm i\varphi}\left[\pm\frac{\partial}{\partial\vartheta} + i\cot(\vartheta)\frac{\partial}{\partial\varphi}\right]$$

If $\hat{L}_\pm$ acts on a state at the edge of the multiplet, the expression vanishes since, for example, a minimal state cannot be further decreased. If $\hat{L}_-$ acts on a spherical harmonic function with $m = -l$, we obtain the trivial expression $\hat{L}_- Y_l^{-l} = 0$. From (5.110) and (5.120) it follows:

$$0 = \hbar e^{-i\varphi}\left[-\frac{\partial}{\partial\vartheta} + i\cot(\vartheta)\frac{\partial}{\partial\varphi}\right]F(\vartheta)e^{-il\varphi} = \hbar e^{-i(l+1)\varphi}\left[-\frac{\partial F(\vartheta)}{\partial\vartheta} + l\cot(\vartheta)F(\vartheta)\right]$$

The expression in the bracket must vanish for the equation to be satisfied. To solve the resulting differential equation, we substitute with $u = \sin(\vartheta)$:

$$\ln(F(\vartheta)) + C = l\int d\vartheta\ \cot(\vartheta) = l\int du\ \frac{1}{\cos(\vartheta)}\frac{\cos(\vartheta)}{u} = \ln(u^l)$$

This way, we have found – without considering normalization – the minimal state of the angular momentum multiplet l! It corresponds to:

$$F(\vartheta) = N\ \sin^l(\vartheta) \tag{5.121}$$

To construct arbitrary spherical harmonic functions Y_l^m from (5.121), we need to know the position representation of a p-fold application of a ladder operator $\hat{L}_\pm$. Simple application gives:

$$\begin{aligned}\hat{L}_\pm Y_l^n &= \hbar e^{\pm i\varphi}\left[\pm\frac{\partial}{\partial\vartheta} + i\cot(\vartheta)\frac{\partial}{\partial\varphi}\right]F(\vartheta)e^{+in\varphi} = \\ &= \hbar e^{i(n\pm1)\varphi}\left[\pm\frac{d}{d\vartheta} - n\cot(\vartheta)\right]F(\vartheta) = \\ &= \hbar e^{i(n\pm1)\varphi}\frac{1}{\sin(\vartheta)}\left[\pm\sin(\vartheta)\frac{d}{d\vartheta} - n\cos(\vartheta)\right]F(\vartheta) = \end{aligned}$$

$$= \pm\hbar e^{i(n\pm1)\varphi}\frac{\sin^{1\pm n}(\vartheta)}{\sin(\vartheta)}\frac{d}{d\vartheta}\left[\sin^{\mp n}(\vartheta)F(\vartheta)\right] =$$
$$= \mp\hbar e^{i(n\pm1)\varphi}\sin^{1\pm n}(\vartheta)\frac{d}{d\cos(\vartheta)}\left[\sin^{\mp n}(\vartheta)F(\vartheta)\right]$$

With the right extensions, one can exploit the fact that we can transform the additive expression in the brackets into a product term using the chain rule. Applying $\hat{L}_\pm$ a second time allows us to ignore the additive representation of $\hat{L}_\pm$ and start directly with the simplified expression in product form. It is easily noticeable that primarily the degree of differentiation within the function changes, as well as the power of the sinus and exponential functions not affected by differentiation. We obtain the following expression:

$$\hat{L}_\pm^2 Y_l^n = \hat{L}_\pm\left\{\mp\hbar e^{i(n\pm1)\varphi}\sin^{1\pm n}(\vartheta)\frac{d}{d\cos(\vartheta)}\left[\sin^{\mp n}(\vartheta)F(\vartheta)\right]\right\} =$$
$$= (\mp\hbar)^2 e^{i(n\pm2)\varphi}\sin^{2\pm n}(\vartheta)\frac{d}{d\cos(\vartheta)}\left\{\sin^{\mp n+1}(\vartheta)\sin^{1\pm n}(\vartheta)\frac{d}{d\cos(\vartheta)}\left[\sin^{\mp n}(\vartheta)F(\vartheta)\right]\right\} =$$
$$= (\mp\hbar)^2 e^{i(n\pm2)\varphi}\sin^{2\pm n}(\vartheta)\frac{d^2}{d\cos(\vartheta)^2}\left[\sin^{\mp n}(\vartheta)F(\vartheta)\right]$$

Iteratively, we can generalize this result for a p-fold application of $\hat{L}_\pm^p$:

$$\hat{L}_\pm^p Y_l^n = (\mp\hbar)^p e^{i(n\pm p)\varphi}\sin^{p\pm n}(\vartheta)\frac{d^p}{d\cos(\vartheta)^p}\left[\sin^{\mp n}(\vartheta)F(\vartheta)\right] \tag{5.122}$$

Let $n = -l$ and $p = l + m$, then we obtain – again without normalization – the spherical harmonic Y_l^m:

$$Y_l^m \propto \left(\frac{\hat{L}_+}{\hbar}\right)^{l+m} Y_l^{-l} = (-1)^{l+m}e^{im\varphi}\sin^m(\vartheta)\frac{d^{l+m}}{d\cos(\vartheta)^{l+m}}\left[\sin^l(\vartheta)F(\vartheta)\right] \overset{(5.121)}{=}$$
$$= (-1)^{l+m}e^{im\varphi}\sin^m(\vartheta)\frac{d^{l+m}}{d\cos(\vartheta)^{l+m}}\sin^{2l}(\vartheta)$$

5.3.3 Symmetry Properties of the Eigenfunctions

We now let the parity operator $\hat{\Pi}$ act on the eigenstate of the relevant angular momentum operators $\hat{\mathbf{L}}^2$ and $\hat{L}_z$ to investigate the corresponding symmetry properties. In general, for a state in position space, it holds:

$$\hat{\Pi}\,|\mathbf{r}\rangle = -\,|\mathbf{r}\rangle$$

In spherical coordinates, we must consider that the radius remains unchanged under a reflection ($r \to r$), as it cannot become negative. However, the angles ϑ and φ will change. The following relationships hold:

$$\vartheta \to \pi - \vartheta \quad \text{and} \quad \varphi \to \pi + \varphi \tag{5.123}$$

Why (5.123) is valid can be easily seen in Figures 26 and 27. But what effect does $\hat{\Pi}$ have on the explicit position representation of the eigenfunction? Let's consider the Rodrigues Formula of the Legendre Differential Equation – our chosen representation of the spherical harmonics Y_l^m in (5.120) – and substitute the relations from (5.123).
It follows for the polar angle ϑ:

$$\vartheta \to \pi - \vartheta : \begin{cases}\cos(\vartheta) \to -\cos(\vartheta)\\ \sin(\vartheta) \to \sin(\vartheta)\end{cases}$$

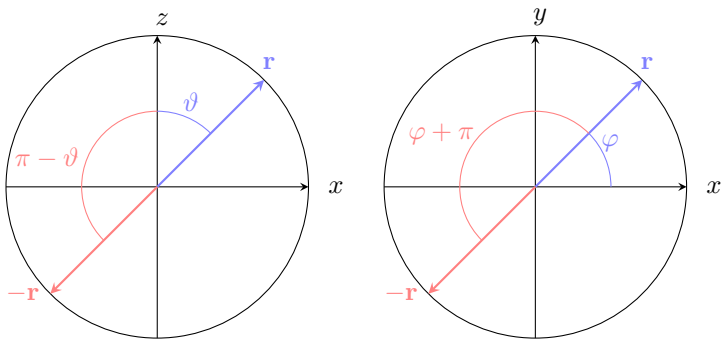


Fig. 26: Parity transformation in spherical coordinates: Only the two-dimensional projection of the angles ϑ and φ is considered.

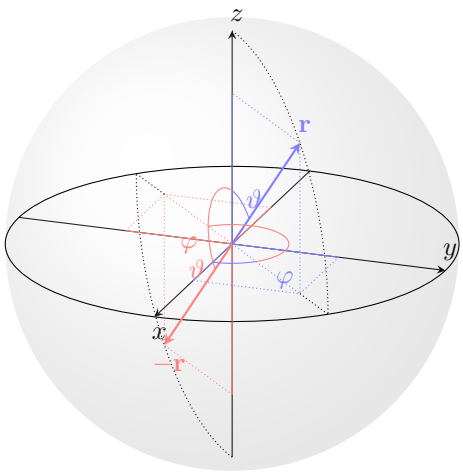


Fig. 27: A three-dimensional representation of the parity transformation of a position vector: While r remains unchanged, only the angles ϑ and φ change.

The azimuthal angle only appears in the phase factor of the spherical harmonics and behaves under the symmetry transformation in the following manner:

$$\varphi \to \pi + \varphi \implies e^{+im\varphi} \to e^{+im\varphi} e^{+im\pi} = (-1)^m e^{+im\varphi}$$

Using these three relations in (5.120), we obtain (noting that only the relevant terms are explicitly written out):

$$\begin{aligned} Y_l^m(\vartheta,\varphi) &= N e^{+im\varphi} \sin^m(\vartheta) \frac{d^{m+l}}{d\cos(\vartheta)^{m+l}} \sin^{2l}(\vartheta) \to \\ &\to N(-1)^m e^{+im\varphi} \sin^m(\vartheta)(-1)^{m+l} \frac{d^{m+l}}{d\cos(\vartheta)^{m+l}} \sin^{2l}(\vartheta) = \\ &= (-1)^l N e^{+im\varphi} \sin^m(\vartheta) \frac{d^{m+l}}{d\cos(\vartheta)^{m+l}} \sin^{2l}(\vartheta) \end{aligned}$$

In short form, we can write the effect of the parity operator on the spherical harmonics $Y_l^m(\vartheta,\varphi)$ as:

$$\hat{\Pi} Y_l^m(\vartheta,\varphi) = (-1)^l Y_l^m(\vartheta,\varphi) \tag{5.124}$$

A parity transformation, i.e., a reflection about the origin of the used spherically symmetric coordinate system, thus depends only on the eigenvalue l and is not influenced by the magnetic quantum number m.

5.3.4 Representations of the Spherical Harmonics

In the following, some spherical harmonics are depicted in two-dimensional polar representation and in three dimensions. In both cases, $Y_0^0(\vartheta, \varphi)$ was omitted, since it can simply be represented by a circle or a sphere.

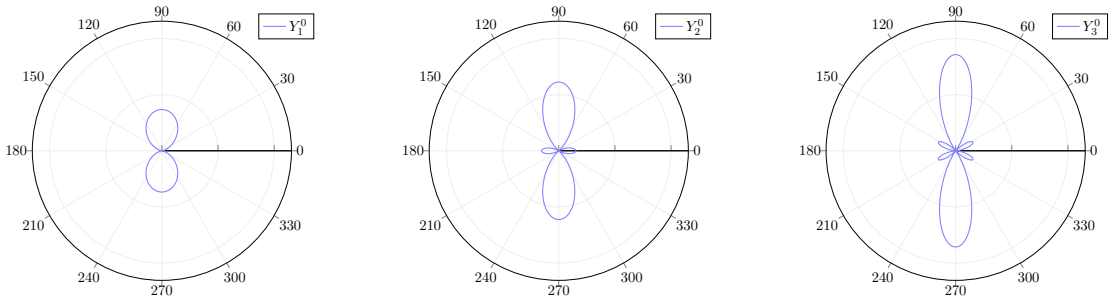


Fig. 28: Polar representations of $Y_l^0(\vartheta, \varphi)$ for $l = 1, 2, 3$.

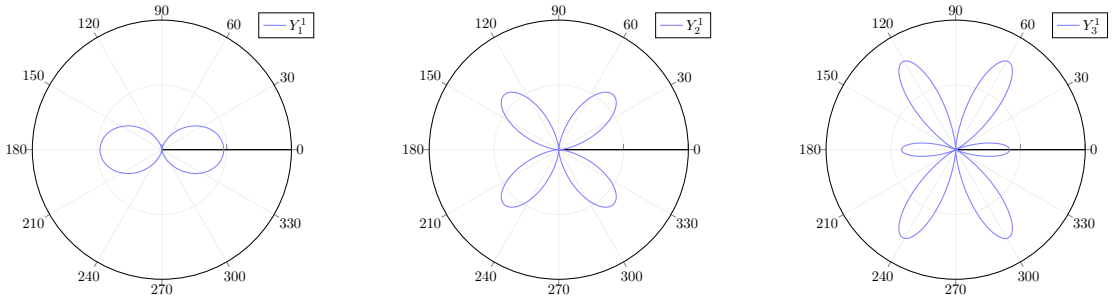


Fig. 29: Polar representations of $Y_l^1(\vartheta, \varphi)$ for $l = 1, 2, 3$.

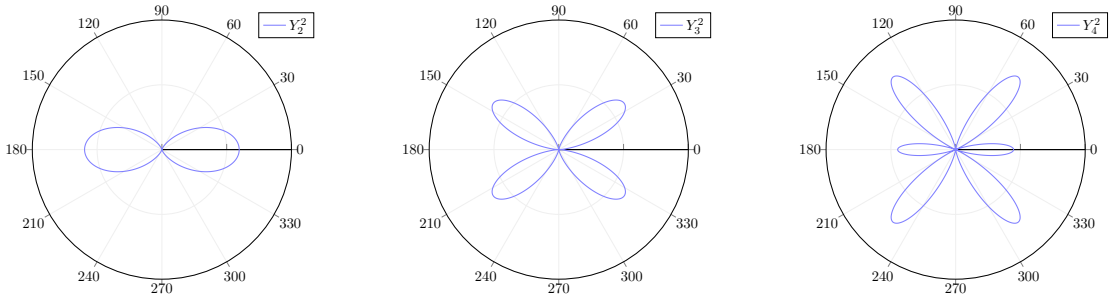


Fig. 30: Polar representations of $Y_l^2(\vartheta, \varphi)$ for $l = 2, 3, 4$.

For example, we can see from this representation that with greater difference $\Delta_{lm} = l - m$, the number of nodes increases. In three dimensions, the same spherical harmonics are shown in Figure 31-33.

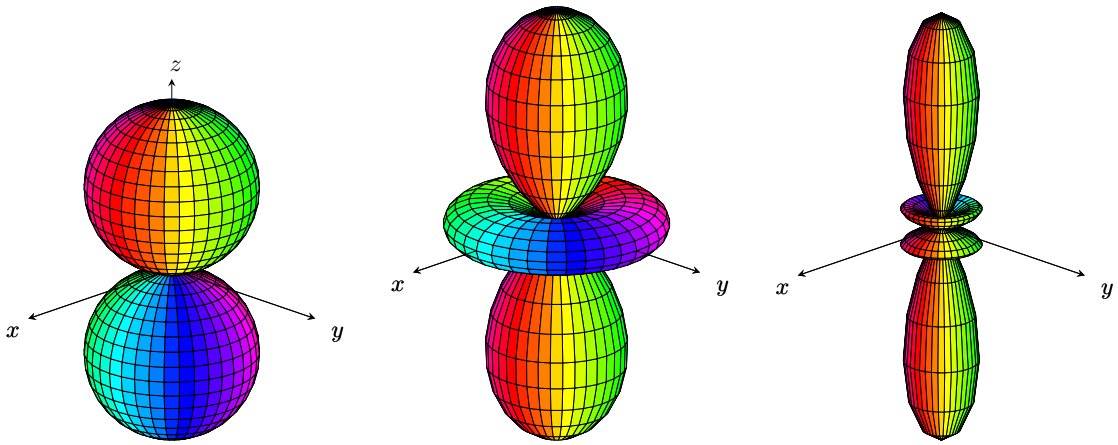


Fig. 31: The spherical harmonics $Y_l^0(\vartheta, \varphi)$ for $l = 1, 2, 3$.

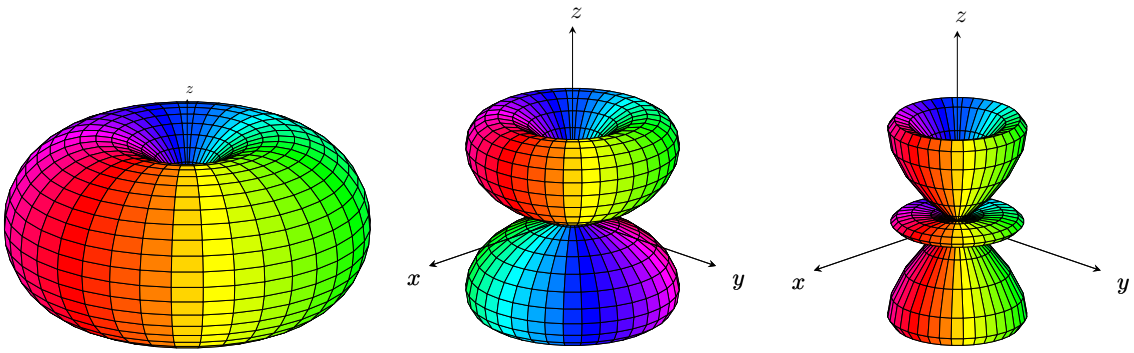


Fig. 32: The spherical harmonics $Y_l^1(\vartheta, \varphi)$ for $l = 1, 2, 3$.

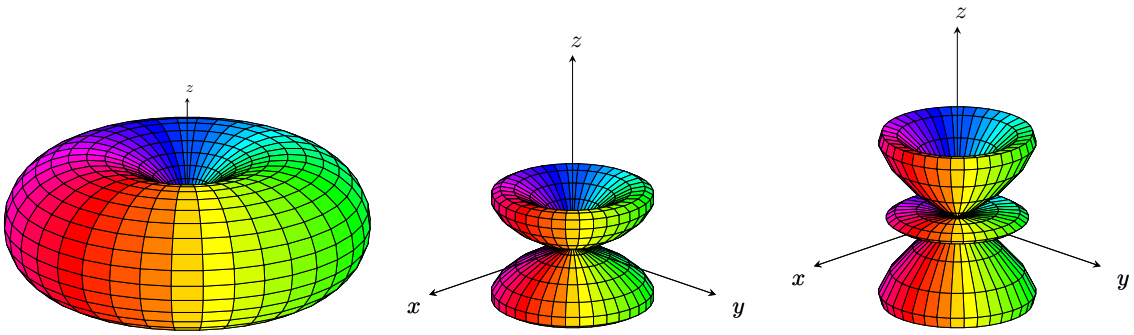


Fig. 33: The spherical harmonics $Y_l^2(\vartheta, \varphi)$ for $l = 2, 3, 4$.

6 Hydrogen Atom

Motivation: The Hydrogen Atom

The good agreement between experimentally measured and quantum mechanically calculated spectral lines in the absorption and emission spectrum of the hydrogen atom was one of the greatest successes of quantum theory since 1925. Besides explaining the series in different wavelength ranges, the solution of the Schrödinger equation for the hydrogen atom also provides the geometric shape of the electron orbitals. This enabled the development of the shell model of electrons, and the understanding of the structure of the periodic table of elements began without further postulates. Additionally, it became possible to explain chemical bonds, which play a central role in the formation of molecules and solids.

The hydrogen atom also played a central role in the extension of quantum theory to relativistic quantum mechanics by Paul Dirac and quantum electrodynamics by Hans Bethe and Edwin Salpeter.

6.1 Schrödinger Equation as a Two-Body Problem

The hydrogen atom consists of a positively charged proton and a negatively charged electron. The proton forms the nucleus of the atom (the "nucleus"). To describe the entire system "hydrogen atom" sensibly, we must consider both the electron and the nucleus in our Hamiltonian operator. Thus, we are dealing with a 6-dimensional two-body problem and a complication of the Schrödinger equation (2.10):

$$\hat{H} = \frac{\hat{\mathbf{p}}_e^2}{2m_e} + \frac{\hat{\mathbf{p}}_N^2}{2m_N} + V(|\mathbf{r}_e - \mathbf{r}_N|) \tag{6.1}$$

While the kinetic terms in $\hat{H}$ can be easily separated, the potential term, which explicitly depends on $\mathbf{r}_e$ and $\mathbf{r}_N$, poses a problem. We will subsequently transform our coordinate system so that we achieve a decoupling of the 6-dimensional problem into two separable three-dimensional problems.

6.1.1 Transformation to the Center of Mass System

We are initially in the laboratory system. Electron and nucleus can be described (at least in the thinking of classical physics) with the position and momentum coordinates $\mathbf{r}_e$ and $\mathbf{r}_N$ as well as $\mathbf{p}_e$ and $\mathbf{p}_N$. The goal is to transition into the center of mass system with the relative coordinates $\mathbf{r}$ and $\mathbf{p}$ and the center of mass coordinates $\mathbf{R}$ and $\mathbf{P}$. This will have the advantage that we can express the relative distance $\mathbf{r}_e - \mathbf{r}_N$ of the electron and proton in our new coordinates very simply.

With the mass of the electron $m_e = 9.109\,383 \times 10^{-31}\,\mathrm{kg}$ and the mass of the nucleus (in other words, the proton mass) $m_N = 1.672\,621 \times 10^{-27}\,\mathrm{kg}$ we can trivially express the total mass as:

$$M = m_N + m_e \approx m_N \tag{6.2}$$

Since the nucleus is approximately 1000 times heavier than the electron, M is practically dominated only by the proton. For the center of mass system, the reduced mass μ is relevant, which is a product of the force balance between the electron and the proton:

$$\mu = \frac{m_e m_N}{m_N + m_e} = \left(\frac{1}{m_e} + \frac{1}{m_N}\right)^{-1} \approx m_e \tag{6.3}$$

For the center of mass coordinate $\mathbf{R}$, the following expression holds, whereby $\mathbf{R}$ lies very close to the nucleus due to the much greater mass m_N:

$$\mathbf{R} = \frac{m_e \mathbf{r}_e}{m_N + m_e} + \frac{m_N \mathbf{r}_N}{m_N + m_e} \approx \frac{m_N}{M} \mathbf{r}_N \tag{6.4}$$

We already know the expression for the relative coordinate $\mathbf{r}$, it corresponds to the difference of both components:

$$\mathbf{r} = \mathbf{r}_e - \mathbf{r}_N \tag{6.5}$$

Thus, after the transformation, the potential term reduces to $V(|\mathbf{r}_e - \mathbf{r}_N|) \to V(|\mathbf{r}|)$ and we achieve a separation of variables. By deriving the position coordinate with respect to time, we also obtain the velocity and thus the momentum. For the total momentum, the following holds:

$$\mathbf{P} = M \frac{d\mathbf{R}}{dt} = M \left(\frac{m_e \mathbf{v}_e}{m_N + m_e} + \frac{m_N \mathbf{v}_N}{m_N + m_e} \right) = m_e \mathbf{v}_e + m_N \mathbf{v}_N = \mathbf{p}_e + \mathbf{p}_N \tag{6.6}$$

The relative motion is determined by the relative mass μ and thus results in:

$$\mathbf{p} = \mu \frac{d\mathbf{r}}{dt} = \mu \left(\mathbf{v}_e - \mathbf{v}_N \right) = \frac{m_N \mathbf{p}_e}{m_N + m_e} - \frac{m_e \mathbf{p}_N}{m_N + m_e} \tag{6.7}$$

In the center of mass system, the relation for the magnitude of the individual momentum components holds: $|\mathbf{p}_e| = |\mathbf{p}_N|$. Through the correspondence principle, we can also associate the relative momentum $\mathbf{p}$ and the center of mass momentum $\mathbf{P}$ with the corresponding operators $\hat{\mathbf{p}}$ and $\hat{\mathbf{P}}$!

Example: Transformation to the Center of Mass System

With (6.4) and (6.5) as well as (6.6) and (6.7), we have found all relevant expressions for the transformation. While we can already write the potential term $V(|\mathbf{r}_e - \mathbf{r}_N|) \to V(|\mathbf{r}|)$ in the new coordinates, we need to rewrite the momentum operators for the two kinetic terms. We will first do this explicitly in the position representation, where $\hat{p}_i = -i\hbar \boldsymbol{\nabla}_i$, and then in an abstract operator notation.

Let's start by expressing the Laplace operators Δ_e and Δ_N in center of mass coordinates. We use an expression similar to (5.87), applying the chain rule:

$$\frac{\partial}{\partial r_{e,i}} = \frac{\partial r_i}{\partial r_{e,i}} \frac{\partial}{\partial r_i} + \frac{\partial R_i}{\partial r_{e,i}} \frac{\partial}{\partial R_i} \overset{(6.5,6.4)}{=} \frac{m_e}{M} \frac{\partial}{\partial R_i} + \frac{\partial}{\partial r_i}$$
$$\frac{\partial}{\partial r_{N,i}} = \frac{\partial r_i}{\partial r_{N,i}} \frac{\partial}{\partial r_i} + \frac{\partial R_i}{\partial r_{N,i}} \frac{\partial}{\partial R_i} \overset{(6.5,6.4)}{=} \frac{m_N}{M} \frac{\partial}{\partial R_i} - \frac{\partial}{\partial r_i}$$

By squaring the upper two expressions, we obtain the individual components of the Laplace operators $\Delta_e = \partial^2_{x_e} + \partial^2_{y_e} + \partial^2_{z_e}$ and $\Delta_N = \partial^2_{x_N} + \partial^2_{y_N} + \partial^2_{z_N}$:

$$\frac{\partial^2}{\partial r^2_{e,i}} = \frac{\partial^2}{\partial r^2_i} + \frac{m_e^2}{M^2} \frac{\partial^2}{\partial R^2_i} + \frac{2m_e}{M} \frac{\partial^2}{\partial r_i \partial R_i}$$
$$\frac{\partial^2}{\partial r^2_{N,i}} = \frac{\partial^2}{\partial r^2_i} + \frac{m_e^2}{M^2} \frac{\partial^2}{\partial R^2_i} - \frac{2m_N}{M} \frac{\partial^2}{\partial r_i \partial R_i}$$

By Schwarz's theorem, we know that the order of derivatives can be exchanged, and thus the upper two expressions simplify.
If we replace the indices i with $\{x, y, z\}$ and move to three dimensions according to Einstein's summation convention, we obtain expressions dependent on the corresponding

Laplace operators Δ_r and Δ_R:

$$\begin{aligned}\Delta_e &= \Delta_r + \frac{m_e^2}{M^2}\Delta_R + \frac{2m_e}{M}\boldsymbol{\nabla}_r\boldsymbol{\nabla}_R \\ \Delta_N &= \Delta_r + \frac{m_N^2}{M^2}\Delta_R - \frac{2m_N}{M}\boldsymbol{\nabla}_r\boldsymbol{\nabla}_R\end{aligned} \tag{6.8}$$

If we substitute these into our Hamiltonian operator $\hat{H}$, which is now independent of the relative potential, we obtain the following simple expression in the position representation:

$$\begin{aligned}\hat{H}(V=0) &= -\frac{\hbar^2}{2m_e}\Delta_e - \frac{\hbar^2}{2m_N}\Delta_N \overset{(6.8)}{=} \\ &= -\frac{\hbar^2}{2}\left(\frac{1}{m_e}\Delta_r + \frac{m_e}{M^2}\Delta_R + \frac{2}{M}\boldsymbol{\nabla}_r\boldsymbol{\nabla}_R + \frac{1}{m_N}\Delta_r + \frac{m_N}{M^2}\Delta_R - \frac{2}{M}\boldsymbol{\nabla}_r\boldsymbol{\nabla}_R\right) = \\ &= -\frac{\hbar^2}{2}\left[\left(\frac{1}{m_e}+\frac{1}{m_N}\right)\Delta_r + \frac{m_e+m_N}{M^2}\Delta_R\right] \overset{(6.2,6.3)}{=} \\ &= -\frac{\hbar^2}{2\mu}\Delta_r - \frac{\hbar^2}{2M}\Delta_R\end{aligned} \tag{6.9}$$

This corresponds in operator notation to the following relation with $\hat{\mathbf{p}}$ and $\hat{\mathbf{P}}$:

$$\hat{H}(V=0) = \frac{\hat{\mathbf{p}}^2}{2\mu} + \frac{\hat{\mathbf{P}}^2}{2M} \tag{6.10}$$

With (6.7) and (6.6), we can derive (6.10) directly through the respective operators. We first rewrite $\hat{\mathbf{p}}_e$ and $\hat{\mathbf{p}}_N$ so that they depend on $\hat{\mathbf{p}}$ and $\hat{\mathbf{P}}$:

$$\begin{aligned}(m_e+m_N)\,\hat{\mathbf{p}}_e &= m_e\hat{\mathbf{P}} + M\hat{\mathbf{p}} \implies \hat{\mathbf{p}}_e = \frac{m_e}{M}\hat{\mathbf{P}} + \hat{\mathbf{p}} \\ (m_e+m_N)\,\hat{\mathbf{p}}_N &= m_N\hat{\mathbf{P}} - M\hat{\mathbf{p}} \implies \hat{\mathbf{p}}_N = \frac{m_N}{M}\hat{\mathbf{P}} - \hat{\mathbf{p}}\end{aligned}$$

Squaring the above expressions, we obtain the kinetic terms of the Hamiltonian (6.1). The center of mass momentum operator $\hat{\mathbf{P}}$ and the relative momentum operator $\hat{\mathbf{p}}$ also commute since they act in distinct spaces ($\{|\mathbf{R}\rangle\}$ and $\{|\mathbf{r}\rangle\}$).

$$\begin{aligned}\hat{\mathbf{p}}_e^2 &= \hat{\mathbf{p}}^2 + \frac{m_e^2}{M^2}\hat{\mathbf{P}}^2 + \frac{2m_e}{M}\hat{\mathbf{P}}\hat{\mathbf{p}} \\ \hat{\mathbf{p}}_N^2 &= \hat{\mathbf{p}}^2 + \frac{m_N^2}{M^2}\hat{\mathbf{P}}^2 - \frac{2m_N}{M}\hat{\mathbf{P}}\hat{\mathbf{p}}\end{aligned} \tag{6.11}$$

(6.11) corresponds to (6.8), and the calculation of $\hat{H}$ is analogous to (6.9).

We obtain a new Hamiltonian operator, with individual terms dependent on a single (center of mass) coordinate:

$$\hat{H} = \frac{\hat{\mathbf{P}}^2}{2M} + \frac{\hat{\mathbf{p}}^2}{2\mu} + V(|\mathbf{r}|) \tag{6.12}$$

Without initially going into the form of the potential $V(|\mathbf{r}|)$, we can provide the stationary Schrödinger equation as:

$$\left(\frac{\hat{\mathbf{P}}^2}{2M} + \frac{\hat{\mathbf{p}}^2}{2\mu} + V(|\mathbf{r}|)\right)\psi(\mathbf{r},\mathbf{R}) = E\psi(\mathbf{r},\mathbf{R}) \tag{6.13}$$

E represents the eigenenergies of the entire system. (6.13) is still a differential equation in 6 dimensions. At first glance, we seemingly have not improved the problem in its complexity

compared to the laboratory system. However, upon closer inspection: It is now possible to separate the variables!

Separation of Variables The potential term $V(|\mathbf{r}|)$ and one of the two kinetic terms now depend only on the relative coordinates, while the second kinetic term is determined by the center of mass component $\hat{\mathbf{R}}$. Through a separation Ansatz, we can separate relative and center of mass coordinates: $\psi(\mathbf{r},\mathbf{R}) = \phi(\mathbf{r})\Psi(\mathbf{R})$. If we substitute this into (6.13), we can write:

$$\begin{aligned} E\phi(\mathbf{r})\Psi(\mathbf{R}) &= \left(\frac{\hat{\mathbf{P}}^2}{2M} + \frac{\hat{\mathbf{p}}^2}{2\mu} + V(|\mathbf{r}|)\right)\phi(\mathbf{r})\Psi(\mathbf{R}) = \\ &= \frac{1}{2M}\hat{\mathbf{P}}^2\left[\phi(\mathbf{r})\Psi(\mathbf{R})\right] + \frac{1}{2\mu}\hat{\mathbf{p}}^2\left[\phi(\mathbf{r})\Psi(\mathbf{R})\right] + V(|\mathbf{r}|)\phi(\mathbf{r})\Psi(\mathbf{R}) \end{aligned}$$

However, since $\hat{\mathbf{P}}^2$ only acts on $\Psi(\mathbf{R})$, and $\hat{\mathbf{p}}^2$ only acts on $\Psi(\mathbf{r})$, in the first term $\phi(\mathbf{r})$ and in the second term $\Psi(\mathbf{R})$ can be pulled before the operator:

$$\begin{aligned} &\frac{\phi(\mathbf{r})}{2M}\hat{\mathbf{P}}^2\Psi(\mathbf{R}) + \frac{\Psi(\mathbf{R})}{2\mu}\hat{\mathbf{p}}^2\phi(\mathbf{r}) + V(|\mathbf{r}|)\phi(\mathbf{r})\Psi(\mathbf{R}) = E\phi(\mathbf{r})\Psi(\mathbf{R}) \implies \\ &\frac{\Psi(\mathbf{R})}{2\mu}\hat{\mathbf{p}}^2\phi(\mathbf{r}) + V(|\mathbf{r}|)\phi(\mathbf{r})\Psi(\mathbf{R}) = E\phi(\mathbf{r})\Psi(\mathbf{R}) - \frac{\phi(\mathbf{r})}{2M}\hat{\mathbf{P}}^2\Psi(\mathbf{R}) \mid \div \Psi(\mathbf{R})\phi(\mathbf{r}) \\ &\frac{1}{\phi(\mathbf{r})}\left(\frac{\hat{\mathbf{p}}^2}{2\mu}\phi(\mathbf{r}) + V(|\mathbf{r}|)\phi(\mathbf{r})\right) = \frac{1}{\Psi(\mathbf{R})}\left(E\Psi(\mathbf{R}) - \frac{\hat{\mathbf{P}}^2}{2M}\Psi(\mathbf{R})\right) \implies \\ &\frac{1}{\phi(\mathbf{r})}\left(\frac{\hat{\mathbf{p}}^2}{2\mu} + V(|\mathbf{r}|)\right)\phi(\mathbf{r}) = \frac{1}{\Psi(\mathbf{R})}\left(E - \frac{\hat{\mathbf{P}}^2}{2M}\right)\Psi(\mathbf{R}) \end{aligned} \tag{6.14}$$

Both sides of the differential equation (6.14) now depend on only one variable each and thus must each result in a constant ε, so that the equation can be satisfied. We first consider the right side of differential equation (6.14), and replace (due to $\hat{\mathbf{P}} = -i\hbar\nabla_R$) the operator $\hat{\mathbf{P}}^2$ by $-\hbar^2\Delta_R$:

$$\begin{aligned} \frac{1}{\Psi(\mathbf{R})}\left(E + \frac{\hbar^2}{2M}\Delta_R\right)\Psi(\mathbf{R}) &= \varepsilon \mid \cdot \Psi(\mathbf{R}) \\ \left(E + \frac{\hbar^2}{2M}\Delta_R\right)\Psi(\mathbf{R}) &= \varepsilon\Psi(\mathbf{R}) \end{aligned} \tag{6.15}$$

The relation (6.15) describes the motion of the center of mass of the atom. Since the equation does not depend on the potential, it can be easily solved by a plane wave:

$$\Psi(\mathbf{R}) = e^{i\mathbf{k}\mathbf{R}} \tag{6.16}$$

We substitute this solution into (6.15) and consider that $\Delta_R\Psi(\mathbf{R}) = \Delta_R e^{i\mathbf{k}\mathbf{R}} = \mathbf{k}^2 e^{i\mathbf{k}\mathbf{R}} = \mathbf{k}^2\Psi(\mathbf{R})$. We then obtain the following expression for the total energy E:

$$\left(E - \frac{\hbar^2\mathbf{k}^2}{2M}\right)\Psi(\mathbf{R}) = \varepsilon\Psi(\mathbf{R}) \implies E = \varepsilon + \frac{\hbar^2\mathbf{k}^2}{2M}$$

The total energy E thus consists (as expected) of the kinetic energy of the center of mass $\hbar^2\mathbf{k}^2/2M$ and ε. The latter corresponds to the relative kinetic energy between the electron and the nucleus – i.e., the interaction energy. We have thus reduced the 6-dimensional problem to a three-dimensional one by separating the center of mass coordinates.

6.1.2 Separation Ansatz in Spherical Coordinates

We now turn to the actually interesting wave function $\phi(\mathbf{r})$ and the associated eigenenergy ε. For this, we now consider the left part of equation (6.14), which must also correspond to the constant energy ε due to the separation Ansatz:

$$\begin{aligned} \frac{1}{\phi(\mathbf{r})}\left(\frac{\hat{\mathbf{p}}^2}{2\mu}+V(|\mathbf{r}|)\right)\phi(\mathbf{r}) &= \varepsilon \mid \cdot\ \phi(\mathbf{r}) \\ \left(\frac{\hat{\mathbf{p}}^2}{2\mu}+V(|\mathbf{r}|)\right)\phi(\mathbf{r}) &= \varepsilon\phi(\mathbf{r}) \end{aligned} \tag{6.17}$$

We now solve the Schrödinger equation from (6.17) in the position space of the relative coordinates, and thus replace (due to $\hat{\mathbf{p}} = -i\hbar\nabla_r$) the operator $\hat{\mathbf{p}}^2$ by $-\hbar^2\Delta_r$:

$$\left(\frac{-\hbar^2}{2\mu}\Delta_r + V(|\mathbf{r}|)\right)\phi(\mathbf{r}) = \varepsilon\phi(\mathbf{r}) \tag{6.18}$$

The potential depends only on the magnitude of the coordinate $|\mathbf{r}| \equiv r$, which is why a transformation into spherical coordinates is suitable. In this case, the angular dependent parts are independent of the potential and can be solved separately. For the transformation, however, the Laplace operator Δ_r must be transformed into spherical coordinates.

The expression for Δ_r in spherical coordinates results in a more or less complex expression, which we can simplify using our knowledge about the spatial representation of angular momentum:

$$\begin{aligned} \Delta_r &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{1}{r^2\sin^2(\vartheta)}\left[\sin(\vartheta)\frac{\partial}{\partial\vartheta}\left(\sin(\vartheta)\frac{\partial}{\partial\vartheta}\right)+\frac{\partial^2}{\partial\varphi^2}\right] \overset{(5.109)}{=} \\ &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{1}{\hbar^2 r^2}\hat{\mathbf{L}}^2 \end{aligned} \tag{6.19}$$

The first term of the operator acts only on the radial component r, while the second term, as shown in (5.109), involves only the angular components ϑ and φ through $\hat{\mathbf{L}}^2$. Incidentally, this separation can also be motivated through the square of the classical angular momentum $\mathbf{L}$:

$$\mathbf{L}^2 = (\mathbf{r}\times\mathbf{p})^2 = (rp\sin(\alpha))^2 = r^2p^2(1-\cos^2(\alpha)) = r^2p^2 - (rp\cos(\alpha))^2 = r^2p^2 - (\mathbf{r}\cdot\mathbf{p})^2$$

This can be transformed into $r^2p^2 = (\mathbf{r}\cdot\mathbf{p})^2 + \mathbf{L}^2$ and finally divided by the radius r:

$$p^2 = \frac{(\mathbf{r}\cdot\mathbf{p})^2}{r^2} + \frac{1}{r^2}\mathbf{L}^2 \tag{6.20}$$

Thus, we have obtained an expression corresponding to the kinetic term $\hat{\mathbf{p}}^2/2\mu$ in (6.18). While the first term of the momentum square projects the momentum vector $\mathbf{p}$ in the direction of $\mathbf{r}$ – thus considering only radial components – the second term corresponds to the angular components due to $\mathbf{L}^2 = \mathbf{L}^2(\vartheta,\varphi)$. Without specifying the potential in more detail, we can rewrite the Schrödinger equation in spherical coordinates by inserting (6.19) into (6.18) as follows:

$$\left[\frac{-\hbar^2}{2\mu r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\hat{\mathbf{L}}^2}{2\mu r^2} + V(r)\right]\phi(\mathbf{r}) = \varepsilon\phi(\mathbf{r}) \tag{6.21}$$

Again, we can perform a separation ansatz to separate the radial wave functions from the angle-dependent part:

$$\phi(\mathbf{r}) = R(r)Y_l^m(\vartheta,\varphi) \tag{6.22}$$

The angular component is completely described by the magnitude operator of angular momentum $\hat{\mathbf{L}}^2$; we already know its eigenfunction in the spatial representation, namely the spherical

harmonics $Y_l^m(\vartheta, \varphi)$. The effect of $\hat{\mathbf{L}}^2$ on the $Y_l^m(\vartheta, \varphi)$ is also known from (5.114) and we can thus write in a simplifying manner:

$$\left[\frac{-\hbar^2}{2\mu r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\hbar^2 l(l+1)}{2\mu r^2} + V(r)\right] R(r)Y_l^m(\vartheta,\varphi) = \varepsilon R(r)Y_l^m(\vartheta,\varphi) \tag{6.23}$$

The spherical harmonics can be canceled out on both sides, and we thereby obtain the differential equation for the radial part of the spherical Schrödinger equation. Only here does it become necessary to consider the specific form of the potential. Due to the electrostatic attractive forces between the electron and proton, $V(r)$ is the Coulomb potential (note the prefactors, depending on the units used a $1/4\pi\varepsilon_0$ might also appear here):

$$V(r) = -\frac{Ze^2}{r} \tag{6.24}$$

Z is the charge number of the nucleus in units of the elementary charge $e = 1.602\,176 \times 10^{-19}\,\mathrm{C}$. In the case of the hydrogen atom, $Z = 1$. We maintain a general Z to be able to describe hydrogen-like atoms (that is, atoms and ions with only one electron).

6.1.3 Effective Potential and Form of the Wave Function

Inserting the Coulomb potential (6.24) into (6.23) gives us the following equation for the radial part $R(r)$ of the wave function:

$$\left[\frac{-\hbar^2}{2\mu r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\hbar^2 l(l+1)}{2\mu r^2} - \frac{Ze^2}{r}\right] R(r) = \varepsilon R(r) \tag{6.25}$$

In (6.25) we call the sum of the angular momentum term and the Coulomb potential the *effective potential*; it corresponds to the following form:

$$V_{\text{eff}}(r) = \underbrace{\frac{\hbar^2 l(l+1)}{2\mu r^2}}_{\substack{\text{Centrifugal-}\\\text{potential}}} - \underbrace{\frac{Ze^2}{r}}_{\substack{\text{Coulomb-}\\\text{potential}}} \tag{6.26}$$

Both terms lie approximately in the same order of magnitude. At small distances, the repulsive centrifugal potential associated with the angular momentum dominates, at large distances the attractive Coulomb potential dominates. This leads in cases where $l > 0$ to electrons hardly being found near the nucleus, since the energy induced by angular momentum increases more strongly at very small distances than the electrostatic attraction. This fact is clearly visible in Figure 34: The larger the orbital angular momentum quantum number l becomes, the greater the influence of angular momentum and the farther the electron will be found from the nucleus.

Before we solve the Schrödinger equation for the central potential, we can make assertions about the properties of the wave function based on the form of the potential. In the case of $l = 0$, the potential term $V(r)$ is consistently negative and thus allows for bound states; if $\varepsilon < 0$, a bound state occurs while an electron with $\varepsilon > 0$ is scattered by the purely electrostatic potential.
Assuming $l > 0$, it is classically forbidden for an electron wave function to penetrate into the positively diverging region of the effective potential – however, due to the tunneling properties of wave functions, we will also find electrons with $l > 0$ at small distances r. Generally, in this chapter, we will only deal with electrons with negative eigenenergy $\varepsilon < 0$, which exhibit a bound state for all l.
Additionally, $V(r)$ is rotationally symmetric. This means that the central potential always has the same value regardless of the angles ϑ and φ. Thus, $[\hat{\Pi}, V(r)] = 0$ and consequently $[\hat{\Pi}, \hat{H}] = 0$.

Parity and Hamiltonian operators are therefore compatible and share the same eigenfunctions. The wave function $\phi(\mathbf{r})$ will thus remain invariant (except for a sign) under point reflections. For the spherical harmonics $Y_l^m(\vartheta, \varphi)$, the behavior under that symmetry transformation has already been shown in (5.123): Applying $\hat{\Pi}$ results in an additional $(-1)^l$ term.

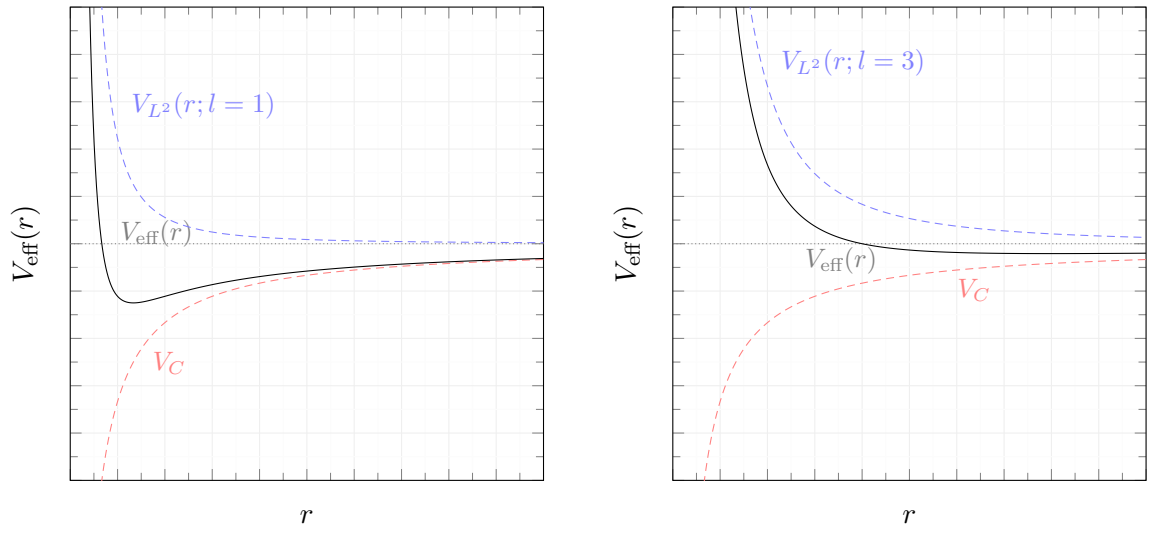


Fig. 34: Effective potential for the cases $l = 1$ and $l = 3$: Dashed lines indicate the pure contributions of the Coulomb potential V_C and the angular momentum term V_{L^2}.

For bound states, the node rule applies again: We will find a quantum number corresponding to the number of "nodes" of the wave function, thus allowing the placement of $\phi(\mathbf{r})$ in an energetic hierarchy.

6.2 Solution for Radial Wave Functions

Let us now discuss how the differential equation for the radial wave function $R(r)$ can be solved. To do this, we first bring the right-hand side of equation (6.25) to the left. Considering that we assume $\varepsilon < 0$ to allow for bound states (we use $\varepsilon = -|\varepsilon|$), we can write this as:

$$\left[-\frac{\hbar^2}{2\mu r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\hbar^2 l(l+1)}{2\mu r^2} - \frac{Ze^2}{r} + |\varepsilon|\right] R(r) = 0 \tag{6.27}$$

The first derivative term with respect to the radius r can, following the product rule, be rewritten as:

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) = \frac{1}{r^2}\left(2r\frac{\partial}{\partial r} + r^2\frac{\partial^2}{\partial r^2}\right) = \left(\frac{2}{r}\frac{\partial}{\partial r} + \frac{\partial^2}{\partial r^2}\right)$$

Inserted into (6.27), we can further simplify the radial wave equation using transformations and substitutions from (6.29) and (6.31):

$$\begin{aligned} 0 &= \left[-\frac{\hbar^2}{2\mu}\left(\frac{2}{r}\frac{\partial}{\partial r} + \frac{\partial^2}{\partial^2 r}\right) + \frac{\hbar^2 l(l+1)}{2\mu r^2} - \frac{Ze^2}{r} + |\varepsilon|\right] R(r) = \Big| \cdot \frac{-2\mu}{\hbar^2} \\ &= \left(\frac{2}{r}\frac{\partial}{\partial r} + \frac{\partial^2}{\partial^2 r} - \frac{l(l+1)}{r^2} + \frac{2Z}{r}\frac{\mu e^2}{\hbar^2} - \frac{2\mu|\varepsilon|}{\hbar^2}\right) R(r) \overset{(6.29,6.31)}{=} \\ &= \left(\frac{2}{r}\frac{\partial}{\partial r} + \frac{\partial^2}{\partial^2 r} - \frac{l(l+1)}{r^2} + \frac{2Z}{a_0 r} - k^2\right) R(r) \end{aligned} \tag{6.28}$$

with

$$a_0 = \frac{\hbar^2}{m_e e^2} \tag{6.29}$$

where we have replaced the reduced mass μ with the electron mass m_e. The difference, however, is only about 0.1 ‰. It turns out that a_0 is the *Bohr radius*, expressed in natural (Planck) units, where all fundamental physical constants are set to 1. The Bohr radius represents the typical radius of the hydrogen atom in its ground state. In SI units, the Bohr radius is given by:

$$a_0 = 4\pi\varepsilon_0 \frac{\hbar^2}{m_e e^2} \tag{6.30}$$

The energy ε and the wave number k are related as follows:

$$|\varepsilon| = \frac{\hbar^2 k^2}{2\mu} \iff k^2 = \frac{2\mu|\varepsilon|}{\hbar^2} \tag{6.31}$$

First Modification of the Radial Wave Function Without knowing the explicit form of the radial wave functions, we know that it must be normalized over the range $r \in [0, \infty)$. When calculating the normalization integral, we must remember that we are working in spherical coordinates and must therefore include the Jacobian determinant when integrating. It follows that:

$$1 = \int_0^\infty dr\, r^2 |R(r)|^2 = \int_0^\infty dr\, |rR(r)|^2 = \int_0^\infty dr\, |u(r)|^2 \quad \text{with} \quad u(r) \overset{\text{def}}{=} rR(r) \tag{6.32}$$

By including the Jacobian determinant and defining a new function $u(r) = rR(r)$, we can interpret the radial wave function $R(r)$ as a sort of spherical wave:

$$R(r) = \frac{u(r)}{r} \tag{6.33}$$

As a side calculation, we now need to explicitly evaluate the first and second derivatives of $R(r) \equiv u/r$:

$$\frac{\partial}{\partial r} R(r) = \frac{\partial}{\partial r}\frac{u}{r} = \frac{u'r - u}{r^2} = \frac{u'}{r} - \frac{u}{r^2} \tag{6.34}$$

$$\begin{aligned} \frac{\partial^2}{\partial r^2} R(r) &= \frac{\partial^2}{\partial r^2}\frac{u}{r} = \frac{\partial}{\partial r}\left(\frac{u'r - u}{r^2}\right) = \frac{(u'r - u)'\, r^2 - (u'r - u)2r}{r^4} \\ &= \frac{(u''r + u' - u')r^2 - 2r^2u' - 2ru}{r^4} = \frac{u''}{r} - 2\frac{u'}{r^2} + 2\frac{u}{r^3} \end{aligned} \tag{6.35}$$

By substituting (6.33), (6.34), and (6.35) into the radial wave equation (6.28), we obtain:

$$\begin{aligned} 0 &= \left[\frac{2}{r}\left(\frac{u'}{r} - \frac{u}{r^2}\right) + \frac{u''}{r} - 2\frac{u'}{r^2} - 2\frac{u}{r^3}\right] + \left(-\frac{l(l+1)}{r^2} + \frac{2Z}{a_0 r} - k^2\right)\frac{u}{r} = \\ &= \frac{1}{r}u'' + \frac{1}{r}\left(-\frac{l(l+1)}{r^2} + \frac{2Z}{a_0 r} - k^2\right)u = \frac{1}{r}\left(\frac{\partial^2}{\partial r^2} - \frac{l(l+1)}{r^2} + \frac{2Z}{a_0 r} - k^2\right)u = \quad | \cdot r \\ &= \left(\frac{\partial^2}{\partial r^2} - \frac{l(l+1)}{r^2} + \frac{2Z}{a_0 r} - k^2\right)u(r) \end{aligned} \tag{6.36}$$

This equation has the form of a single-particle Schrödinger equation with the effective potential as the potential term. However, in this equation, the derivative is not taken with respect to Cartesian coordinates but with respect to the radius. To obtain physical solutions, we introduce the following two Dirichlet boundary conditions for $u(r)$:

$$u(r \to \infty) \longrightarrow 0 \quad \text{and} \quad u(r \to 0) \longrightarrow 0 \tag{6.37}$$

The first boundary condition is necessary for the radial wave function to be normalizable. The second boundary condition must be fulfilled because otherwise the radial wave function $R(r) = u(r)/r$ would diverge at $r = 0$.

Second Modification of the Radial Wave Function We have now transformed the initial differential equation into the following form:

$$\left(\frac{\partial^2}{\partial r^2} - \frac{l(l+1)}{r^2} + \frac{2Z}{a_0 r} - k^2\right) u(r) = 0$$

We can find another ansatz for $u(r)$ that modifies the differential equation and brings us closer to a solution:

$$u(r) = e^{-kr}\, y(r) \tag{6.38}$$

The form of $y(r) \equiv y$ will be discussed in the following sections. For now, we substitute the ansatz (6.38) into the above differential equation:

$$\begin{aligned}
0 &= \left(\frac{\partial^2}{\partial r^2} - \frac{l(l+1)}{r^2} + \frac{2Z}{a_0 r} - k^2\right) y e^{-kr} = \\
&= \frac{\partial^2}{\partial r^2}\left(y e^{-kr}\right) - \frac{l(l+1)}{r^2} y e^{-kr} + \frac{2Z}{a_0 r} y e^{-kr} - k^2 y e^{-kr} = \\
&= \frac{\partial}{\partial r}\left(y' e^{-kr} - k y e^{-kr}\right) - \frac{l(l+1)}{r^2} y e^{-kr} + \frac{2Z}{a_0 r} y e^{-kr} - k^2 y e^{-kr} = \\
&= y'' e^{-kr} - k y' e^{-kr} - k y' e^{-kr} + k^2 y e^{-kr} - \frac{l(l+1)}{r^2} y e^{-kr} + \frac{2Z}{a_0 r} y e^{-kr} - k^2 y e^{-kr} = \\
&= \left[y'' - 2ky' - \left(\frac{l(l+1)}{r^2} - \frac{2Z}{a_0 r}\right) y\right] e^{-kr}
\end{aligned}$$

We can now cancel out the exponential function e^{-kr}; similarly, the k^2 term has vanished and the wave number k now appears as an exponential damping term in (6.38). This leads to the following differential equation:

$$\left[\frac{\partial^2}{\partial r^2} - 2k\frac{\partial}{\partial r} - \frac{1}{r}\left(\frac{l(l+1)}{r} - \frac{2Z}{a_0}\right)\right] y(r) = 0 \tag{6.39}$$

To determine $y(r)$, two different methods can be chosen. We will first solve the above differential equation using the so-called Frobenius method by inserting a power series ansatz. Later, we will also introduce the *Laguerre differential equation*, for which ready-made solution formulas exist.

6.2.1 Solution Using Frobenius Method

As an ansatz solution, starting with the Frobenius method, we will choose a power series at this point. We can represent it as:

$$y(r) = r^{\alpha+1} \sum_{i=0}^{\infty} a_i r^i = \sum_{i=0}^{\infty} a_i r^{\alpha+1+i} \tag{6.40}$$

In (6.39), we need the first and second derivatives of our ansatz:

$$\frac{\partial y}{\partial r} = \frac{\partial}{\partial r}\left(\sum_{i=0}^{\infty} a_i r^{\alpha+1+i}\right) = \sum_{i=0}^{\infty} a_i \frac{\partial}{\partial r} r^{\alpha+1+i} = \sum_{i=0}^{\infty} a_i (\alpha+i+1) r^{\alpha+i} \tag{6.41}$$

$$\frac{\partial^2 y}{\partial r^2} = \frac{\partial}{\partial r}\left(\frac{\partial y}{\partial r}\right) \overset{(6.41)}{=} \frac{\partial}{\partial r}\left(\sum_{i=0}^{\infty} a_i (\alpha+i+1) r^{\alpha+i}\right) = \sum_{i=0}^{\infty} a_i (\alpha+i+1)(\alpha+i) r^{\alpha-1+i} \tag{6.42}$$

If we substitute the ansatz (6.40) and the derived derivative expressions (6.42) and (6.41) into (6.39), we obtain:

$$\begin{aligned}
0 &= \left[\frac{\partial^2}{\partial r^2} - 2k\frac{\partial}{\partial r} - \frac{1}{r}\left(\frac{l(l+1)}{r} - \frac{2Z}{a_0}\right)\right] r^{\alpha+1}\sum_{i=0}^{\infty} a_i r^i \overset{(6.41)(6.42)}{=} \\
&= \sum_{i=0}^{\infty} a_i \left[(\alpha+i+1)(\alpha+i)r^{\alpha-1+i} - 2k(\alpha+i+1)r^{\alpha+i} - \frac{1}{r}\left(\frac{l(l+1)}{r} - \frac{2Z}{a_0}\right) r^{\alpha+1+i}\right] = \\
&= \sum_{i=0}^{\infty} a_i \left[(\alpha+i+1)(\alpha+i)r^{\alpha-1+i} - 2k(\alpha+i+1)r^{\alpha+i} - \left(l(l+1)r^{-1} - \frac{2Z}{a_0}\right) r^{\alpha+i}\right] = \\
&= \sum_{i=0}^{\infty} a_i \left[(\alpha+i+1)(\alpha+i)r^{\alpha-1+i} - 2k(\alpha+i+1)r^{\alpha+i} - l(l+1)r^{\alpha-1+i} + \frac{2Z}{a_0} r^{\alpha+i}\right] = \\
&= \sum_{i=0}^{\infty} a_i \left\{\Big[(\alpha+i+1)(\alpha+i) - l(l+1)\Big] r^{\alpha-1+i} + \left[\frac{2Z}{a_0} - 2k(\alpha+i+1)\right] r^{\alpha+i}\right\} \qquad (6.43)
\end{aligned}$$

We find that r appears with the powers $r^{\alpha-1+i}$ and $r^{\alpha+i}$. If we explicitly write out the sum, each individual r power term must vanish to satisfy the differential equation. This allows us to determine α. Let's first consider the term with the smallest power of r: This occurs at $i = 0$. In this case, there is a term with $r^{\alpha-1}$ and one with r^{α}. Obviously, $r^{\alpha-1}$ is the smallest power term that can occur. It must also be zero, which means:

$$\begin{aligned}
0 &= a_0\left[(\alpha+1)\alpha - l(l+1)\right] r^{\alpha-1} = \quad | \div a_0 r^{\alpha-1} \\
&= \alpha(\alpha+1) - l(l+1) = \alpha^2 + \alpha - l(l+1)
\end{aligned}$$

We can now solve a simple quadratic equation for α, obtaining two solutions:

$$\alpha_{1,2} = -\frac{1}{2} \pm \sqrt{\frac{1}{4} + l(l+1)} = -\frac{1}{2} \pm \frac{1}{2}\sqrt{(2l+1)^2} = \frac{-1 \pm (2l+1)}{2}$$

Only one of these solutions will have actual physical relevance. We find:

$$\alpha \implies \begin{cases} \alpha_1 = l \\ \alpha_2 = -(l+1) \quad \text{(unphysical)} \end{cases} \qquad (6.44)$$

How do we know that only the solution α_1 is physical? For this, we must first link the original radial wave function $R(r)$ over $u(r)$ and $y(r)$ with our ansatz (6.40):

$$R(r) \overset{(6.33)}{=} \frac{u(r)}{r} \overset{(6.38)}{=} \frac{1}{r} e^{-kr}\, y(r) \overset{(6.40)}{=} \frac{1}{r} e^{-kr}\, r^{\alpha+1} \sum_{i=0}^{\infty} a_i r^i = e^{-kr} \sum_{i=0}^{\infty} a_i r^{\alpha+i} \qquad (6.45)$$

Let's try inserting the two possible solutions for α from (6.44) into (6.45) and check if we can find solutions over the entire existence range of the radius r. We distinguish:

- $\boldsymbol{\alpha = l}$: Replacing α with the orbital angular momentum quantum number $\alpha_1 = l$, we obtain:

$$R(r) = e^{-kr} \sum_{i=0}^{\infty} a_i r^{l+i} \implies u(r) = r e^{-kr} \sum_{i=0}^{\infty} a_i r^{l+i} = e^{-kr} \sum_{i=0}^{\infty} a_i r^{l+i+1}$$

  For all sum terms $i \geq 0$, it holds that $R(r)$ and $u(r)$ remain normalizable since there can be no singularity in the allowed interval of r. Especially $r \to 0$ poses no problems here, not even for $i = 0$ and $l = 0$.

- $\boldsymbol{\alpha = -(l+1)}$: When we now insert our found result $\alpha_2 = -(l+1)$ for α, it follows:

$$R(r) = e^{-kr} \sum_{i=0}^{\infty} a_i r^{i-l-1} \implies u(r) = re^{-kr} \sum_{i=0}^{\infty} a_i r^{i-l-1} = e^{-kr} \sum_{i=0}^{\infty} a_i r^{i-l}$$

We recognize that, in the case of $r \to 0$ and $i = 0$, $l \geq 1$, a singularity arises in $u(r)$, and $u(r)$ would not be normalizable in this form. Considering the special case $l = 0$, although there is no problem with normalizing $u(r)$, the actual radial wave function $R(r)$ would have a singularity at the sum term $i = 0$:

$$R(r \to 0) = e^{-kr} \sum_{i=0}^{\infty} a_i r^{i-1} \longrightarrow \frac{1}{r}$$

At the transition to the limit, the r^{-1} term dominates! $R(r)$ is then inserted into the Schrödinger equation (6.23), resulting in a Δr^{-1} term that needs further discussion. For this, it holds:

$$\Delta \frac{1}{r} = -4\pi\delta(r)$$

Thus, we obtain a δ function in the Schrödinger equation, preventing a comprehensive solution of the differential equation. Therefore, the special case $l = 0$ is not a possible solution for the radial wave functions, and $\alpha_2 = -(l+1)$ must be excluded overall.

Example: Δr^{-1}

In the previous argument, we used the following relation (6.46), but we now want to derive in more detail how this expression arises. Initially, it holds:

$$\Delta \frac{1}{r} = -4\pi\delta(r) \tag{6.46}$$

r represents the radius in a spherical coordinate system; we must therefore also adjust our Laplace operator Δ according to (6.19) to the chosen coordinates. When we insert Δ in spherical coordinates into (6.46) for the case of $r \neq 0$, we obtain:

$$\Delta \frac{1}{r} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) \frac{1}{r} = -\frac{1}{r^2} \frac{\partial}{\partial r} \left(\frac{r^2}{r^2} \right) = 0$$

If we now perform the integration over an arbitrary sphere with radius r_0, using Gauss's theorem, or the Nabla operator $\boldsymbol{\nabla}$ in polar coordinates, we can derive the following expression:

$$\int_V dV\, \Delta \frac{1}{r} = \int_V dV\, \boldsymbol{\nabla}\boldsymbol{\nabla} \frac{1}{r} = \int_{\partial V} dS\, \mathbf{r}\boldsymbol{\nabla} \frac{1}{r} = \int_{\partial V} r_0^2 d\Omega\, \mathbf{rr} \frac{\partial}{\partial r} \frac{1}{r} \Big|_{r=r_0} = -\int_{\partial V} d\Omega = -4\pi$$

As we previously recognized that in the case of $r \neq 0$, the relation $\Delta r^{-1} = 0$ must hold, we only obtain this result of the integration if $r = 0$. With the delta function, we can enforce the corresponding behavior, and it follows:

$$\Delta \frac{1}{r} = -4\pi\delta(r) \quad \square$$

Only $\alpha = l$ is therefore physically meaningful as a solution. We further require that the Dirichlet boundary condition, i.e., $u(r \to 0) \longrightarrow 0$, must be fulfilled. We now use the found relation $\alpha = l$

in our sum equation from (6.43) and further simplify it to obtain a recursive relationship between the development coefficients a_i:

$$
\begin{aligned}
0 &= \sum_{i=0}^{\infty} a_i \left\{ \Big[(l+i+1)(l+i) - l(l+1)\Big] r^{l-1+i} + \left[\frac{2Z}{a_0} - 2k(l+i+1)\right] r^{l+i} \right\} = \\
&= \sum_{i=0}^{\infty} a_i \left\{ \Big[l^2 + li + l + li + i^2 + i + i - i - l(l+1)\Big] r^{l-1+i} + \left[\frac{2Z}{a_0} - 2k(l+i+1)\right] r^{l+i} \right\} = \\
&= \sum_{i=0}^{\infty} a_i \left\{ \Big[l^2 + l + 2li + 2i + i^2 - i - l(l+1)\Big] r^{l-1+i} + \left[\frac{2Z}{a_0} - 2k(l+i+1)\right] r^{l+i} \right\} = \\
&= \sum_{i=0}^{\infty} \left\{ a_i \Big[l(l+1) + \underbrace{2i(l+1) + i(i-1)}_{\text{vanishes at } i=0} - l(l+1)\Big] r^{l-1+i} + a_i \left[\frac{2Z}{a_0} - 2k(l+i+1)\right] r^{l+i} \right\} = \\
&= \sum_{i=1}^{\infty} a_i \Big[2i(l+1) + i(i-1)\Big] r^{l-1+i} - 2\sum_{i=0}^{\infty} a_i \left[k(l+i+1) - \frac{Z}{a_0}\right] r^{l+i} = \quad |i = n+1 \\
&= \sum_{n=0}^{\infty} a_{n+1} \Big[2(n+1)(l+1) + (n+1)n\Big] r^{l+n} - \sum_{i=0}^{\infty} 2a_i \left[k(l+i+1) - \frac{Z}{a_0}\right] r^{l+i} = \quad |n \to i \\
&= \sum_{i=0}^{\infty} \left\{ a_{i+1} \Big[2(l+1)(i+1) + i(i+1)\Big] - 2a_i \left[k(l+i+1) - \frac{Z}{a_0}\right] \right\} r^{l+i}
\end{aligned}
\tag{6.47}
$$

We have managed to manipulate both sum terms in such a way that instead of two different powers of r, we now have two different development coefficients a_{i+1} and a_i. A coefficient comparison is now very easy: To fulfill the equation, the entire expression in the curly brackets must vanish for each power r^{l+i}:

$$
\begin{aligned}
a_{i+1}\left[2(l+1)(i+1) + i(i+1)\right] - 2a_i \left[k(l+i+1) - \frac{Z}{a_0}\right] = 0 \implies \\
a_{i+1}\left[2(l+1)(i+1) + i(i+1)\right] = 2a_i \left[k(l+i+1) - \frac{Z}{a_0}\right]
\end{aligned}
$$

From this, we can derive the following recursion relation between a_{i+1} and a_i:

$$
a_{i+1} = a_i \frac{2[k(l+i+1) - Z/a_0]}{2(l+1)(i+1) + i(i+1)} \tag{6.48}
$$

This corresponds to a two-term recursion formula for the Laguerre functions. If we know the starting value a_0, we can write down all radial wave functions using (6.48) and a subsequent normalization.

Recursion Formula We must check again whether it will diverge at any point. For this, we calculate the ratio between two successive development coefficients by dividing the recursion equation (6.48) by a_i:

$$
\begin{aligned}
\frac{a_{i+1}}{a_i} = \frac{2[k(l+i+1) - Z/a_0]}{2(l+1)(i+1) + i(i+1)} = \frac{2[ki + k(l+1) - Z/a_0]}{2(l+1)(i+1) + i(i+1)} = \\
= \frac{2ki}{i^2 + i + 2(l+1)(i+1)} + \frac{2[k(l+1) - Z/a_0]}{i^2 + i + 2(l+1)(i+1)}
\end{aligned}
$$

Letting i go to infinity, we can immediately neglect the second term and also recognize that in the denominator, the quadratic term will dominate. Therefore:

$$
\frac{a_{i+1}}{a_i} \xrightarrow{i\to\infty} \frac{2ki}{i^2} = \frac{2k}{i} \tag{6.49}
$$

Similar to the harmonic oscillator, we now try to determine which function corresponds to this asymptotic behavior (6.49). We find:

$$e^{2kr} = \sum_{i=0}^{\infty} \frac{(2kr)^i}{i!} \equiv \sum_{i=0}^{\infty} b_i r^i \tag{6.50}$$

Again, we form the ratio between two successive development coefficients based on this Taylor expansion:

$$\frac{b_{i+1}}{b_i} = \frac{(2k)^{i+1}}{(i+1)!}\frac{i!}{(2k)^i} = \frac{2k}{i+1} \xrightarrow{i\to\infty} \frac{2k}{i}$$

We have found a function that exhibits the same behavior as the development coefficients a_i of the radial wave function (6.45) in the limit $i \to \infty$. The i-th term of the series (6.45) reads:

$$R^{(i)}(r) = e^{-kr} a_i \, r^{l+i}$$

Setting i to infinity, we can, as just shown, replace a_i with e^{2kr}:

$$R(r) = e^{-kr} \sum_{i=0}^{\infty} a_i \, r^{l+i} \xrightarrow{i\to\infty} \frac{2k}{i} \sum_{i=0}^{\infty} e^{-kr} a_i \, r^{l+i} \overset{(6.50)}{=} \sum_{i=0}^{\infty} e^{-kr} e^{2kr} \, r^{l+i} = \sum_{i=0}^{\infty} e^{kr} \, r^{l+i}$$

We now have the problem that with our exponential ansatz for the development coefficients, the Dirichlet boundary condition $u(r \to \infty) = R(r \to \infty) \longrightarrow 0$ cannot be satisfied, as the exponential function e^{kr} diverges. To avoid divergence in the limit case, the recursion relationship must be truncated at some point. We therefore define a cut-off condition $i = n_r$ for the recursion relation from (6.48), where we refer to n_r as the *radial quantum number*. For all $i > n_r$, all a_i must vanish. For this condition to be met:

$$0 \overset{!}{=} a_{n_r+1} \overset{(6.48)}{=} a_{n_r} \frac{2[k(l+n_r+1) - Z/a_0]}{2(l+1)(n_r+1) + n_r(n_r+1)} \implies \underbrace{k\,(l+n_r+1)}_{n} - \frac{Z}{a_0} = 0 \tag{6.51}$$

Since n_r indicates the maximum polynomial degree of the power series (6.45), the radial quantum number naturally also determines the number of zeros or the number of nodes of the radial wave function. We can now, as shown in (6.51), define the *energy* or *principal quantum number* n:

$$n = l + n_r + 1 \quad \text{with} \quad n \geq 1 \tag{6.52}$$

Unlike the energy quantum number in the harmonic oscillator (which is also represented by n), in the case of the hydrogen atom, n cannot be zero. The orbital angular momentum quantum number l can be written according to (6.52) as:

$$l = n - n_r - 1 \quad \text{with} \quad 0 \leq l \leq n-1 \tag{6.53}$$

We already know that $l \geq 0$ must hold from the theory of angular momentum. $l = n-1$ applies in the special case that $n_r = 0$. By simple transformation of (6.51), we can also express the wave number $k \equiv k_n$ as a discrete function of the principal quantum number n:

$$k_n = \frac{Z}{a_0 n} \tag{6.54}$$

Substituting this into (6.31), we can assign an energy ε_n to each principal quantum number n:

$$\varepsilon_n = -\frac{\hbar^2 k_n^2}{2\mu} = -\frac{\hbar^2}{2\mu}\frac{Z^2}{a_0^2 n^2} \tag{6.55}$$

The principal quantum number n thus represents different (discrete) energy levels of the hydrogen atom! We can summarize the constants in (6.55) in the so-called *Rydberg constant* Ry and obtain the following compact expression for the eigenenergies, which increases quadratically with the principal quantum number n:

$$\varepsilon_n = -\mathrm{Ry}\frac{Z^2}{n^2} \quad \text{with} \quad \mathrm{Ry} = \frac{\hbar^2}{2\mu a_0^2} \tag{6.56}$$

The energy ε_n is now written without absolute value, reminding us that as a bound state, negative eigenenergies *must* have. Through the inverse proportionality to $1/n^2$, we approach with increasing energy quantum number $\varepsilon_n \longrightarrow 0$; the energy necessary to lift an electron from a level n to $\varepsilon_n = 0$ is called *ionization energy*. Reaching this energy, the electron can „escape“ from the atom, leaving a proton behind in the case of hydrogen: The hydrogen atom H is ionized to H^+.
Applying less energy results in a spectral transition, formally described earlier in (1.44). By adding energy, we lift the electron from a lower level to a higher one; the electron can also fall from a higher-energy level to a lower one. For $n > m$, we convert the following energy in the form of a photon:

$$\hbar\omega_{nm} = \varepsilon_n - \varepsilon_m = \mathrm{Ry}\left(\frac{Z^2}{m^2} - \frac{Z^2}{n^2}\right)$$

The radial eigenfunctions can be expressed by using the expression (6.45), remembering that $\alpha = l$, and letting the series run only up to $i = n_r$. We then obtain:

$$R^l_{n_r}(r) = e^{-kr}\sum_{i=0}^{n_r} a_i r^{l+i} \tag{6.57}$$

The development coefficients must be calculated successively using the recursion relation from (6.48). In the following section, we will deal with a solution strategy that leads directly to a wave function for n_r and l.

6.2.2 Solution via Laguerre Differential Equation

We have so far treated the differential equation of the radial wave function using the Frobenius method with a power series ansatz and successfully determined the eigenenergies. At this point, we want to abandon the power series ansatz and instead switch to the following ansatz with a general function $q(r) \equiv q$:

$$y(r) = r^{\alpha+1}q(r) \tag{6.58}$$

Before we substitute this ansatz into the initial equation (6.39), we calculate the first and second derivatives of our ansatz from (6.58) as a side calculation:

$$\begin{aligned}
\frac{d}{dr}\left(r^{\alpha+1}q\right) &= (\alpha+1)r^{\alpha}q + r^{\alpha+1}\frac{dq}{dr} = r^{\alpha+1}\left[\frac{(\alpha+1)}{r}q + \frac{dq}{dr}\right] \\
\frac{d^2}{dr^2}\left(r^{\alpha+1}q\right) &= \frac{d}{dr}\left[(\alpha+1)r^{\alpha}q + r^{\alpha+1}\frac{dq}{dr}\right] = \\
&= \alpha(\alpha+1)r^{\alpha-1}q + (\alpha+1)r^{\alpha}\frac{dq}{dr} + (\alpha+1)r^{\alpha}\frac{dq}{dr} + r^{\alpha+1}\frac{d^2q}{dr^2} = \\
&= r^{\alpha+1}\left[\frac{\alpha(\alpha+1)}{r^2}q + \frac{2(\alpha+1)}{r}\frac{dq}{dr} + \frac{d^2q}{dr^2}\right]
\end{aligned} \tag{6.59}$$

Inserting now (6.58) and the two expressions from (6.59) into the initial differential equation (6.39), we obtain:

$$\begin{aligned}
0 &= \left[\frac{d^2}{dr^2} - 2k\frac{d}{dr} - \frac{1}{r}\left(\frac{l(l+1)}{r} - \frac{2Z}{a_0}\right)\right] r^{\alpha+1}q(r) = \\
&= \frac{d^2}{dr^2}\left(r^{\alpha+1}q\right) - 2k\frac{d}{dr}\left(r^{\alpha+1}q\right) - r^{\alpha+1}\frac{l(l+1)}{r^2}q + r^{\alpha+1}\frac{2Z}{a_0 r}q \overset{(6.59)}{=} \\
&= r^{\alpha+1}\left[\frac{\alpha(\alpha+1)}{r^2}q + \frac{2(\alpha+1)}{r}\frac{dq}{dr} + \frac{d^2q}{dr^2} - 2k\frac{(\alpha+1)}{r}q - 2k\frac{dq}{dr} - \frac{l(l+1)}{r^2}q + \frac{2Z}{a_0 r}q\right] = \\
&= r^{\alpha+1}\left[\frac{\alpha(\alpha+1) - l(l+1)}{r^2}q + \frac{2(\alpha+1)\frac{dq}{dr} - 2k(\alpha+1)q}{r} + \frac{d^2q}{dr^2} - 2k\frac{dq}{dr} + \frac{2Z}{a_0 r}q\right] = \\
&= r^{\alpha+1}\left[\frac{\alpha(\alpha+1) - l(l+1)}{r^2}q + 2(\alpha+1)\frac{\frac{dq}{dr} - kq}{r} + \frac{d^2q}{dr^2} - 2k\frac{dq}{dr} + \frac{2Z}{a_0 r}q\right]
\end{aligned} \tag{6.60}$$

At this point, we perform the following substitution:

$$r = \frac{y}{k} \implies \frac{dr}{dy} = \frac{1}{k} \implies \frac{dy}{dr} = k \tag{6.61}$$

With this substitution, the function $q \equiv q(r)$ becomes a function $q \equiv q(y)$, and we must appropriately consider the following transformation:

$$\frac{dq}{dr} = \frac{dq}{dy}\frac{dy}{dr} \overset{(6.61)}{=} \frac{dq}{dy}k \tag{6.62}$$

We again substitute (6.61) and (6.62) into (6.60):

$$\begin{aligned}
0 &= \left(\frac{y}{k}\right)^{\alpha+1}\left[\frac{\alpha(\alpha+1) - l(l+1)}{(y/k)^2}q + 2(\alpha+1)\frac{k\frac{dq}{dy} - kq}{y/k} + k^2\frac{d^2q}{dy^2} - 2k^2\frac{dq}{dy} + \frac{2Zk}{a_0 y}q\right] = \\
&= \left(\frac{y}{k}\right)^{\alpha+1}\left[\frac{\alpha(\alpha+1) - l(l+1)}{y^2}k^2q + 2(\alpha+1)\frac{k^2}{y}\left(\frac{dq}{dy} - q\right) + k^2\frac{d^2q}{dy^2} - 2k^2\frac{dq}{dy} + \frac{2Z}{a_0 k}q\right] = \\
&= \frac{k^2}{y}\left(\frac{y}{k}\right)^{\alpha+1}\left[\frac{\alpha(\alpha+1) - l(l+1)}{y}q + 2(\alpha+1)\left(\frac{dq}{dy} - q\right) + y\frac{d^2q}{dy^2} - 2y\frac{dq}{dy} + \frac{2Z}{a_0 k}q\right] = \\
&= \frac{y^\alpha}{k^{\alpha-1}}\left\{y\frac{d^2q}{dy^2} + [2(\alpha+1) - 2y]\frac{dq}{dy} + \left[\frac{\alpha(\alpha+1) - l(l+1)}{y} - 2(\alpha+1) + 2\frac{Z}{a_0 k}\right]q\right\}
\end{aligned} \tag{6.63}$$

Correspondingly, we can, through a further simple transformation – corresponding to the last term in (6.63) – from (6.54) write:

$$n = \frac{Z}{a_0 k} \tag{6.64}$$

We need to substitute again to achieve a simplification. We define:

$$y = \frac{x}{2} \implies \frac{dy}{dx} = \frac{1}{2} \implies \frac{dx}{dy} = 2 \tag{6.65}$$

With this substitution, the function $q \equiv q(y)$ becomes a function $q \equiv q(x)$, and we must appropriately consider the following transformation:

$$\frac{dq}{dy} = \frac{dq}{dx}\frac{dx}{dy} \overset{(6.65)}{=} 2\frac{dq}{dx} \tag{6.66}$$

We now insert (6.64), (6.65), and (6.66) into (6.63):

$$\begin{aligned}
0 &= \frac{(x/2)^\alpha}{k^{\alpha-1}} \left\{ \frac{x}{2} 4 \frac{d^2q}{dx^2} + [2(\alpha+1) - x]\, 2\frac{dq}{dx} + \left[\frac{\alpha(\alpha+1) - l(l+1)}{x/2} - 2(\alpha+1) + 2n \right] q \right\} = \\
&= \frac{x^\alpha}{2^\alpha k^{\alpha-1}} \left\{ 2x \frac{d^2q}{dx^2} + 2\,[2(\alpha+1) - x] \frac{dq}{dx} + 2\left[\frac{\alpha(\alpha+1) - l(l+1)}{x} - (\alpha+1) + n \right] q \right\} = \\
&= \frac{x^\alpha}{(2k)^{\alpha-1}} \left\{ x \frac{d^2q}{dx^2} + [2(\alpha+1) - x] \frac{dq}{dx} + \left[\frac{\alpha(\alpha+1) - l(l+1)}{x} - (\alpha+1) + n \right] q \right\}
\end{aligned}$$

Since the entire expression should vanish, we can cancel the term before the curly brackets:

$$\left\{ x \frac{d^2}{dx^2} + [2(\alpha+1) - x] \frac{d}{dx} + \left[\frac{\alpha(\alpha+1) - l(l+1)}{x} - (\alpha+1) + n \right] \right\} q(x) = 0$$

Additionally, we know from (6.44) that $\alpha = l$. Therefore, another term cancels out:

$$\begin{aligned}
0 &= \left\{ x \frac{d^2}{dx^2} + [2(l+1) - x] \frac{d}{dx} + \left[\frac{l(l+1) - l(l+1)}{x} - (l+1) + n \right] \right\} q(x) = \\
&= \left\{ x \frac{d^2}{dx^2} + [\underbrace{2l+1}_{\beta} + 1 - x] \frac{d}{dx} + \underbrace{(n-l-1)}_{w} \right\} q(x) = \\
&= xq'' + (\beta + 1 - x)q' + wq \quad \text{with} \quad \beta = 2l+1 \quad \text{and} \quad w = n-l-1
\end{aligned} \tag{6.67}$$

With the substitutions $\beta = 2l + 1$ and $w = n - l - 1$, we have brought this equation into a form corresponding to the Laguerre differential equation from (10.28). In Chapter 10.4 of the appendix, it is derived that the Laguerre polynomials $L_w^\beta(x)$ represent the solutions of this differential equation, which can be expressed, for example, through the Rodrigues formula (10.31):

$$L_w^\beta(x) = \frac{1}{w!} x^{-\beta} e^x \frac{d^w(x^{w+\beta} e^{-x})}{dx^w} \tag{6.68}$$

The solutions of the differential equation (6.67) depending on the quantum numbers n and l can be written as:

$$q(x) = L_w^\beta(x) \quad \text{with} \quad \beta = 2l+1 \quad \text{and} \quad w = n-l-1 \tag{6.69}$$

To determine the radial wave function $R(r) \equiv R_n^l(r)$, the substitutions (6.33), (6.38), (6.58), and (6.44) must first be reversed before we can use the solution (6.69):

$$\begin{aligned}
R_n^l(r) &\overset{(6.33)}{=} \frac{1}{r} u(r) \overset{(6.38)}{=} \frac{1}{r} e^{-kr} y(r) \overset{(6.58)}{=} \frac{1}{r} e^{-kr} r^{\alpha+1} q(r) \overset{(6.44)}{=} e^{-kr} r^l q(r) \overset{(6.69)}{\Longrightarrow} \\
R_n^l(r) &= e^{-kr} r^l L_w^\beta(x(r))
\end{aligned} \tag{6.70}$$

By reversing the substitutions (6.65), (6.61), and (6.64), we obtain $x = 2y$, $y = kr$, and $k = Z/(a_0 n)$. Thus, we achieve that the right side of equation (6.70) directly depends on r:

$$x \overset{(6.65)}{=} 2y \overset{(6.61)}{=} 2kr \overset{(6.64)}{=} \frac{2Zr}{a_0 n} \tag{6.71}$$

$$R_n^l(r) \overset{(6.70)}{=} e^{-kr} r^l L_w^\beta(x) \overset{(6.64)}{=} e^{-\frac{Zr}{a_0 n}} r^l L_w^\beta(x) \overset{(6.71)}{=} e^{-\frac{Zr}{a_0 n}} r^l L_w^\beta\left(\frac{2Zr}{a_0 n}\right) \tag{6.72}$$

We substitute for x in $L_w^\beta(x)$ the expression $(2Zr)/(a_0 n)$. Thus, we must also replace the w-th derivative with respect to x in the Rodrigues formula (6.68) with the w-th derivative with respect to r. For this, we calculate:

$$\frac{dx}{dr} \overset{(6.71)}{=} \frac{2Z}{a_0 n} \implies \frac{dr}{dx} = \frac{a_0 n}{2Z} \tag{6.73}$$

and

$$\begin{aligned}
\frac{d^w(x^{w+\beta}e^{-x})}{dx^w} &= \frac{d^w(x^{w+\beta}e^{-x})}{dr^w}\frac{d^w r}{dx^w} \overset{(6.73)}{=} \frac{d^w(x^{w+\beta}e^{-x})}{dr^w}\left(\frac{a_0 n}{2Z}\right)^w \overset{(6.71)}{=} \\
&= \left(\frac{a_0 n}{2Z}\right)^w \frac{d^w}{dr^w}\left[\left(\frac{2Zr}{a_0 n}\right)^{w+\beta} e^{-\frac{2Zr}{a_0 n}}\right] = \\
&= \left(\frac{a_0 n}{2Z}\right)^w \left(\frac{2Z}{a_0 n}\right)^w \left(\frac{2Z}{a_0 n}\right)^\beta \frac{d^w}{dr^w}\left[r^{w+\beta} e^{-\frac{2Zr}{a_0 n}}\right] = \\
&= \left(\frac{2Z}{a_0 n}\right)^\beta \frac{d^w}{dr^w}\left[r^{w+\beta} e^{-\frac{2Zr}{a_0 n}}\right]
\end{aligned} \tag{6.74}$$

Finally, substituting $\beta = 2l+1$ and $w = n-l-1$ into (6.72), and further considering the normalization $\int_0^\infty r^2 |R_n^l(r)|^2 dr = 1$, we can express the following expression for the radial wave function:

$$\begin{aligned}
R_n^l(r) &= \frac{2}{n^2}\sqrt{\frac{(n-l-1)!}{(n+l)!}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}}\left(\frac{2Zr}{a_0 n}\right)^l e^{-\frac{Zr}{a_0 n}} L_{n-l-1}^{2l+1}\left(\frac{2Zr}{a_0 n}\right) \\
\text{with } L_w^\beta(x) &\overset{(10.30)}{=} \sum_{m=0}^{w} (-1)^m \frac{(w+\beta)!}{(w-m)!(\beta+m)!m!} x^m
\end{aligned} \tag{6.75}$$

Using the Rodrigues formula (6.68) and the substitution (6.74), the radial wave function (6.75) can be rewritten as follows:

$$R_n^l(r) = \underbrace{\frac{2}{n^2}\sqrt{\frac{(n-l-1)!Z^3}{(n+l)!\quad a_0^3}}\left(\frac{2Zr}{a_0 n}\right)^l e^{-\frac{Zr}{a_0 n}}}_{\text{corresponding to (6.75)}} \underbrace{\frac{1}{(n-l-1)!}\left(\frac{2Zr}{a_0 n}\right)^{-2l-1} e^{\frac{2Zr}{a_0 n}} \frac{d^{n-l-1}}{dr^{n-l-1}}\left[r^{n+l} e^{-\frac{2Zr}{a_0 n}}\right]}_{L_{n-l-1}^{2l+1}\left(\frac{2Zr}{a_0 n}\right) \text{ according to (6.68) and (6.74)}}$$

After canceling, the following alternative expression for the radial wave function remains:

$$R_n^l(r) = \frac{2}{n^2}\sqrt{\frac{1}{(n-l-1)!(n+l)!}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}}\left(\frac{2Z}{a_0 n}\right)^l r^{-l-1} e^{\frac{Zr}{a_0 n}} \frac{d^{n-l-1}}{dr^{n-l-1}}\left[r^{n+l} e^{-\frac{2Zr}{a_0 n}}\right] \tag{6.76}$$

We recognize that only for $l = 0$ at $r = 0$ is the radial wave function R_n^0 different from zero; if the electron has an angular momentum component $l > 0$, the wave function vanishes at $r = 0$.

As a reminder: To fully represent an eigenfunction $\phi_{nlm}(r, \vartheta, \varphi)$ of the hydrogen atom in the state n, l, m, we must only use the result (6.75) or (6.76), and then reverse the separation (6.22):

$$\phi_{nlm}(r, \vartheta, \varphi) \overset{(6.22)}{=} R_n^l(r)\, Y_l^m(\vartheta, \varphi) \tag{6.77}$$

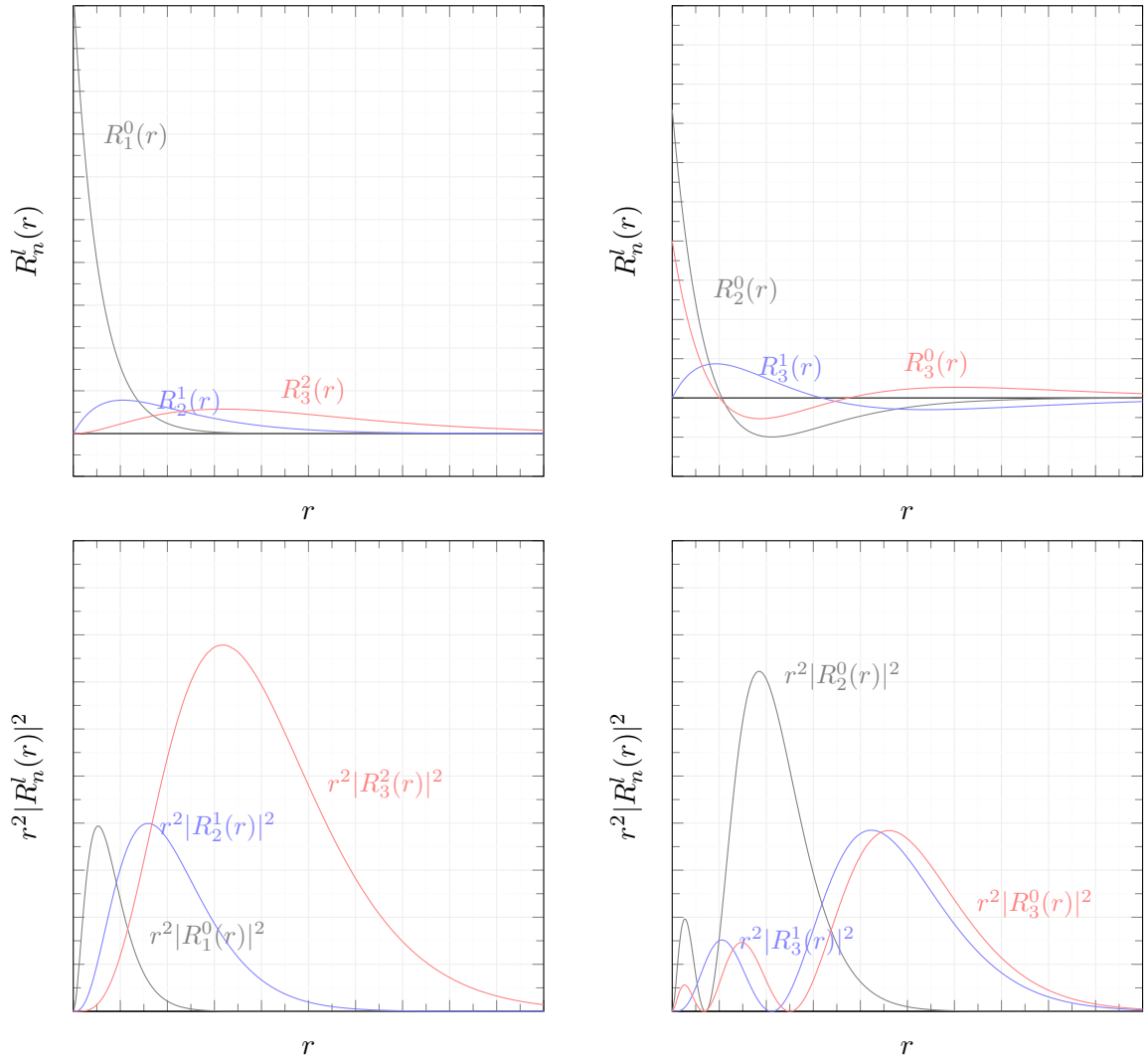


Fig. 35: Selected examples of radial wave functions $R_n^l(x)$ (top) and their probability densities $|R_n^l(x)|$ (bottom). (left) No nodes. (right) One (R_2^0, R_3^1) or two (R_3^0) radial nodes.

6.3 Energy Levels and Notation

The energy levels in the hydrogen atom follow a relatively simple law of the form:

$$\varepsilon_n = -\mathrm{Ry}\frac{1}{n^2}$$

The Rydberg constant $\mathrm{Ry} \approx 13.605\,\mathrm{eV}$ sets the lowest energy value, i.e., the ground state, at $n = 1$; due to the quadratic dependence on the principal quantum number n, the energy increases with increasing n. We notice that the energy ε_n *does not* depend on the other quantum numbers used to describe the angular components of the hydrogen atom: the orbital angular momentum and the magnetic quantum number l and m. Our energy levels are thus degenerate in l and m! Although this seems quite convenient at first glance, it is actually due to a "flaw" in our derivation: We calculated the eigenenergies of the hydrogen atom non-relativistically!
The flaw can be corrected by solving the *Dirac equation* instead of the Schrödinger equation, as it can merge quantum theory and relativity theory. In this case, however, we will find that our "solution" is an approximation of the actual solution, which sometimes lifts the degeneracy of l. However, the degeneracy of the magnetic quantum number m remains without further assumptions – only when we place our hydrogen atom in an external magnetic field $\mathbf{B}$ does it lead to

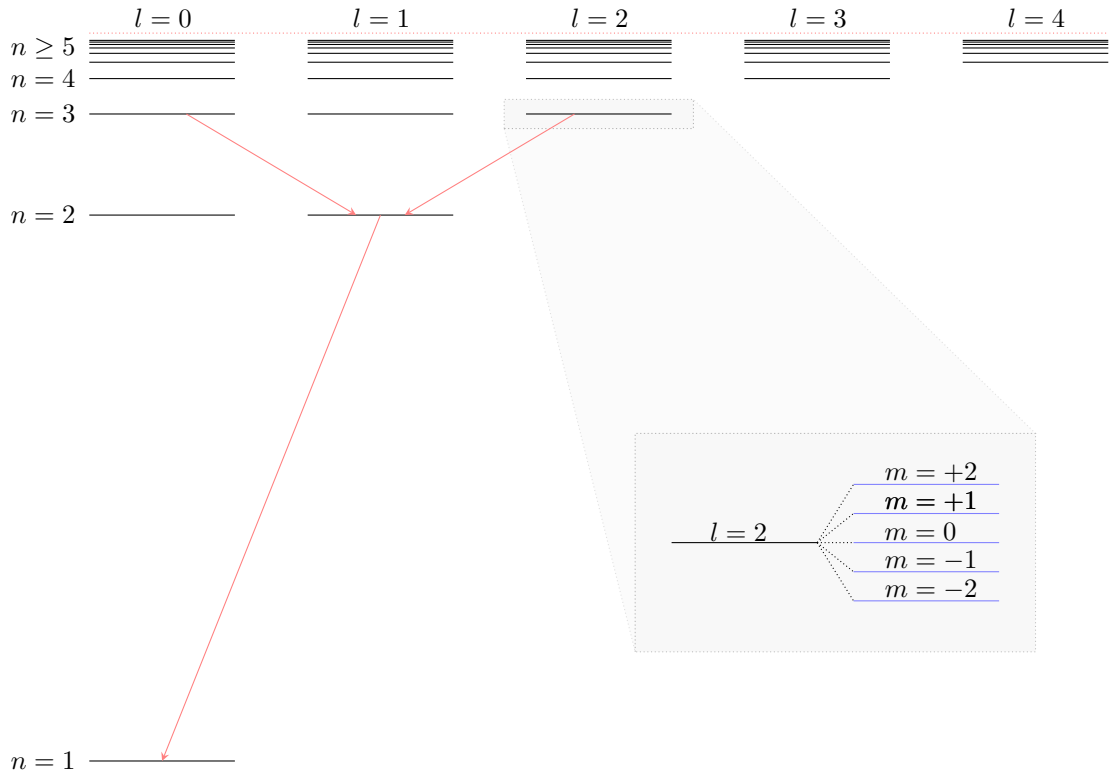


Fig. 36: Schematic representation of the energy levels of the hydrogen atom: Additionally highlighted is the state $|n, l\rangle = |3, 2\rangle$ and its magnetic splitting according to m.

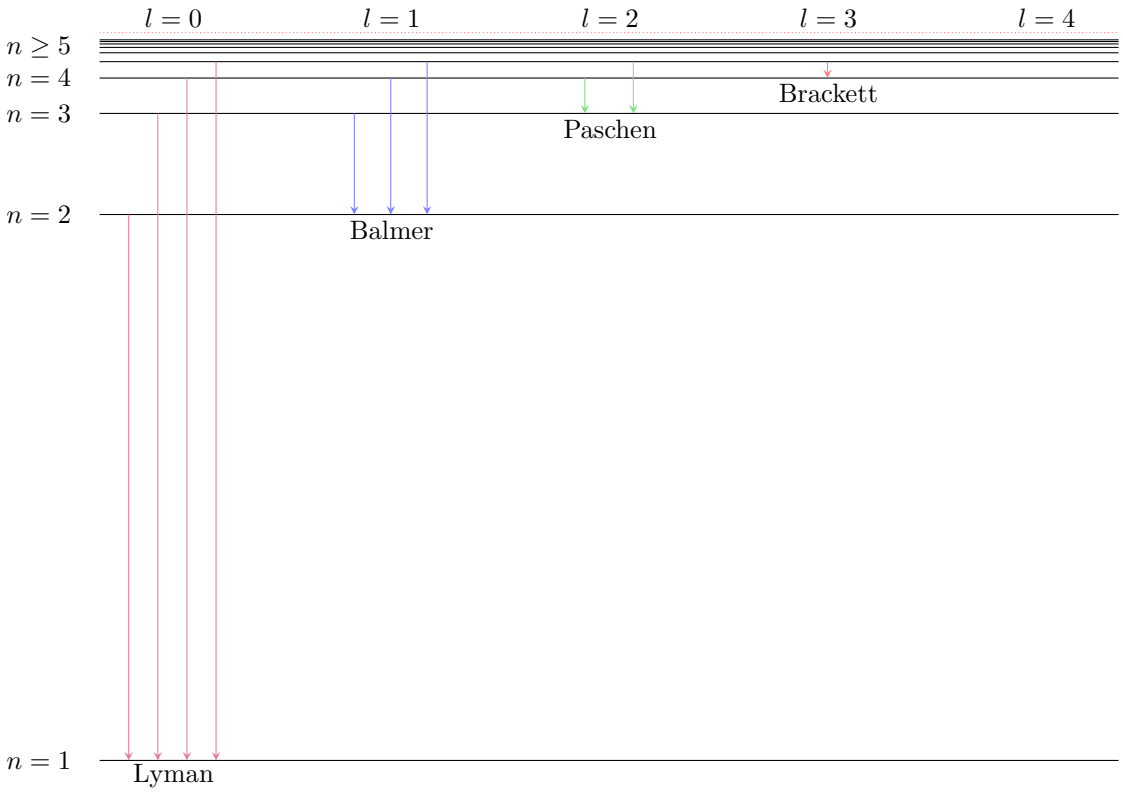


Fig. 37: Emission sequences of spectral transitions: According to (1.44), the energy of the emitted photon depends on the spacing of the spectral lines.

an energetic splitting according to m.

If an electron now finds itself in an orbital $n > 1$, it wants to fall back to the ground state $n = 1$ (or another more favorable state) for energetic reasons. This releases energy, which is emitted as a photon with frequency $\nu_{nm} = \Delta\varepsilon/h$ and corresponds to the difference in energy levels:

$$\Delta\varepsilon = \mathrm{Ry}\left(\frac{1}{m^2} - \frac{1}{n^2}\right)$$

We find various emission series characterized by the final state n of the decay (for example, a decay in the Balmer series *always* leads to $n = 2$).

So far, we have identified three quantum numbers to characterize the state of an electron bound in a hydrogen-like system: the principal quantum number n, the orbital angular momentum quantum number l, and the magnetic quantum number m. While n determines the energy of the electron and its maximum achievable angular momentum ($n-1 \geq l$), l determines the parity of the state through a $(-1)^l$ term from (5.123). m actually plays a role energetically only if an external magnetic field is present – otherwise, the magnetic quantum number with l influences the actual geometry of our state.

We can introduce a new notation for the individual angular momentum quantum numbers l:

$$l = 0, 1, 2, 3, 4, \ldots \implies l = s, p, d, f, g, \ldots \tag{6.78}$$

Let's consider an example of a state: Our electron is characterized by the quantum numbers $n = 3$ and $l = 1$. We designate this state in our (spectroscopic) shorthand notation as a $3p$-state.

7 Spin

Motivation: Spin

According to the Bohr's atomic model, electrons move in orbits around the nucleus, to which a quantized angular momentum can be assigned. Now the question arises whether the electron – similar to a planet – also has an intrinsic angular momentum (spin), or whether this is even possible for a point-like particle.

The Stern-Gerlach experiment conducted by Otto Stern and Walther Gerlach in 1922 represents the central experiment in this context, in which (unknowingly) the spin of the electron was first demonstrated (see Figure 39). A beam of silver atoms passes through an inhomogeneous magnetic field. The electron configuration of a silver atom consists of a closed shell plus a $5s$-electron. Due to this configuration, the total angular momentum of a silver atom should actually be zero. However, it is observed that the incoming atom beam is split into two outgoing beams in the inhomogeneous magnetic field. If the electron were in a p-orbital ($l = 1$), there would have to be three outgoing beams due to the degeneracy $(2l + 1)$. The observation of two beams can only be explained by a half-integer angular momentum of the electron. This internal angular momentum or spin of the electron was theoretically explained by Paul Dirac in 1928 using relativistic quantum mechanics.

7.1 Magnetic Moment

In the previous chapters, we have already discussed the electron in the hydrogen atom and calculated the geometric shape of the wave functions analytically in the position representation. Another quantity, which is related to the angular momentum $\mathbf{L}$, however, we have not yet discussed: the magnetic (dipole) moment $\boldsymbol{\mu}$.

Before we specifically describe the magnetic moment, let's recall the classical angular momentum $\mathbf{L}$. It can be written as follows:

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} = m\,(\mathbf{r} \times \mathbf{v}) = m\,(\mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r})) \tag{7.1}$$

Using the "bac-cab" rule $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}\,(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}\,(\mathbf{a} \cdot \mathbf{b})$ and $|\mathbf{r}| \equiv r$, this can be converted to:

$$\mathbf{L} = m[\boldsymbol{\omega}(\mathbf{r} \cdot \mathbf{r}) - \mathbf{r}(\mathbf{r} \cdot \boldsymbol{\omega})] = mr^2\boldsymbol{\omega} \tag{7.2}$$

If we are interested in the magnitude of the angular momentum $|\mathbf{L}|$, we get with $|\boldsymbol{\omega}| \equiv \omega$ and the rotation period $T = 2\pi/\omega$ or $\omega = 2\pi/T$:

$$|\mathbf{L}| = mr^2\frac{2\pi}{T} \tag{7.3}$$

The classical magnetic (dipole) moment is given in classical physics as the product of the current I and the area $\mathbf{A}$:

$$\boldsymbol{\mu} = I\mathbf{A} \tag{7.4}$$

Let's specifically consider a charged point particle (e.g., an electron) with charge q on a circular orbit with rotation period T – the current I then results as $I = q/T$. For the circular area, $|\mathbf{A}| = \pi r^2$ holds, where $\mathbf{A}$ points in the direction of the surface normal $\mathbf{e}_n$ and has the same orientation as $\mathbf{L}$ or $\boldsymbol{\omega}$. By inserting into (7.4), we get:

$$\boldsymbol{\mu} = \frac{q}{T}\pi r^2\,\mathbf{e}_n \tag{7.5}$$

Rearranging the magnitude of the angular momentum $|\mathbf{L}|$ from (7.3) for T and inserting into (7.5) provides a relationship between $\boldsymbol{\mu}$ and $\boldsymbol{L}$.

$$\boldsymbol{\mu} = \frac{q|\mathbf{L}|}{2m\pi r^2}\pi r^2\,\mathbf{e}_n = \frac{q}{2m}\mathbf{L} = -\frac{e}{2m_\mathrm{e}}\mathbf{L} \tag{7.6}$$

In the last step, since we mainly deal with the electron, we have set the charge q as the negative elementary charge $-e$, and m as the electron mass m_e. Due to the negative charge, the magnetic moment $\boldsymbol{\mu}$ points in the opposite direction of the angular momentum $\mathbf{L}$. The relative length of $\boldsymbol{\mu}$ with respect to $\mathbf{L}$ is called the *gyromagnetic ratio* γ_L and can be calculated as follows:

$$\gamma_\mathrm{L} = \frac{|\boldsymbol{\mu}|}{|\mathbf{L}|} \tag{7.7}$$

A consequence of the semi-classical Bohr model of the atom for the electron is the *Bohr magneton* $\mu_\mathrm{B} = 9.274 \times 10^{-24}\,\mathrm{J\,T^{-1}}$. It is defined as the magnitude of the magnetic moment produced by an electron with the lowest orbital angular momentum quantum number ($l = 1$ with $|\mathbf{L}| = 1\hbar$) in the Bohr model on its circular orbit (with Bohr radius a_0) around the nucleus:

$$\mu_\mathrm{B} = \frac{e\hbar}{2m_\mathrm{e}} \tag{7.8}$$

For the gyromagnetic ratio γ_L of the electron, we find using the Bohr magneton μ_B the relation:

$$\gamma_\mathrm{L} = \frac{e}{2m_\mathrm{e}}\frac{\hbar}{\hbar} \overset{(7.8)}{=} \frac{\mu_B}{\hbar} \tag{7.9}$$

So far the classical consideration. The transition from classical to quantum mechanics is achieved via the correspondence principle; the angular momentum operator $\hat{\mathbf{L}}$ has already been defined in (5.1) and is given by:

$$\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$$

Also, through the correspondence principle, we can now form a new operator for the observable of the magnetic moment $\hat{\boldsymbol{\mu}}$. Here we initially introduce an additional factor g_L without further assumptions, which we will generally refer to as the *Landé factor* of the orbital angular momentum. Later we will also find different Landé factors for other angular momentum sizes; in the case of the orbital angular momentum $\hat{\mathbf{L}}$, however, it trivially holds that $g_\mathrm{L} = 1$. Considering g_L and with the help of (7.6), we find:

$$\hat{\boldsymbol{\mu}} = \mathrm{sgn}(q)\frac{\mu_B}{\hbar}g_\mathrm{L}\,\hat{\mathbf{L}} \tag{7.10}$$

Here $\mathrm{sgn}(q)$ is the sign of the charge q of the considered particle; in the case of the electron, $\mathrm{sgn}(q) = \mathrm{sgn}(-e) = -1$. An equivalent, shortened formulation would be $\hat{\boldsymbol{\mu}} = \mathrm{sgn}(q)\gamma_\mathrm{L}\,\hat{\mathbf{L}}$. It can be seen from (7.10): If $\hat{\mathbf{L}}$ is quantized in units of $\hbar$, then $\hat{\boldsymbol{\mu}}$ is also quantized in units of μ_B. Also, if we now let $\hat{\boldsymbol{\mu}}$ act on an eigenstate of the angular momentum operator, we can derive simple results using the already known relations (5.59), (5.60), and (5.61).

In the previous chapters, we always chose the z-direction as the preferred axis of the angular momentum; in the x- and y-directions, we cannot obtain "sharp" measurement values due to the uncertainty relation. However, we can always rotate our coordinate system such that we can use the simple eigenvalue relationships (5.60) and (5.115). Also within this chapter, in the case of spin and general angular momenta, we will adhere to this convention and arbitrarily set the z-direction as the preferred axis.

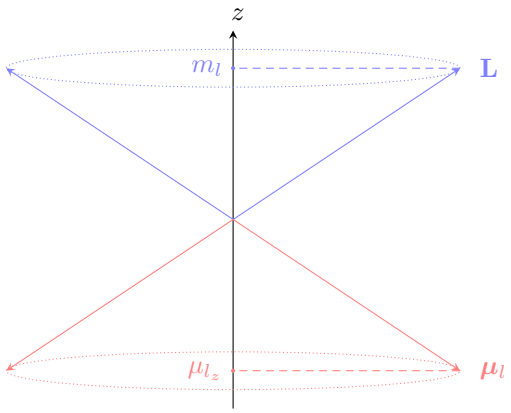


Fig. 38: Schematic representation of angular momentum $\hat{\mathbf{L}}$ and magnetic moment $\hat{\boldsymbol{\mu}}_l$, where for simplicity the gyromagnetic ratio $\gamma_\mathrm{L} = 1$ is set.

7.1.1 Energy in the Magnetic Field

The existence of a magnetic moment is interesting in the sense that in the presence of a magnetic field $\mathbf{B}$, there is a change in energy. Accordingly, the eigenenergies and eigenfunctions of the Hamiltonian operator $\hat{H}$ will change. Let's first consider the classical case of a magnetic dipole in an external magnetic field $\mathbf{B}$. The potential, magnetic energy is given by:

$$E_\mu = -\boldsymbol{\mu}\mathbf{B} \overset{(7.6)}{=} \frac{\mu_B}{\hbar}\mathbf{L}\mathbf{B} \tag{7.11}$$

If there is an inhomogeneous magnetic field, there is also a force effect $\mathbf{F}_\mu$ on the dipole moment:

$$\mathbf{F}_\mu = -\boldsymbol{\nabla}E_\mu \overset{(7.11)}{=} -\frac{\mu_B}{\hbar}\boldsymbol{\nabla}(\mathbf{L}\mathbf{B}) \tag{7.12}$$

For simplicity, we choose our coordinate system so that our magnetic field only points in the z-direction: $\mathbf{B} = B\mathbf{e}_z$. The potential energy E_μ from (7.11) simplifies so that we can neglect the x- and y-directions of the angular momentum (and hence in a following quantum mechanical description the angular momentum operators $\hat{L}_x$ and $\hat{L}_y$):

$$E_{\mu,z} = \frac{\mu_B B}{\hbar}L_z \tag{7.13}$$

Through the correspondence principle, we can define from (7.13) an energy operator $\hat{H}_{B,z}$ for an electron in the directed magnetic field (where we also take the Landé factor g_L into account). At the same time, we will write $\hat{H}_{B,z}$ in spherical coordinates – we already know $\hat{L}_z$ in these coordinates from (5.108):

$$\hat{H}_{B,z} = \frac{\mu_B g_\mathrm{L} B}{\hbar}\hat{L}_z \overset{(5.108)}{=} -i\mu_B g_\mathrm{L} B\frac{\partial}{\partial\varphi} \tag{7.14}$$

The eigenfunctions of $\hat{H}_{B,z}$ are in the position representation the spherical harmonics. Starting from an abstract eigenstate of the angular momentum operator $|l,m\rangle$, according to (5.60), the eigenvalues can be easily calculated to be:

$$\hat{H}_{B,z}\,|l,m\rangle = \frac{\mu_\mathrm{B} B}{\hbar}g_\mathrm{L}\hat{L}_z\,|l,m\rangle \overset{(5.60)}{=} \mu_\mathrm{B} B g_\mathrm{L} m\,|l,m\rangle \tag{7.15}$$

If an external magnetic field is present, the states, which are otherwise degenerate, characterized by the magnetic quantum number m, can split! This is referred to as the *normal Zeeman effect*.

7.2 Stern-Gerlach Device

Let's consider the setup of a Stern-Gerlach device as shown in Figure 39. A beam of silver atoms with a well-defined speed is passed through a strongly inhomogeneous magnetic field $\mathbf{B}$ with an orientation $\mathbf{n}$ given by the pole shoes of the magnet.

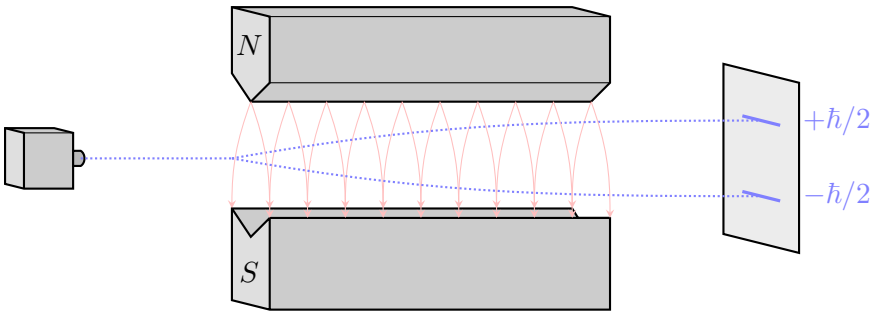


Fig. 39: Sketch of the Stern-Gerlach experiment with the double splitting of the particle beam.

Analogous to (7.12), the following force $\mathbf{F}$ acts on each particle:

$$\mathbf{F} = -\nabla\left(-\boldsymbol{\mu}\cdot\mathbf{B}\right) = \nabla\left(\boldsymbol{\mu}\cdot\mathbf{B}\right) \tag{7.16}$$

$\boldsymbol{\mu}$ represents the magnetic moment. It is important to understand that this deflecting force is *not* the Lorentz force, but arises from the magnetic moment. Therefore, it is also important to use neutral particles for the experiment, as the Lorentz force would otherwise completely overlay the result. Assuming the orientation of the magnetic field as the z-direction, expression (7.16) simplifies to:

$$\mathbf{F} = \left(\mu_z \frac{\partial B}{\partial z}\right)\mathbf{e}_z \tag{7.17}$$

For a finite force on the particles, it is essential for the experiment that the magnetic field indeed has a finite gradient $\frac{\partial B}{\partial z} \neq 0$. However, if the magnetic field is aligned along any direction $\mathbf{n}$, expression (7.16) can also be written as:

$$\mathbf{F} = \left(|\boldsymbol{\mu}|\frac{\partial B}{\partial n}\cos(\alpha)\right)\mathbf{n} \tag{7.18}$$

Here, $\frac{\partial B}{\partial n}$ should be the derivative of $B \equiv |\mathbf{B}|$ in the direction of $\mathbf{n}$, and α represents the angle between $\boldsymbol{\mu}$ and $\mathbf{B}$.

In the Stern-Gerlach experiment, a splitting of the beam of silver atoms into two sub-beams was observed. When attempting to interpret the splitting classically, one might conclude that the particles' angular momentum is not statistically distributed, but that only two groups of particles are generated in the particle source: those with angular momentum parallel to the magnetic field and those with angular momentum antiparallel to the magnetic field. However, the problem with this interpretation is that the measurement result is completely independent of the orientation of the magnetic field. No matter how the magnetic field is oriented, one always observes the division into two sub-beams! From a classical perspective, this makes no sense, and Stern and Gerlach were confused by the result of their experiment. They hoped to demonstrate the (observed) quantization of the magnetic moment, but an orbital angular momentum with an integer angular momentum quantum number $l = 1$ should have resulted in three sub-beams according to the triple magnetic degeneracy $p(1) = (2 \cdot 1 + 1) = 3$ according to (5.64).
A possible solution to this problem offers itself when we assume *half-integer* angular momenta.

7.3 Postulate of Spin

In the chapter on angular momentum, we decided to rename the angular momentum quantum number j from (5.56) to l for the orbital angular momentum, and to allow only positive integer values. However, for the general angular momentum quantum number j, it is also permissible to choose positive, *half-integer* values. Such half-integer angular momentum quantum numbers we will now assign to "spin", where we want to focus at this point on the spin of a single electron. Instead of an integer angular momentum j, for spin $j = s = \frac{1}{2}$ is chosen, in accordance with the double splitting of the atomic beam in the Stern-Gerlach experiment. The associated operators are labeled $\hat{\mathbf{S}}^2$ and $\hat{S}_z$. They function analogous to the orbital angular momentum operators $\hat{\mathbf{L}}^2$ and $\hat{L}_z$, and follow the same eigenvalue relationships with the eigenfunctions $|s, m_s\rangle$.

To actually explain the two lines in the Stern-Gerlach experiment, a magnetic (spin) moment $\hat{\boldsymbol{\mu}}_s$ must also exist, analogous to the orbital angular momentum. Depending on the actual value of the spin (here $s = \frac{1}{2}$), we obtain the following number of realization possibilities for the magnetic component m_s:

$$2s + 1 = 2$$

In general, for the magnetic component m_s of spin s, it must lie in the value range $m_s = -s, -s+1, \ldots, s-1, s$. In the case of a single spin with $s = \frac{1}{2}$, m_s can therefore only take two values, which we can write as:

$$m_s = \pm\frac{1}{2} \tag{7.19}$$

A clear interpretation of the spin is difficult; it is suggested to understand the electron spin as a kind of "intrinsic angular momentum", but upon closer inspection, this classical interpretation fails. This is because the speed v with which the "equator" of an electron with classical electron radius would have to rotate would exceed the speed of light ($v \gg c$).

Spin has another astonishing peculiarity: Only after a rotation by $\varphi = 4\pi$ does a wave function return to itself again. For orbital angular momentum, we find $\varphi = 2\pi$, which corresponds to classical expectations. It is as if an electron "has to rotate twice around its own axis" before it "looks the same again". This can be explained by the magnetic quantum number m_s of spin, which is half-integer and thus lets the phase of a wave function rotate only half as fast:

$$\exp(im_s\varphi) = \exp\left(i\frac{\varphi}{2}\right) \stackrel{!}{=} \exp\left[i\left(\frac{\varphi}{2} + 2\pi\right)\right] = \exp\left[i\frac{(\varphi + 4\pi)}{2}\right]$$

A coherent description of spin becomes possible only through the solution of the Dirac equation, which is a relativistic extension of the Schrödinger equation. Instead of postulating the spin in the non-relativistic case and introducing it artificially, the solution of the Dirac equation naturally provides both the spin and antiparticles.

7.4 Spin Eigen System

In the following chapter, we will deal with a pure spin system $s = \frac{1}{2}$ (i.e., there is no orbital angular momentum $\hat{\mathbf{L}}$) to learn about the relevant eigenvalue equations, operators, and representations. Let's begin with a consideration of the magnetic moment of spin $\hat{\boldsymbol{\mu}}_s$:

$$\hat{\boldsymbol{\mu}}_s = \mathrm{sgn}(q)\frac{\mu_B}{\hbar} g_s \hat{\mathbf{S}} \tag{7.20}$$

Here, $\mathrm{sgn}(q)$ is the sign of the charge q of the considered particle; in the case of the electron, $\mathrm{sgn}(q) = \mathrm{sgn}(-e) = -1$. For the electron, the Landé factor g_s no longer takes the trivial value 1; it specifically holds for electron spin:

$$g_s = 2.002319... \approx 2 \tag{7.21}$$

The approximately correct factor $g_s = 2$ can be derived from the solution of the Dirac equation for fermions. A consideration within the framework of quantum field theory allows for an even clearer picture and the determination of the decimal places.

7.4.1 Spin Operators and Commutators

As an angular momentum operator, the spin operator $\hat{\mathbf{S}}$ satisfies the same commutator relations as the orbital angular momentum $\hat{L}$ and is also a generator of rotations. When calculating the commutator between two spin operators, the following relation is always fulfilled (instead of $\hat{S}_j$, any vector operator $\hat{V}_j$ could be used here):

$$[\hat{S}_i, \hat{S}_j] = i\hbar\varepsilon_{ijk}S_k \tag{7.22}$$

We can simultaneously measure two spin-related quantum numbers exactly: the observable $\hat{\mathbf{S}}^2$ and a single component of spin $\hat{S}_i$, where we here, as usual, choose the z-direction. $\hat{\mathbf{S}}^2$ and $\hat{S}_z$ are compatible and can therefore be measured simultaneously:

$$[\hat{\mathbf{S}}^2, \hat{S}_z] = 0 \tag{7.23}$$

Relation (7.23) again allows us to find a common eigen system between the magnitude of the spin operator $\hat{\mathbf{S}}^2$ and an arbitrary spin component $\hat{S}_z$. (7.22) also forbids the possibility of simultaneously measuring different components of spin, so that our choice of the z-component leads to the fact that we cannot simultaneously measure the x- and y-components exactly.

We now know which observables we can determine simultaneously. But what is the action of the associated operators? Since we only renamed the general angular momentum quantum number j as s in the case of spin (instead of, as with the orbital angular momentum, to l), we already know the answer: The already known relations (5.59), (5.60), and (5.61) apply, only we replace the orbital angular momentum quantum number l with the spin quantum number s, and the magnetic angular momentum quantum number m with the magnetic spin quantum number m_s. The magnitude operator of spin $\hat{\mathbf{S}}^2$ thus acts on a spin eigenstate $|s, m_s\rangle$ in the following way:

$$\hat{\mathbf{S}}^2 |s, m_s\rangle = \hbar^2 s(s+1) |s, m_s\rangle \tag{7.24}$$

The z-component of spin leads us to an eigenvalue, which depends on the magnetic spin quantum number. It holds:

$$\hat{S}_z |s, m_s\rangle = \hbar m_s |s, m_s\rangle \tag{7.25}$$

We also find variations equivalent to the angular momentum for raising and lowering operators $\hat{S}_\pm$:

$$\hat{S}_\pm |s, m_s\rangle = \hbar\sqrt{s(s+1) - m_s(m_s \pm 1)}\, |s, m_s \pm 1\rangle \tag{7.26}$$

For the spin of a single electron, we already know $s = \frac{1}{2}$ or $m_s = \pm\frac{1}{2}$. We can thus rewrite the eigenstate anew, using the following notation variants optionally:

$$|s, m_s\rangle = \begin{cases} |\frac{1}{2}, +\frac{1}{2}\rangle = |\uparrow\rangle = |+\rangle = |0\rangle \\ |\frac{1}{2}, -\frac{1}{2}\rangle = |\downarrow\rangle = |-\rangle = |1\rangle \end{cases} \tag{7.27}$$

We thus obtain a clearer picture for (7.24) and (7.25). For a single electron, the eigenvalue of the magnitude operator $\hat{\mathbf{S}}^2$ is always the same, independent of the spin direction:

$$\begin{aligned}\hat{\mathbf{S}}^2 |\uparrow\rangle &\overset{(7.27)}{=} \hat{\mathbf{S}}^2 |\tfrac{1}{2}, +\tfrac{1}{2}\rangle \overset{(7.24)}{=} \hbar\tfrac{1}{2}\left(\tfrac{1}{2}+1\right)|\tfrac{1}{2}, +\tfrac{1}{2}\rangle \overset{(7.27)}{=} \tfrac{3}{4}\hbar^2 |\uparrow\rangle \\ \hat{\mathbf{S}}^2 |\downarrow\rangle &\overset{(7.27)}{=} \hat{\mathbf{S}}^2 |\tfrac{1}{2}, -\tfrac{1}{2}\rangle \overset{(7.24)}{=} \hbar\tfrac{1}{2}\left(\tfrac{1}{2}+1\right)|\tfrac{1}{2}, -\tfrac{1}{2}\rangle \overset{(7.27)}{=} \tfrac{3}{4}\hbar^2 |\downarrow\rangle\end{aligned} \tag{7.28}$$

The eigenvalue of $\hat{S}_z$ allows for a bit more variation; depending on the spin orientation, it results in a positive or negative sign:

$$\begin{aligned}\hat{S}_z |\uparrow\rangle &\overset{(7.27)}{=} \hat{S}_z |\tfrac{1}{2}, +\tfrac{1}{2}\rangle \overset{(7.26)}{=} +\tfrac{1}{2}\hbar |\tfrac{1}{2}, +\tfrac{1}{2}\rangle \overset{(7.27)}{=} +\tfrac{1}{2}\hbar |\uparrow\rangle \\ \hat{S}_z |\downarrow\rangle &\overset{(7.27)}{=} \hat{S}_z |\tfrac{1}{2}, -\tfrac{1}{2}\rangle \overset{(7.26)}{=} -\tfrac{1}{2}\hbar |\tfrac{1}{2}, -\tfrac{1}{2}\rangle \overset{(7.27)}{=} -\tfrac{1}{2}\hbar |\downarrow\rangle\end{aligned} \tag{7.29}$$

The ladder operators also behave as expected, with $\hat{S}_+ |\uparrow\rangle = \hat{S}_- |\downarrow\rangle = 0$ trivially disappearing:

$$\begin{aligned}\hat{S}_+ |\uparrow\rangle &\overset{(7.27)}{=} \hat{S}_+ |\tfrac{1}{2}, +\tfrac{1}{2}\rangle \overset{(7.26)}{=} \hbar\sqrt{\tfrac{1}{2}\left(\tfrac{1}{2}+1\right) - \tfrac{1}{2}\left(\tfrac{1}{2}+1\right)}\,|\tfrac{1}{2}, +\tfrac{3}{2}\rangle = 0 \\ \hat{S}_+ |\downarrow\rangle &\overset{(7.27)}{=} \hat{S}_+ |\tfrac{1}{2}, -\tfrac{1}{2}\rangle \overset{(7.26)}{=} \hbar |\uparrow\rangle \\ \hat{S}_- |\uparrow\rangle &\overset{(7.27)}{=} \hat{S}_- |\tfrac{1}{2}, +\tfrac{1}{2}\rangle \overset{(7.26)}{=} \hbar\sqrt{\tfrac{1}{2}\left(\tfrac{1}{2}+1\right) - \tfrac{1}{2}\left(\tfrac{1}{2}-1\right)}\,|\tfrac{1}{2}, -\tfrac{1}{2}\rangle \overset{(7.27)}{=} \hbar |\downarrow\rangle \\ \hat{S}_- |\downarrow\rangle &\overset{(7.27)}{=} \hat{S}_- |\tfrac{1}{2}, -\tfrac{1}{2}\rangle \overset{(7.26)}{=} \hbar\sqrt{\tfrac{1}{2}\left(\tfrac{1}{2}+1\right) + \tfrac{1}{2}\left(-\tfrac{1}{2}-1\right)}\,|\tfrac{1}{2}, -\tfrac{3}{2}\rangle = 0\end{aligned} \tag{7.30}$$

7.4.2 Matrix Representation and Pauli Matrices

So far, all our wave functions could be represented in the position space $\mathcal{H}_r$. The representation of the newly added spin eigenstates occurs exclusively in a spin Hilbert space $\mathcal{H}_s$. To describe all the characteristics of an electron, we introduce a total Hilbert space $\mathcal{H}$, which consists of a position component and a spin component:

$$\mathcal{H} = \mathcal{H}_r \otimes \mathcal{H}_s \quad \text{with} \quad \dim \mathcal{H} = 2 \cdot \dim \mathcal{H}_r \tag{7.31}$$

The fact that we consider spin at all when solving the Schrödinger equation is an *ad hoc* assumption, as only the Dirac equation naturally provides spin eigenstates (so-called *spinors*). With a single electron, only two spin states ($|+\rangle$ and $|-\rangle$), or their superposition, are realized. According to (7.31), the corresponding spin Hilbert space is therefore two-dimensional.

Let $|\psi\rangle_s \in \mathcal{H}_s$ be an eigenstate of the operators $\hat{\mathbf{S}}^2$ and $\hat{S}_z$, then orthogonality holds for the individual eigenstates:

$$\langle\uparrow|\uparrow\rangle = \langle\downarrow|\downarrow\rangle = 1 \quad \text{and} \quad \langle\downarrow|\uparrow\rangle = \langle\uparrow|\downarrow\rangle = 0 \tag{7.32}$$

The eigen system must also be complete; for our two-dimensional Hilbert space, the simple relationship holds:

$$\mathbb{1} = |\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow| \tag{7.33}$$

We can represent any spin state as a superposition of $|\uparrow\rangle$ and $|\downarrow\rangle$, where the coefficients $\alpha, \beta \in \mathbb{C}$ satisfy:

$$|\psi\rangle_s = \alpha|\uparrow\rangle + \beta|\downarrow\rangle \tag{7.34}$$

α and β must satisfy the relationship: $|\alpha|^2 + |\beta|^2 = 1$. This can be quickly shown by calculating the inner product:

$$\begin{aligned}\langle\psi|\psi\rangle_s &= \left(\alpha^*\langle\uparrow| + \beta^*\langle\downarrow|\right)\left(\alpha|\uparrow\rangle + \beta|\downarrow\rangle\right) = \\ &= \alpha^*\alpha\langle\uparrow|\uparrow\rangle + \beta^*\beta\langle\downarrow|\downarrow\rangle = |\alpha|^2 + |\beta|^2 = 1 \quad \square\end{aligned}$$

In analogy to $\mathbb{R}^2$ as a two-dimensional vector space, we can also represent the spin eigenstates $|+\rangle$ and $|-\rangle$ as vectors (where $|\phi_i\rangle \in \{|+\rangle, |-\rangle\}$):

$$|\uparrow\rangle \xrightarrow{\{|\phi_i\rangle\}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |\downarrow\rangle \xrightarrow{\{|\phi_i\rangle\}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{7.35}$$

We could also define the vectorial representation of $|\uparrow\rangle$ and $|\downarrow\rangle$ the other way around, but the notation must remain consistent. The general spin state from (7.34) is given in this vector notation as:

$$\psi_s = \begin{pmatrix} \langle\uparrow|\psi\rangle_s \\ \langle\downarrow|\psi\rangle_s \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{7.36}$$

In the eigen space of the operators $\hat{\mathbf{S}}^2$ and $\hat{S}_z$, the eigenstates can be represented as vectors; a spectral decomposition allows representing the operators as matrices as well:

$$\hat{\mathbf{S}}^2 = \hat{\mathbf{S}}^2 \mathbb{1} = \hat{\mathbf{S}}^2 \left(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|\right) \overset{(7.28)}{=} \frac{3\hbar^2}{4} \left(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|\right) \xrightarrow{\{|\phi_i\rangle\}}$$
$$\xrightarrow{\{|\phi_i\rangle\}} \frac{3\hbar^2}{4} \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} \right] = \frac{3\hbar^2}{4} \left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right] = \frac{3\hbar^2}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

In the spin eigenbasis, we can thus represent $\hat{\mathbf{S}}^2$ as a matrix of the following form:

$$\mathbf{S}^2 = \frac{3\hbar^2}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{7.37}$$

Analogously, using (7.29) for the z-component of the spin operator $\hat{S}_z$:

$$\hat{S}_z = \hat{S}_z \mathbb{1} = \hat{S}_z \left(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|\right) \overset{(7.28)}{=} \frac{\hbar}{2} \left(|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|\right) \xrightarrow{\{|\phi_i\rangle\}}$$
$$\xrightarrow{\{|\phi_i\rangle\}} \frac{\hbar}{2} \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} \right] = \frac{\hbar}{2} \left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right] = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

We can give the matrix representation of $\hat{S}_z$ in the spin eigenbasis and find the corresponding representation as a matrix as:

$$S_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{7.38}$$

As expected, $\hat{\mathbf{S}}^2$ and $\hat{S}_z$ are diagonal in matrix representation – since the description takes place in the common eigen system of the two operators, we can identify the diagonal elements as eigenvalues.

Example: Matrix Representation of $\hat{S}_x$ and $\hat{S}_y$

The representation of $\hat{S}_x$ and $\hat{S}_y$ as matrices is somewhat trickier compared to the z-component, as we are *not* in the spin eigen system of the two operators. Expressing $\hat{S}_x$ and $\hat{S}_y$ in analogy to (5.67) and (5.68) via $\hat{S}_+$ and $\hat{S}_-$, we have:

$$\hat{S}_x = \frac{1}{2}\left(\hat{S}_+ + \hat{S}_-\right) \quad \text{and} \quad \hat{S}_y = \frac{1}{2i}\left(\hat{S}_+ - \hat{S}_-\right) \tag{7.39}$$

Let's first consider what the matrix representations of the ladder operators $\hat{S}_+$ and $\hat{S}_-$

look like. As in (7.37) and (7.38), we introduce a complete identity $\mathbb{1}$ and write

$$\hat{S}_+ = \hat{S}_+\mathbb{1} = \hat{S}_+ |\uparrow\rangle \langle\uparrow| + \hat{S}_+ |\downarrow\rangle \langle\downarrow| \overset{(7.30)}{=} \hbar |\uparrow\rangle \langle\downarrow| \xrightarrow{\{|\phi_i\rangle\}} \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$\hat{S}_- = \hat{S}_-\mathbb{1} = \hat{S}_- |\uparrow\rangle \langle\uparrow| + \hat{S}_- |\downarrow\rangle \langle\downarrow| \overset{(7.30)}{=} \hbar |\downarrow\rangle \langle\uparrow| \xrightarrow{\{|\phi_i\rangle\}} \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

We have utilized the action of the ladder operators on a multiplet state: The highest state $|\uparrow\rangle$ cannot be increased further, just as the lowest state $|\downarrow\rangle$ cannot be decreased further. $\hat{S}_x$ and $\hat{S}_y$ result from the appropriate summation of the ladder operators according to the relations (7.39):

$$\hat{S}_x \overset{(7.39)}{=} \frac{1}{2}\left(\hat{S}_+ + \hat{S}_-\right) \xrightarrow{\{|\phi_i\rangle\}} \frac{\hbar}{2}\left[\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\right] = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{7.40}$$

$$\hat{S}_y \overset{(7.39)}{=} \frac{1}{2i}\left(\hat{S}_+ - \hat{S}_-\right) \xrightarrow{\{|\phi_i\rangle\}} \frac{\hbar}{2i}\left[\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\right] = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{7.41}$$

These matrices are *not* diagonal. Letting $\hat{S}_x$ and $\hat{S}_y$ act on an eigenstate $|\uparrow\rangle$ or $|\downarrow\rangle$, the eigenstate changes. For example, for $\hat{S}_x |\downarrow\rangle = (\hbar/2) |\uparrow\rangle$, we obtain:

$$\hat{S}_x |\downarrow\rangle = \frac{\hbar}{2} |\uparrow\rangle \xrightarrow{\{|\phi_i\rangle\}} \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Pauli Matrices $\hat{S}_x$, $\hat{S}_y$, and $\hat{S}_z$ are usually expressed with the *Pauli matrices* σ_x, σ_y, and σ_z (often called σ_1, σ_2, and σ_3 in literature), which have the following form:

$$\sigma_x \equiv \sigma_1 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \sigma_y \equiv \sigma_2 \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{and} \quad \sigma_z \equiv \sigma_3 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{7.42}$$

It is clear from (7.38), (7.40), and (7.41) that the components of the spin operator $\hat{\mathbf{S}}$ and the Pauli matrices differ only by a constant factor $\hbar/2$, which corresponds to the eigenvalue of $\hat{S}_z$:

$$\hat{S}_i = \frac{\hbar}{2}\sigma_i \tag{7.43}$$

Since the difference is only in a multiplicative factor, all eigenvalue relations and commutators for the Pauli matrices function the same as for $\hat{S}_x$, $\hat{S}_y$, and $\hat{S}_z$, differing only by the factor $\hbar/2$. For the commutator from (7.22), we find the equivalent relation:

$$[\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k \tag{7.44}$$

The correctness of the commutator (7.44) can be easily verified by substituting (7.43) into (7.22). Furthermore, for the anticommutator (in relativistic quantum theory, this relationship is a determining part of the Clifford algebra), the following holds:

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij} \tag{7.45}$$

(7.45) can again be verified by substituting the actual Pauli matrices. We can use (7.44) and (7.45) to find a generally valid relationship for arbitrary vector operators $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$:

$$(\boldsymbol{\sigma}\hat{\mathbf{A}})(\boldsymbol{\sigma}\hat{\mathbf{B}}) = \hat{\mathbf{A}}\hat{\mathbf{B}} + i\boldsymbol{\sigma}(\hat{\mathbf{A}} \times \hat{\mathbf{B}}) \tag{7.46}$$

Example: Product of Pauli Matrices and Vector Operators

Let's start with the anticommutator relation $\{\sigma_i, \sigma_j\} = 2\delta_{ij}$ and manipulate it. We find:

$$\{\sigma_i, \sigma_j\} = \sigma_i\sigma_j + \sigma_j\sigma_i = \sigma_i\sigma_j - \sigma_j\sigma_i + \sigma_j\sigma_i + \sigma_j\sigma_i = [\sigma_i, \sigma_j] + 2\sigma_j\sigma_i$$

Rearranging this equation allows us to represent the product of two Pauli matrices through the commutator and the anticommutator:

$$\sigma_j\sigma_i = \frac{1}{2}\left(\{\sigma_i, \sigma_j\} + [\sigma_i, \sigma_j]\right) \tag{7.47}$$

We can use (7.47) to derive $(\boldsymbol{\sigma}\hat{\mathbf{A}})(\boldsymbol{\sigma}\hat{\mathbf{B}}) = \hat{\mathbf{A}}\hat{\mathbf{B}} + i\boldsymbol{\sigma}(\hat{\mathbf{A}} \times \hat{\mathbf{B}})$ from (7.46). $\hat{A}$ and $\hat{B}$ should be arbitrary vector operators. To facilitate this derivation, we switch to index notation:

$$\begin{aligned}
(\boldsymbol{\sigma}\hat{\mathbf{A}})(\boldsymbol{\sigma}\hat{\mathbf{B}}) &= \sigma_i\sigma_j\hat{A}_i\hat{B}_j \overset{(7.47)}{=} \\
&= \frac{1}{2}\left(\{\sigma_i, \sigma_j\} + [\sigma_i, \sigma_j]\right)\hat{A}_i\hat{B}_j \overset{(7.44,7.45)}{=} \\
&= \frac{1}{2}\left(2\delta_{ij} + 2i\varepsilon_{ijk}\sigma_k\right)\hat{A}_i\hat{B}_j = \\
&= \delta_{ij}\hat{A}_i\hat{B}_j + i\varepsilon_{ijk}\sigma_k\hat{A}_i\hat{B}_j = \\
&= \hat{A}_i\hat{B}_i + i\sigma_k\varepsilon_{ijk}\hat{A}_i\hat{B}_j = \\
&= (\boldsymbol{\sigma}\hat{\mathbf{A}})(\boldsymbol{\sigma}\hat{\mathbf{B}}) + i\boldsymbol{\sigma}(\hat{\mathbf{A}} \times \hat{\mathbf{B}}) \quad \square
\end{aligned}$$

In the last step, it was used that $\varepsilon_{ijk}\hat{A}_i\hat{B}_j$ corresponds to the cross product of the quantities $\hat{A}$ and $\hat{B}$.

Properties of the Pauli Matrices All Pauli matrices must satisfy the following properties; that these statements are true can be easily confirmed by explicitly inserting the Pauli matrices from (7.42):

- **Hermiticity**: Each of the Pauli matrices is hermitian, thus:

$$\sigma_i^\dagger = \sigma_i \tag{7.48}$$

- **Unitarity**: Likewise, all Pauli matrices are unitary, hence:

$$\sigma_i^{-1} = \sigma_i^\dagger \tag{7.49}$$

- **Determinant**: The determinant of any Pauli matrix is always:

$$\det\{\sigma_i\} = -1 \tag{7.50}$$

- **Trace**: The trace of any Pauli matrix vanishes:

$$\mathrm{Tr}\{\sigma_i\} = 0 \tag{7.51}$$

- **Magnitude squared**: The magnitude squared of each component of the vector of Pauli matrices σ_i results in unity (which follows from the hermiticity and unitarity of σ_i):

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \mathbb{1} \tag{7.52}$$

7.4.3 Bloch Sphere

So far, to describe spin, we have always assumed $\hat{\mathbf{S}}^2$ and $\hat{S}_z$ to be the sharply measurable observables, while $\hat{S}_x$ and $\hat{S}_y$ are "smeared". Let us now abandon this simplification and assume spin states oriented parallel or antiparallel in the spin space to an arbitrarily oriented unit vector $\mathbf{n} = (n_x, n_y, n_z)^\intercal$. We seek a spin operator $\hat{\mathbf{S}}_\mathbf{n}$, for which such states are eigenstates. A description in spherically symmetric coordinates (for a unit sphere $r = 1$) is suitable, where for the spin operator in the direction $\mathbf{n}$ the following applies: $\hat{\mathbf{S}}_\mathbf{n} \equiv \hat{\mathbf{S}} \cdot \mathbf{n}$:

$$\begin{aligned}\hat{\mathbf{S}}_\mathbf{n} = \hat{\mathbf{S}} \cdot \mathbf{n} = \frac{\hbar}{2}\boldsymbol{\sigma} \cdot \mathbf{n} = \frac{\hbar}{2}\begin{pmatrix}\boldsymbol{\sigma}_x\\ \boldsymbol{\sigma}_y\\ \boldsymbol{\sigma}_z\end{pmatrix} \cdot \begin{pmatrix}n_x\\ n_y\\ n_z\end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix}\boldsymbol{\sigma}_x\\ \boldsymbol{\sigma}_y\\ \boldsymbol{\sigma}_z\end{pmatrix} \cdot \begin{pmatrix}\sin(\vartheta)\cos(\varphi)\\ \sin(\vartheta)\sin(\varphi)\\ \cos(\vartheta)\end{pmatrix} = \\ = \frac{\hbar}{2}\left[\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\sin(\vartheta)\cos(\varphi) + \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\sin(\vartheta)\sin(\varphi) + \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\cos(\vartheta)\right]\end{aligned} \tag{7.53}$$

The expression from (7.53) can be simplified using Euler's formula $e^{\pm i\varphi} = \cos(\varphi) \pm i\sin(\varphi)$. Therefore, we find the following expression for $\hat{\mathbf{S}}_\mathbf{n}$:

$$\hat{\mathbf{S}}_\mathbf{n} \equiv \hat{\mathbf{S}}_\mathbf{n}(\vartheta, \varphi) = \frac{\hbar}{2}\begin{pmatrix}\cos(\vartheta) & \sin(\vartheta)e^{-i\varphi}\\ \sin(\vartheta)e^{+i\varphi} & -\cos(\vartheta)\end{pmatrix} \tag{7.54}$$

Thus, we have found a compact expression for the spin operator $\hat{\mathbf{S}}_\mathbf{n}$ for an arbitrary orientation $\mathbf{n}$ of the spin. However, an eigenstate of $\hat{\mathbf{S}}_\mathbf{n}$ can always be expressed in the eigenbasis of $\hat{S}_z$! If we solve the eigenvalue problem with respect to $\hat{\mathbf{S}}_\mathbf{n}$, we find the eigenvalues and eigenstates (7.55) with the help of the following calculation.

The eigenvalues of $\hat{\mathbf{S}}_\mathbf{n}$ are calculated using $\det(\hat{\mathbf{S}}_\mathbf{n} - \mathbb{1}\lambda) = 0$, which yields the following characteristic polynomial:

$$0 = -\left(\frac{\hbar}{2}\cos(\vartheta) + \lambda\right)\left(\frac{\hbar}{2}\cos(\vartheta) - \lambda\right) - \frac{\hbar^2}{4}\sin^2(\vartheta) = \lambda^2 - \frac{\hbar^2}{4}$$

We find two eigenvalues $\lambda_1 = +\hbar/2$ and $\lambda_2 = -\hbar/2$, just like with the already known operator $\hat{S}_z$. To calculate the corresponding eigenstates, we use the two trigonometric relations $\cos^2(\vartheta/2) = (1 + \cos(\vartheta))/2$ and $\sin^2(\vartheta/2) = (1 - \cos(\vartheta))/2$ as well as $\sin(\vartheta) = 2\sin(\vartheta/2)\cos(\vartheta/2)$. It follows:

$$\begin{aligned}\lambda_1 &: \left(\begin{array}{cc|c}\cos(\vartheta) - 1 & \sin(\vartheta)e^{-i\varphi} & 0\\ \sin(\vartheta)e^{+i\varphi} & -\cos(\vartheta) - 1 & 0\end{array}\right) = \left(\begin{array}{cc|c}-\sin^2(\vartheta/2) & \sin(\vartheta/2)\cos(\vartheta/2)e^{-i\varphi} & 0\\ \sin(\vartheta/2)\cos(\vartheta/2)e^{+i\varphi} & -\cos^2(\vartheta/2) & 0\end{array}\right)\\ \lambda_2 &: \left(\begin{array}{cc|c}\cos(\vartheta) + 1 & \sin(\vartheta)e^{-i\varphi} & 0\\ \sin(\vartheta)e^{+i\varphi} & -\cos(\vartheta) + 1 & 0\end{array}\right) = \left(\begin{array}{cc|c}\cos^2(\vartheta/2) & \sin(\vartheta/2)\cos(\vartheta/2)e^{-i\varphi} & 0\\ \sin(\vartheta/2)\cos(\vartheta/2)e^{+i\varphi} & \sin^2(\vartheta/2) & 0\end{array}\right)\end{aligned}$$

By solving these two systems of equations, we obtain for the new eigenstates $|\uparrow_\mathbf{n}\rangle$ and $|\downarrow_\mathbf{n}\rangle$, expressed in the eigenbasis $\{|\uparrow\rangle, |\downarrow\rangle\}$ of $\hat{S}_z$, the two general states:

$$\begin{aligned}|\uparrow_\mathbf{n}\rangle &= \begin{pmatrix}\cos(\vartheta/2)e^{-i\varphi}\\ \sin(\vartheta/2)\end{pmatrix} = \cos(\vartheta/2)e^{-i\varphi}|\uparrow\rangle + \sin(\vartheta/2)|\downarrow\rangle \quad \text{for} \quad \lambda_1 = +\frac{\hbar}{2}\\ |\downarrow_\mathbf{n}\rangle &= \begin{pmatrix}\sin(\vartheta/2)e^{-i\varphi}\\ -\cos(\vartheta/2)\end{pmatrix} = \sin(\vartheta/2)e^{-i\varphi}|\uparrow\rangle - \cos(\vartheta/2)|\downarrow\rangle \quad \text{for} \quad \lambda_1 = -\frac{\hbar}{2}\end{aligned} \tag{7.55}$$

A general spin state $|\uparrow_\mathbf{n}\rangle$ or $|\downarrow_\mathbf{n}\rangle$ is thus developed into the eigenstates $|\uparrow\rangle$ or $|\downarrow\rangle$ of $\hat{S}_z$, where the development coefficients describe the position on the surface of a unit sphere. This can be illustrated graphically using the *Bloch sphere* (also called *Poincaré sphere*), on which each spin

state $|\psi\rangle = |\uparrow_{\mathbf{n}}\rangle$ can be uniquely represented on the sphere surface by the angles ϑ and φ. At the same time, this representation indicates the spatial angles in which the corresponding spin state can be measured sharply.
The fact that each state in a two-dimensional complex vector space can be uniquely associated with a spatial angle ultimately follows from the isomorphism between the Lie algebras for the SU(2) and the SO(3). For spin quantum numbers higher than $s = \frac{1}{2}$, this unique correspondence no longer holds.

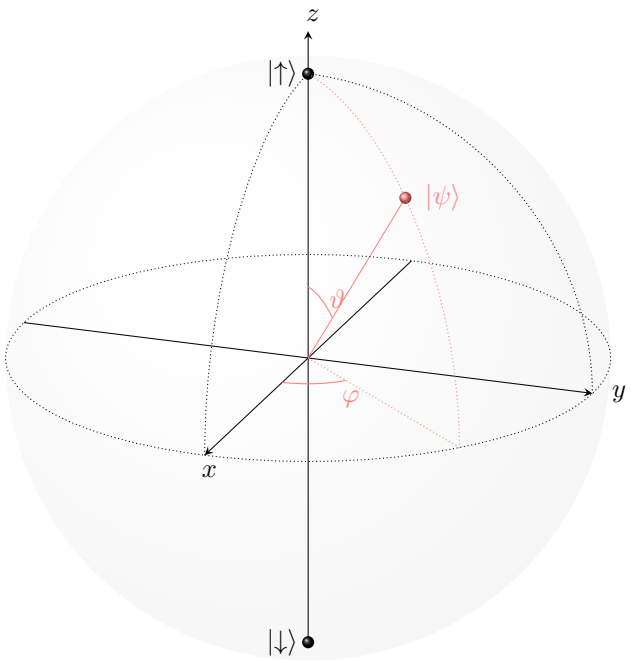


Fig. 40: Schematic representation of the Bloch sphere with an arbitrary state $|\psi\rangle = |\uparrow_{\mathbf{n}}\rangle$ with the spatial angles ϑ and φ.

Example: Rotation of Spin Eigenstates

To illustrate spin states in arbitrary directions, we choose the simple case for a fixed azimuth angle $\varphi = 0$ and a variable polar angle ϑ. Substituting into the finished formula for the spin eigenfunctions (7.55), we find for a general ϑ the state $|\uparrow_{\vartheta,0}\rangle$, respectively for $\vartheta = \pi/2$ the state $|\uparrow_x\rangle$:

$$|\uparrow_{\vartheta,0}\rangle = \cos(\vartheta/2)\,|\uparrow\rangle + \sin(\vartheta/2)\,|\downarrow\rangle \xrightarrow{\vartheta=\pi/2} |\uparrow_x\rangle = \frac{1}{\sqrt{2}}\,(|\uparrow\rangle + |\downarrow\rangle)$$

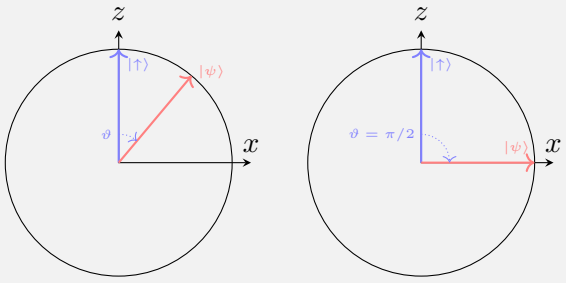


Fig. 41: Rotation of an eigenstate of the $\hat{S}_z$ operator by (left) the arbitrary angle ϑ and (right) by $\vartheta = \pi/2$. In both cases, $\varphi = 0$ applies.

With $\vartheta = \pi/2$, we have rotated the general spin state $|\uparrow_{\vartheta,0}\rangle$ directly to the x-axis. We can now calculate the probability with which we measure $+\hbar/2$ for the observable $\hat{S}_z$ for a system in the state $|\uparrow_{\vartheta,0}\rangle$ or $|\uparrow_x\rangle$:

$$\langle \hat{P}_{\uparrow\vartheta,0}\rangle = |\langle\uparrow|\uparrow_{\vartheta,0}\rangle|^2 = \cos^2(\vartheta/2) \xrightarrow{\vartheta=\pi/2} \langle \hat{P}_{\uparrow x}\rangle = \frac{1}{2}$$

In-Depth: Modern Interpretation of the Stern-Gerlach Experiment

When there is an interaction between the magnetic field in the Stern-Gerlach apparatus and the spin of the particle, its energy changes. The energy during such a process can be expressed using the Hamiltonian operator $\hat{H}$, which has the following form:

$$\hat{H} = -\hat{\boldsymbol{\mu}} \cdot \mathbf{B} \approx 2\frac{\mu_B}{\hbar} B\, \hat{\mathbf{S}} \cdot \mathbf{n} = \mu_B B\, \hat{\boldsymbol{\sigma}} \cdot \mathbf{n} = \alpha\, \hat{\boldsymbol{\sigma}} \cdot \mathbf{n} \tag{7.56}$$

For an arbitrary direction $\mathbf{n}$ of the magnetic field, we can write (7.56) as a matrix in the eigenbasis of $\hat{\sigma}_z$ analogous to (7.54), expressing $\mathbf{n}$ in spherically symmetric coordinates according to (5.83) (assuming a unit sphere with $r = 1$):

$$\hat{H} = \alpha \begin{pmatrix} \cos(\vartheta) & \sin(\vartheta)e^{-i\varphi} \\ \sin(\vartheta)e^{+i\varphi} & -\cos(\vartheta) \end{pmatrix} \tag{7.57}$$

Depending on the polar and azimuth angles ϑ and φ, we can manipulate the structure of the Hamiltonian operator. Three configurations of the two angles $\{\vartheta, \varphi\}$ should be highlighted: $\{\pi/2, 0\}$, $\{\pi/2, \pi/2\}$, and $\{0, 0\}$, which correspond to the x-, y-, and z-directions of the magnetic field $\mathbf{B}$ – it is evident that these configurations each generate one of the Pauli matrices.

All particles that are located in the "upper" sub-beam after the measurement are particles in the state $|\uparrow_{\mathbf{n}}\rangle$, and have an eigenvalue of $+\hbar/2$. So if a particle takes the upper path, and the Stern-Gerlach apparatus is interpreted as a measuring device, then one can say that the state $|\uparrow_{\mathbf{n}}\rangle$ has been measured on the particle. If one interprets it as a preparation device, one can also say that the particle has been prepared in the state $|\uparrow_{\mathbf{n}}\rangle$. Conversely, all particles in the "lower" sub-beam are in the state $|\downarrow_{\mathbf{n}}\rangle$, and have an eigenvalue of $-\hbar/2$. The Stern experiment is depicted in Figure 39.

Only with quantum physics and the concept of spin does the result of the Stern-Gerlach experiment make sense: The Stern-Gerlach apparatus is a measuring device that measures the spin of the particles in a certain direction (e.g., in the z-direction). If a particle has not been prepared before the measurement, then it does not have a well-defined spin orientation – in a very fundamental sense. It is not just that we do not know the orientation; it does not actually exist yet. Only through the measurement does the observed reality "arise" in a way, where the spin-$\frac{1}{2}$ particle is either aligned parallel or antiparallel to the measurement direction – and this regardless of which measurement direction was chosen.

However, *after* the measurement with the Stern-Gerlach apparatus, there are two sub-beams, and in each sub-beam, the particles have a well-defined spin orientation along the chosen axis. If one then measures along this axis again, the measurement result can be predicted with certainty. This is an example that in quantum mechanics every measurement process is also a preparation process.

7.5 Total Angular Momentum

Motivation: Total Angular Momentum

We now want to move on to the total angular momentum of a system that has more than one angular momentum as a degree of freedom. This can involve single particles with, for example, spin and orbital angular momentum, as well as the total angular momentum of multiple particles, which may have different intrinsic angular momenta/spins. To understand the behavior of such systems, it is necessary to perform the addition of the individual angular momentum operators and find a quantum mechanical representation of the new eigenstates.

A relevant example of total angular momentum is once again the hydrogen atom. The electron in the hydrogen atom has both an orbital angular momentum and a spin. In an external magnetic field, one observes a splitting of the emission and absorption lines – this splitting and the corresponding number of allowed transitions can only be described by a total angular momentum that takes into account the coupling of the individual angular momentum operators.

Another example is the helium atom, which consists of two electrons and the nucleus. In the case where the orbital angular momentum of both electrons is neglected or $l = 0$ is assumed, describing the total angular momentum of both electrons requires taking into account the coupling of the spins. In the coupled basis, one obtains a so-called singlet state with total spin $S = 0$ (parahelium) and a triplet state with $S = 1$ (orthohelium). These states can also be used for understanding the so-called quantum entanglement.

Let us now transition from a single, free electron to the hydrogen atom from the previous chapter, and try to find a more complete picture for the electron bound in this system using the knowledge about spin we have gained here. The electron bound in the central potential (without spin) can be characterized by the quantum numbers $\{n, l, m_l\}$; describing the spin state requires the additional quantum numbers $\{s, m_s\}$. From (7.31), we know that the orbital angular momentum and spin exist in different spaces $\mathcal{H}_r$ and $\mathcal{H}_s$ – a product state $|\psi\rangle \in \mathcal{H}$ is given accordingly by:

$$|\psi\rangle = |n, l, m_l\rangle \otimes |s, m_s\rangle \equiv |n, l, m_l; s, m_s\rangle \tag{7.58}$$

For the inner product of two normalized states $|\psi\rangle$ and $|\psi'\rangle$, orthogonality follows. It holds:

$$\langle\psi|\psi'\rangle = \langle n, l, m_l; s, m_s|n', l', m_l'; s', m_s'\rangle = \delta_{nn'}\delta_{ll'}\delta_{m_l m_l'}\delta_{ss'}\delta_{m_s m_s'} \tag{7.59}$$

For simplicity, we will always assume $n = n'$ in the following, as we are initially only interested in the relations between orbital angular momentum and spin. Both $\hat{\mathbf{L}}$ and $\hat{\mathbf{S}}$ are angular momentum operators and can be described analogously. The difference lies in the space in which these operators exist. The goal now is to transform $|\psi\rangle$ in a way that allows the description of a total angular momentum.

7.5.1 Product Basis and Coupled Basis

So far, in position space, we have found the following compatible operators: $\{\hat{H}, \hat{\mathbf{L}}^2, \hat{L}_z\}$. We are thus able to measure the energy, as well as the magnitude of the angular momentum and its z-component simultaneously with certainty. We identify this set of observables with the following "good" quantum numbers:

$$\{\hat{H}, \hat{\mathbf{L}}^2, \hat{L}_z\} \implies \{n, l, m_l\}$$

By adding the spin system, we now have two more operators available, which can also be measured simultaneously with certainty: $\{\hat{\mathbf{S}}^2, \hat{S}_z\}$. Again, we find a set of "good" quantum numbers that can be determined simultaneously:

$$\{\hat{\mathbf{S}}^2, \hat{S}_z\} \implies \{s, m_s\}$$

However, the question arises as to whether $\{\hat{H}, \hat{\mathbf{L}}^2, \hat{L}_z\}$ and $\{\hat{\mathbf{S}}^2, \hat{S}_z\}$ are compatible operators, making a simultaneous measurement of the respective observables possible. Since the operators of the position space $\hat{\mathbf{L}}^2$ and $\hat{L}_z$ cannot act on the spin part of a state and vice versa, we immediately recognize that a simultaneous measurement of these quantities will be possible. For $\hat{H}$, such a clear statement is not *a priori* possible, as the Hamiltonian operator can indeed contain terms that act on the spin (for example, in the hydrogen atom in a magnetic field). If this is the case, and there are terms that do not commute with $\hat{\mathbf{S}}^2$ and $\hat{S}_z$, a common eigenstate system cannot generally be found.

Product Basis Let's assume, for simplicity, that the Hamiltonian is entirely independent of spin, whereby $\{\hat{H}, \hat{\mathbf{L}}^2, \hat{L}_z, \hat{\mathbf{S}}^2, \hat{S}_z\}$ is a set of compatible observables, and a characterization of a wave function in the form (7.58) is possible. In this case, we speak of the *product basis* with the quantum numbers:

$$\{\hat{\mathbf{L}}^2, \hat{L}_z; \hat{\mathbf{S}}^2, \hat{S}_z\} \implies \{l, m_l; s, m_s\} \tag{7.60}$$

In this representation, $\hat{\mathbf{L}}$ and $\hat{\mathbf{S}}$ are independent of each other, and, as shown in Figure 42, a separate measurement of m_l and m_s is possible!

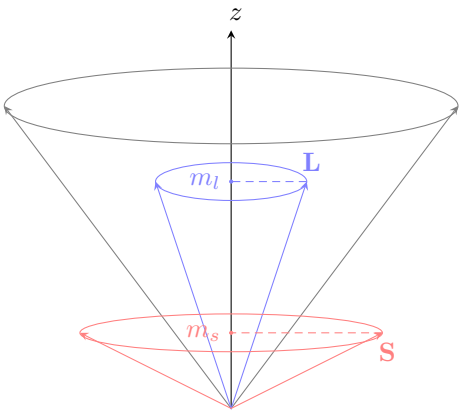


Fig. 42: Schematic representation of the product basis: orbital angular momentum $\hat{\mathbf{L}}$ and spin $\hat{\mathbf{S}}$ are independent of each other and a separate measurement of the projections in the z-direction is possible (m_l and m_s).

However, this set of compatible observables is not unique, as we can build another set of compatible observables using the *total angular momentum* $\hat{\mathbf{J}}$:

$$\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}} = \hat{\mathbf{L}} \otimes \mathbb{1}_s + \mathbb{1}_r \otimes \hat{\mathbf{S}} \tag{7.61}$$

$\hat{\mathbf{J}}$ thus acts both on eigenstates of the orbital angular momentum and on eigenstates of the spin, and generally still fulfills the function of an angular momentum operator. Again, in all generality, the already known commutation relation holds:

$$[\hat{J}_i, \hat{J}_j] = i\hbar\varepsilon_{ijk}\hat{J}_k \tag{7.62}$$

That this must be satisfied can be easily verified by using the already known commutators (5.23) and (7.22). We show:

$$\begin{aligned}
[\hat{J}_i, \hat{J}_j] &= [\hat{L}_i + \hat{S}_i, \hat{L}_j + \hat{S}_j] = \\
&= [\hat{L}_i, \hat{L}_j] + [\hat{L}_i, \hat{S}_j] + [\hat{S}_i, \hat{L}_j] + [\hat{S}_i, \hat{S}_j] \overset{(5.23,7.22)}{=} \\
&= i\hbar\varepsilon_{ijk}\hat{L}_k + i\hbar\varepsilon_{ijk}\hat{S}_k = \\
&= i\hbar\varepsilon_{ijk}(\hat{L}_k + \hat{S}_k) \overset{(7.61)}{=} \\
&= i\hbar\varepsilon_{ijk}\hat{J}_k \quad \square
\end{aligned} \tag{7.63}$$

Additionally, it holds that the operator for the magnitude of the total angular momentum operator $\hat{\mathbf{J}}^2$ and the operator for the z-component $\hat{J}_z$ are compatible, and thus can form a common eigenstate system:

$$[\hat{\mathbf{J}}^2, \hat{J}_z] = 0 \tag{7.64}$$

However, it is no longer clear in advance whether the set of compatible observables $\{\hat{\mathbf{J}}^2, \hat{J}_z\}$ can be added to the observables of the product basis (7.60). To address this, we ask whether $\hat{\mathbf{J}}^2$ and $\hat{J}_z$ commute with the observables $\{\hat{\mathbf{L}}^2, \hat{L}_z, \hat{\mathbf{S}}^2, \hat{S}_z\}$ – $\hat{H}$ has been omitted here as the Hamiltonian will have terms that are at most linear combinations of $\hat{\mathbf{L}}$ and $\hat{\mathbf{S}}$.

Coupled Basis For $\hat{J}_z = \hat{L}_z + \hat{S}_z$, it can be quickly ensured that this observable is compatible with all other observables $\{\hat{\mathbf{L}}^2, \hat{L}_z, \hat{\mathbf{S}}^2, \hat{S}_z\}$. But what about $\hat{\mathbf{J}}^2$? $\hat{\mathbf{J}}^2$ can be expressed as:

$$\hat{\mathbf{J}}^2 = (\hat{\mathbf{L}} + \hat{\mathbf{S}})^2 = \hat{\mathbf{L}}^2 + \hat{\mathbf{S}}^2 + 2\hat{\mathbf{L}}\hat{\mathbf{S}} = \hat{\mathbf{L}}^2 + \hat{\mathbf{S}}^2 + 2\left(\hat{L}_x\hat{S}_x + \hat{L}_y\hat{S}_y + \hat{L}_z\hat{S}_z\right) \tag{7.65}$$

The term $\hat{\mathbf{L}}^2 + \hat{\mathbf{S}}^2$ poses no problem, as it is again compatible with all other observables. However, the cross components $\hat{L}_x\hat{S}_x + \hat{L}_y\hat{S}_y$ do not commute with the observables $\{\hat{L}_z, \hat{S}_z\}$! We thus find *no* common eigenstate system of the form $\{\hat{\mathbf{L}}^2, \hat{L}_z, \hat{\mathbf{S}}^2, \hat{S}_z, \hat{\mathbf{J}}^2, \hat{J}_z\}$!

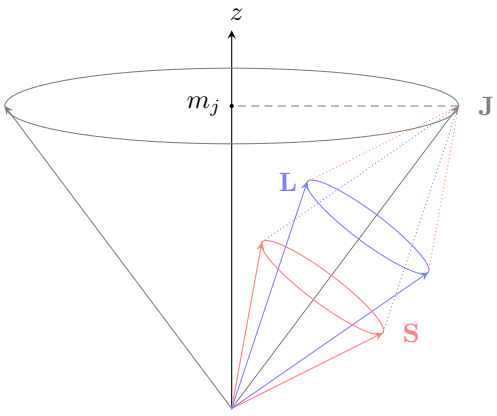


Fig. 43: Schematic representation of the total angular momentum $\hat{\mathbf{J}}$ and its components $\hat{\mathbf{L}}$ and $\hat{\mathbf{S}}$. It is evident that we cannot measure orbital angular momentum and spin sharply, as they precess around the total angular momentum.

Without proof, the following theorem holds: In a system with more than one set of simultaneously diagonalizable operators, the number of compatible and independent operators is a conserved quantity. For our case, this means that we can only find four mutually compatible observables. We choose the set:

$$\{\hat{\mathbf{L}}^2, \hat{\mathbf{S}}^2, \hat{\mathbf{J}}^2, \hat{J}_z\} \implies \{l, s, j, m_j\} \tag{7.66}$$

While $\hat{J}_z$ was already identified as compatible with the other operators, the problematic operators $\{\hat{L}_z, \hat{S}_z\}$ have now been removed. $\hat{\mathbf{J}}^2$ is compatible with $\hat{\mathbf{L}}^2$ and $\hat{\mathbf{S}}^2$, and of course, with $\hat{J}_z$. In the case of (7.66), we speak of the *coupled basis*.

In the coupled basis $\{\hat{\mathbf{L}}^2, \hat{\mathbf{S}}^2, \hat{\mathbf{J}}^2, \hat{J}_z\}$, only the magnetic component of the total angular momentum m_j can actually be measured "sharply" simultaneously with l, s, and $j - m_l$ and m_s, on the other hand, cannot. On the other side, it is not possible in the product basis $\{\hat{\mathbf{L}}^2, \hat{L}_z, \hat{\mathbf{S}}^2, \hat{S}_z\}$ to measure the magnitude of the total angular momentum and its z-component simultaneously with the other observables of the product basis. The choice of basis depends on the respective problem. In a strong external magnetic field, the interactions between the individual angular momenta and the magnetic field are so great that the total spin quantum number j is no longer a good quantum number to describe the states – the angular momenta $\hat{\mathbf{L}}$ and $\hat{\mathbf{S}}$ each couple individually to the external magnetic field, which allows the quantum numbers l and s to be determined (Paschen-Back effect). The product basis is therefore favorable for measurements with a strong external magnetic field. In a weak external magnetic field, on the other hand, the interaction between spin and orbital angular momentum is the dominant influence on the system behavior, and therefore the coupled basis is preferable. Here, the total angular momentum $\hat{\mathbf{J}}$ couples to the magnetic field.

7.5.2 States in the Product Basis and Coupled Basis

Given two general angular momenta $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$. For example, $\hat{\mathbf{J}}_1$ could be the orbital angular momentum $\hat{\mathbf{L}}$, and $\hat{\mathbf{J}}_2$ the spin $\hat{\mathbf{S}}$ of a hydrogen atom. However, we deliberately choose a general formulation here to clarify that the following relations hold for any angular momentum operators. The total angular momentum of the system thus becomes:

$$\hat{\mathbf{J}} = \hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2 \tag{7.67}$$

In what interval do the quantum numbers j and m_j of the operators $\hat{\mathbf{J}}^2$ and $\hat{J}_z$ lie? While the magnitudes of $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$ are fixed, their direction relative to each other can change – it is therefore possible for the individual angular momenta to point in the same direction, or in the opposite direction. It follows:

$$|j_1 - j_2| = j_{\min} \leq j \leq j_{\max} = j_1 + j_2 \tag{7.68}$$

The calculation of m_j is simpler, as each combination of $m_{j_1} + m_{j_2}$ must exist. The magnetic quantum number of the total angular momentum is constrained by the maximum (anti-)parallel alignment of $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$. It holds:

$$m_{j,\max} = m_{j_1,\max} + m_{j_2,\max} \tag{7.69}$$

For each j_i, $(2j_i + 1)$ states can be found, which differ only in their magnetic quantum number. Considering a system composed of two angular momenta $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$ with the quantum numbers j_1 and j_2, in the context of the product basis, we obtain the following total number of states:

$$p(j_1, j_2) = (2j_1 + 1)(2j_2 + 1) \tag{7.70}$$

The angular momenta $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$ are independent of each other when viewed in the product basis and each provides for itself a certain number of states $(2j_1 + 1)$ and $(2j_2 + 1)$. At the same time, exactly the same number of states must exist in the coupled basis; for every possible value of j between $j_{\min} \leq j \leq j_{\max}$, there should again be exactly $(2j + 1)$ realizable states. We thus need to verify that this is consistent with the total number of allowed values for j according to

(7.70). For $j_1 \geq j_2$, it holds:

$$\begin{aligned}
p(j_1, j_2) &= \sum_{j=j_1-j_2}^{j=j_1+j_2} (2j+1) = \\
&= 2\sum_{k=0}^{2j_2} \left[2\left(k + j_1 - j_2\right) + 1\right] = \\
&= 2\sum_{k=0}^{2j_2} k + \sum_{k=0}^{2j_2} (2j_1 - 2j_2 + 1) = \quad \Big| \sum_{k=0}^{n} k = \frac{n}{2}(n+1) \\
&= 2\frac{2j_2}{2}(2j_2+1) + (2j_2+1)(2j_1 - 2j_2 + 1) = \\
&= 2j_2(2j_2+1) - (2j_2+1)\,2j_2 + (2j_2+1)(2j_1+1) = \\
&= (2j_1+1)(2j_2+1) \quad \square
\end{aligned} \tag{7.71}$$

As expected, we obtain the same number of realizable states in both the product basis and the coupled basis. The total number of states remains conserved during the transformation into the new basis!

Example: Product Basis and Coupled Basis for a System with $j_1 = \frac{5}{2}$ and $j_2 = 1$

To illustrate the concept of the coupled basis, let's consider a concrete example with $j_1 = \frac{5}{2}$ and $j_2 = 1$. In the sense of the product basis (7.70), we obtain a total of $(2\cdot\frac{5}{2}+1)\cdot(2\cdot 1+1) = 6 \cdot 3 = 18$ basis vectors, each with different quantum numbers $\{j, m_j\}$. For the magnitude quantum number j of the operator $\hat{\mathbf{J}}^2$, it holds according to (7.68):

$$\left|\frac{5}{2} - 1\right| = \frac{3}{2} \leq j \leq \frac{7}{2} = \frac{5}{2} + 1$$

The schematic representation below shows how the angular momenta can align in the extreme cases $j_{\max} = j_1 + j_2$ and $j_{\min} = |j_1 - j_2|$.

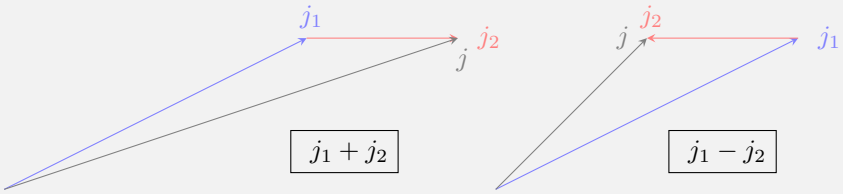


Fig. 44: Maximum and minimum total angular momentum quantum number j.

We find all values for j by starting from $j_{\min} = \frac{3}{2}$ and repeatedly adding $+1$, until we reach $j_{\max} = \frac{7}{2}$. In total, the possible values for j are

$$j \in \left\{\frac{3}{2}, \frac{5}{2}, \frac{7}{2}\right\}$$

All values for m_j must now fall within the interval $-j \leq m_j \leq j$ for each j. Starting at $m_{j,\min}$, successive increasing by $+1$ yields all possible values of the magnetic quantum number. Since in this example there are three values for j, we need to repeat this procedure three times for each individual j. Thus, we find all possible states in the coupled basis to

be:

$$j = \frac{3}{2} : \; m_j \in \left\{-\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}\right\}$$
$$j = \frac{5}{2} : \; m_j \in \left\{-\frac{5}{2}, -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}, \frac{5}{2}\right\}$$
$$j = \frac{7}{2} : \; m_j \in \left\{-\frac{7}{2}, -\frac{5}{2}, -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \frac{7}{2}\right\}$$

We can also graphically identify all possible m_j: Using m_{j_1} and m_{j_2} as coordinates, we mark every pair $\{m_{j_1}, m_{j_2}\}$ as shown in the left graphic of Figure 45. We find $p(j_1, j_2) = 18$ distinct points, and by summing $m_j = m_{j_1} + m_{j_2}$ at each point, it immediately becomes clear that all points on the diagonals have the same m_j value.

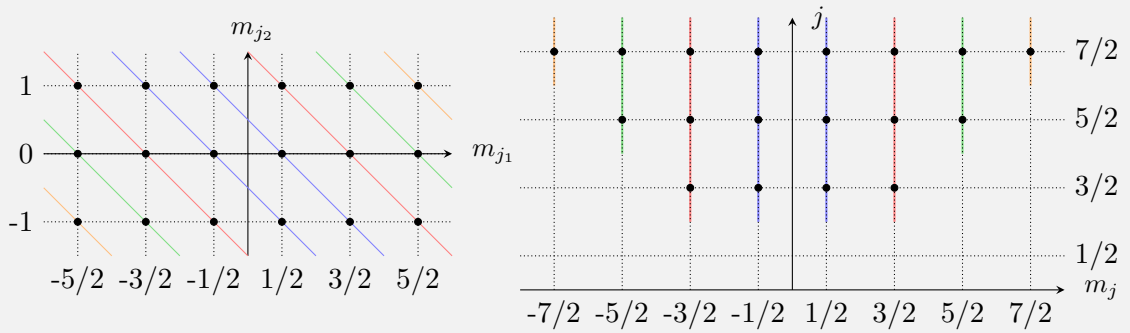


Fig. 45: (Left) Representation of all states in the product basis. (Right) The same system is shown in the coupled basis.

We draw another coordinate system, as shown in the right graphic of Figure 45, this time with the axes m_j and j. Here, we can once again plot all 18 points: At the height of $j = \frac{3}{2}$, we plot horizontally all points from $m_j = -\frac{3}{2}$ to $m_j = +\frac{3}{2}$; at the height $j = \frac{5}{2}$, there are horizontal points from $m_j = -\frac{5}{2}$ to $m_j = +\frac{5}{2}$, and finally, at $j = \frac{7}{2}$, m_j runs from $-\frac{7}{2}$ to $+\frac{7}{2}$. Both graphs can be directly transformed into each other: In the left graph (product basis), the diagonal lines become vertical coordinate lines in the right graph (coupled basis) – a strong indication that we can switch from the product basis to the coupled basis through linear transformations. Except for $m_j = \pm\frac{7}{2}$, there are multiple values for j for every m_j.

7.5.3 Clebsch-Gordan Coefficients

Let us consider a system whose eigenstates $|l, m_l; s, m_s\rangle$ are known in the product basis for the respective observables $\{\hat{\mathbf{L}}^2, \hat{L}_z; \hat{\mathbf{S}}^2, \hat{S}_z\}$. We want to represent the eigenstate $|l, s, j, m_j\rangle$ in the coupled basis $\{\hat{\mathbf{L}}^2, \hat{\mathbf{S}}^2, \hat{\mathbf{J}}^2, \hat{J}_z\}$ as a linear combination of the eigenstates of the product basis. For this, we perform the following linear transformation:

$$|l, s, j, m_j\rangle = \mathbb{1}\,|l, s, j, m_j\rangle = \sum_{m_l, m_s} \langle l, m_l; s, m_s | l, s, j, m_j\rangle\, |l, m_l; s, m_s\rangle \tag{7.72}$$

Thus $|l, s, j, m_j\rangle$ is expanded in the eigenfunctions $|l, m_l; s, m_s\rangle$: The expansion coefficient $\langle l, m_l; s, m_s | l, s, j, m_j\rangle$ is called the *Clebsch-Gordan coefficient*. For the eigenstates of the coupled basis, the same applies as for the product basis in (7.59):

$$\langle l, s, j, m_j | l', s', j', m_j'\rangle = \delta_{ll'}\delta_{ss'}\delta_{jj'}\delta_{m_j m_j'} \tag{7.73}$$

If we carry out the transformation explicitly, we must note that we keep the quantum numbers l and s the same and only transform m_l and m_s into the corresponding j and m_j. For (7.73) it

can be further elaborated as follows:

$$\begin{aligned}\delta_{jj'}\delta_{m_j m'_j} &= \langle l,s,j,m_j|\mathbb{1}|l,s,j',m'_j\rangle = \\ &= \sum_{m_l,m_s} \langle l,s,j,m_j|l,m_l;s,m_s\rangle \langle l,m_l;s,m_s|l,s,j',m'_j\rangle = \\ &= \sum_{m_l,m_s} \langle l,s,j,m_j|l,m_l;s,m_s\rangle \langle l,s,j',m'_j|l,m_l;s,m_s\rangle \end{aligned} \tag{7.74}$$

We use $\langle l,m_l;s,m_s|l,s,j',m'_j\rangle = \langle l,s,j',m'_j|l,m_l;s,m_s\rangle^* = \langle l,s,j',m'_j|l,m_l;s,m_s\rangle$ for *real* Clebsch-Gordan coefficients. Associating $\langle l,s,j,m_j|l,m_l;s,m_s\rangle$ with the real transformation matrix A_{ij} and forming the matrix product $A_{ij}A_{kj}$ – with the same assumptions as in (7.74) we obtain:

$$\sum_j A_{ij}A_{kj} = \sum_j A_{ij}A^T_{jk} \overset{!}{=} \delta_{ik} = \mathbb{1} = AA^{-1}$$

From the comparison with the transformation matrices A, we immediately recognize that the transformation matrix from the product basis to the coupled basis is represented by an orthogonal matrix $A^{-1} = A^T$.
An analogous derivation can be made for the transformation from the coupled basis to the product basis. This gives the inverse transformation matrix (in this case (7.59) would apply). If we already know A, we can obtain the inverse by simply transposing the matrix due to orthogonality!

So far, the calculation of the Clebsch-Gordan coefficients or the explicit transformation between the two representations has been treated very abstractly: In the following examples, we will take a closer look at the transition between the product and the coupled basis. We will deal with very simple cases here, but the same procedures can easily be transferred to more complex systems.

Example: Clebsch-Gordan coefficients for a system with $s_1 = \frac{1}{2}$ and $s_2 = \frac{1}{2}$

In this example, the angular momenta $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$, previously referred to generally, are the spins $\hat{\mathbf{S}}_1$ and $\hat{\mathbf{S}}_2$ of two spin-1/2 particles. The product basis is thus $\{\hat{\mathbf{S}}_1^2, \hat{S}_{z,1}; \hat{\mathbf{S}}_2^2, \hat{S}_{z,2}\}$, and a corresponding eigenstate is generally represented as $|s_1,m_{s,1};s_2,m_{s,2}\rangle$ with four quantum numbers. Since the spin quantum numbers $s_1 = \frac{1}{2}$ and $s_2 = \frac{1}{2}$ are fixed (and remain the same in the coupled basis), we do not have to write them explicitly all the time, and we can therefore indicate all states in the following, shortened notation:

$$|s_1,m_{s,1};s_2,m_{s,2}\rangle \equiv |m_{s,1};m_{s,2}\rangle$$

We can indicate the product basis concisely as $\{\hat{S}_{z,1}, \hat{S}_{z,2}\}$. According to (7.70), the total number of states is $(2\cdot\frac{1}{2}+1)\cdot(2\cdot\frac{1}{2}+1) = 2\cdot 2 = 4$. The realizable states are thus given by:

$$|m_{s,1};m_{s,2}\rangle \in \{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\} = \{|\phi_i\rangle\}$$

Moreover, it should be noted that an operator from the Hilbert space $\mathcal{H}_s^{(1)}$ can only act on the first state. For example, $\hat{S}_z^{(1)}$ acts as follows:

$$\hat{S}_z^{(1)}|m_{s,1};m_{s,2}\rangle \equiv (\hat{S}_z^{(1)}\otimes\mathbb{1}^{(2)})(|m_{s,1}\rangle\otimes|m_{s,2}\rangle) \overset{(7.25)}{=} \hbar m_{s,1}|m_{s,1};m_{s,2}\rangle$$

Analogously, we can determine the effect for $\hat{S}_z^{(2)}$, an operator from $\mathcal{H}_s^{(2)}$. The other state is thus not affected by the operator. But what if we act on a state from $|m_{s,1};m_{s,2}\rangle$ with

$\hat{\mathbf{S}}^2$ and $\hat{S}_z$? Let's start with $\hat{S}_z = \hat{S}_z^{(1)} + \hat{S}_z^{(2)}$:

$$\begin{aligned}
\hat{S}_z\,|\uparrow\uparrow\rangle &= (\hat{S}_z^{(1)} + \hat{S}_z^{(2)})\,|\uparrow\uparrow\rangle = \hat{S}_z^{(1)}\,|\uparrow\uparrow\rangle + \hat{S}_z^{(2)}\,|\uparrow\uparrow\rangle \overset{(7.29)}{=} \tfrac{\hbar}{2}\,|\uparrow\uparrow\rangle + \tfrac{\hbar}{2}\,|\uparrow\uparrow\rangle = \hbar\,|\uparrow\uparrow\rangle \\
\hat{S}_z\,|\uparrow\downarrow\rangle &= (\hat{S}_z^{(1)} + \hat{S}_z^{(2)})\,|\uparrow\downarrow\rangle = \hat{S}_z^{(1)}\,|\uparrow\downarrow\rangle + \hat{S}_z^{(2)}\,|\uparrow\downarrow\rangle \overset{(7.29)}{=} \tfrac{\hbar}{2}\,|\uparrow\downarrow\rangle - \tfrac{\hbar}{2}\,|\uparrow\downarrow\rangle = 0 \\
\hat{S}_z\,|\downarrow\uparrow\rangle &= (\hat{S}_z^{(1)} + \hat{S}_z^{(2)})\,|\downarrow\uparrow\rangle = \hat{S}_z^{(1)}\,|\downarrow\uparrow\rangle + \hat{S}_z^{(2)}\,|\downarrow\uparrow\rangle \overset{(7.29)}{=} -\tfrac{\hbar}{2}\,|\downarrow\uparrow\rangle + \tfrac{\hbar}{2}\,|\downarrow\uparrow\rangle = 0 \\
\hat{S}_z\,|\downarrow\downarrow\rangle &= (\hat{S}_z^{(1)} + \hat{S}_z^{(2)})\,|\downarrow\downarrow\rangle = \hat{S}_z^{(1)}\,|\downarrow\downarrow\rangle + \hat{S}_z^{(2)}\,|\downarrow\downarrow\rangle \overset{(7.29)}{=} -\tfrac{\hbar}{2}\,|\downarrow\downarrow\rangle - \tfrac{\hbar}{2}\,|\downarrow\downarrow\rangle = -\hbar\,|\downarrow\downarrow\rangle
\end{aligned}$$

We can represent the effect of $\hat{S}_z$ in the product basis as a matrix:

$$\hat{S}_z \begin{pmatrix} |\uparrow\uparrow\rangle \\ |\uparrow\downarrow\rangle \\ |\downarrow\uparrow\rangle \\ |\downarrow\downarrow\rangle \end{pmatrix} \xrightarrow{\{|\phi_i\rangle\}} \hbar \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} |\uparrow\uparrow\rangle \\ |\uparrow\downarrow\rangle \\ |\downarrow\uparrow\rangle \\ |\downarrow\downarrow\rangle \end{pmatrix} \tag{7.75}$$

The blue-colored area of the matrix corresponds to a degenerate subspace. For the eigenstates of the product basis $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$, we find the eigenvalue 0 in both cases. A similar degeneracy will also be found for the modulus operator $\hat{\mathbf{S}}^2 = (\hat{\mathbf{S}}^{(1)} + \hat{\mathbf{S}}^{(2)})^2$, where we must first bring the operator $\hat{\mathbf{S}}^2$ to a usable form:

$$\begin{aligned}
\hat{\mathbf{S}}^2 &= (\hat{\mathbf{S}}^{(1)} + \hat{\mathbf{S}}^{(2)})^2 = \hat{\mathbf{S}}^{(1)2} + \hat{\mathbf{S}}^{(2)2} + 2\hat{\mathbf{S}}^{(1)}\hat{\mathbf{S}}^{(2)} = \\
&= \hat{\mathbf{S}}^{(1)2} + \hat{\mathbf{S}}^{(2)2} + 2\left(\hat{S}_x^{(1)}\hat{S}_x^{(2)} + \hat{S}_y^{(1)}\hat{S}_y^{(2)} + \hat{S}_z^{(1)}\hat{S}_z^{(2)}\right) \\
&= \hat{\mathbf{S}}^{(1)2} + \hat{\mathbf{S}}^{(2)2} + 2\hat{S}_z^{(1)}\hat{S}_z^{(2)} + 2\Sigma
\end{aligned}$$

The last term is problematic insofar as $\hat{S}_x^{(1)}\hat{S}_x^{(2)}$ and $\hat{S}_y^{(1)}\hat{S}_y^{(2)}$ will definitely lead to off-diagonal elements. With the help of (7.39), we can rewrite Σ:

$$\begin{aligned}
\Sigma &= \hat{S}_x^{(1)}\hat{S}_x^{(2)} + \hat{S}_y^{(1)}\hat{S}_y^{(2)} \overset{(7.39)}{=} \\
&= \frac{1}{2}\left(\hat{S}_+^{(1)} + \hat{S}_-^{(1)}\right)\frac{1}{2}\left(\hat{S}_+^{(2)} + \hat{S}_-^{(2)}\right) + \frac{1}{2i}\left(\hat{S}_+^{(1)} - \hat{S}_-^{(1)}\right)\frac{1}{2i}\left(\hat{S}_+^{(2)} - \hat{S}_-^{(2)}\right) = \\
&= \frac{1}{4}\left(\hat{S}_+^{(1)}\hat{S}_+^{(2)} + \hat{S}_+^{(1)}\hat{S}_-^{(2)} + \hat{S}_-^{(1)}\hat{S}_+^{(2)} + \hat{S}_-^{(1)}\hat{S}_-^{(2)}\right) \\
&\quad - \frac{1}{4}\left(\hat{S}_+^{(1)}\hat{S}_+^{(2)} - \hat{S}_+^{(1)}\hat{S}_-^{(2)} - \hat{S}_-^{(1)}\hat{S}_+^{(2)} + \hat{S}_-^{(1)}\hat{S}_-^{(2)}\right) = \\
&= \frac{1}{4}\Big(\hat{S}_+^{(1)}\hat{S}_+^{(2)} + \hat{S}_+^{(1)}\hat{S}_-^{(2)} + \hat{S}_-^{(1)}\hat{S}_+^{(2)} + \hat{S}_-^{(1)}\hat{S}_-^{(2)} \\
&\quad - \hat{S}_+^{(1)}\hat{S}_+^{(2)} + \hat{S}_+^{(1)}\hat{S}_-^{(2)} + \hat{S}_-^{(1)}\hat{S}_+^{(2)} - \hat{S}_-^{(1)}\hat{S}_-^{(2)}\Big) = \\
&= \frac{1}{4}\left(\hat{S}_+^{(1)}\hat{S}_-^{(2)} + \hat{S}_-^{(1)}\hat{S}_+^{(2)}\right)
\end{aligned}$$

Plugging Σ back into our expression for $\hat{\mathbf{S}}^2$ finally takes us to

$$\hat{\mathbf{S}}^2 = \hat{\mathbf{S}}^{(1)2} + \hat{\mathbf{S}}^{(2)2} + 2\hat{S}_z^{(1)}\hat{S}_z^{(2)} + \hat{S}_+^{(1)}\hat{S}_-^{(2)} + \hat{S}_-^{(1)}\hat{S}_+^{(2)} \tag{7.76}$$

With (7.76) as well as (7.28) and (7.30), the effect of $\hat{\mathbf{S}}^2$ on all product states results in:

$$\begin{aligned}
\hat{\mathbf{S}}^2 |\uparrow\uparrow\rangle &\overset{(7.28,7.30)}{=} \tfrac{3}{4}\hbar^2 |\uparrow\uparrow\rangle + \tfrac{3}{4}\hbar^2 |\uparrow\uparrow\rangle + 2\tfrac{\hbar}{2}\tfrac{\hbar}{2} |\uparrow\uparrow\rangle + 0 + 0 = 2\hbar^2 |\uparrow\uparrow\rangle \\
\hat{\mathbf{S}}^2 |\uparrow\downarrow\rangle &\overset{(7.28,7.30)}{=} \tfrac{3}{4}\hbar^2 |\uparrow\downarrow\rangle + \tfrac{3}{4}\hbar^2 |\uparrow\downarrow\rangle + 2\tfrac{\hbar}{2}\left(\tfrac{\hbar}{2}\right) |\uparrow\downarrow\rangle + 0 + \hbar\hbar |\downarrow\uparrow\rangle = \hbar^2 |\uparrow\downarrow\rangle + \hbar^2 |\downarrow\uparrow\rangle \\
\hat{\mathbf{S}}^2 |\downarrow\uparrow\rangle &\overset{(7.28,7.30)}{=} \tfrac{3}{4}\hbar^2 |\downarrow\uparrow\rangle + \tfrac{3}{4}\hbar^2 |\downarrow\uparrow\rangle + 2\left(-\tfrac{\hbar}{2}\right)\tfrac{\hbar}{2} |\downarrow\uparrow\rangle + \hbar\hbar |\uparrow\downarrow\rangle + 0 = \hbar^2 |\downarrow\uparrow\rangle + \hbar^2 |\uparrow\downarrow\rangle \\
\hat{\mathbf{S}}^2 |\downarrow\downarrow\rangle &\overset{(7.28,7.30)}{=} \tfrac{3}{4}\hbar^2 |\downarrow\downarrow\rangle + \tfrac{3}{4}\hbar^2 |\downarrow\downarrow\rangle + 2\left(-\tfrac{\hbar}{2}\right)\left(-\tfrac{\hbar}{2}\right) |\downarrow\downarrow\rangle + 0 + 0 = 2\hbar^2 |\downarrow\downarrow\rangle
\end{aligned}$$

If we represent $\hat{\mathbf{S}}^2$ as a matrix in the product basis, we immediately recognize that we now have to deal with off-diagonal elements – thus $\{|m_{s,1}; m_{s,2}\rangle\}$ are *not* eigenfunctions of $\hat{\mathbf{S}}^2$:

$$\hat{\mathbf{S}}^2 \begin{pmatrix} |\uparrow\uparrow\rangle \\ |\uparrow\downarrow\rangle \\ |\downarrow\uparrow\rangle \\ |\downarrow\downarrow\rangle \end{pmatrix} \xrightarrow{\{|\phi_i\rangle\}} \hbar^2 \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} |\uparrow\uparrow\rangle \\ |\uparrow\downarrow\rangle \\ |\downarrow\uparrow\rangle \\ |\downarrow\downarrow\rangle \end{pmatrix} \tag{7.77}$$

This is a consequence of the non-vanishing commutator $[\hat{\mathbf{S}}^2, S_z^{(i)}] \neq 0$. The goal is now to diagonalize $\hat{\mathbf{S}}^2$ and $\hat{S}_z$. We have already found two eigenstates of the coupled basis, namely $|\uparrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$, as they are already diagonal for $\hat{\mathbf{S}}^2$ and $\hat{S}_z$! However, the degenerate subspace (in blue) from (7.77) still needs to be diagonalized.

Let's first make some general considerations about the coupled basis $\{\hat{\mathbf{S}}^2, \hat{S}_z\}$, whose eigenstates we will denote as $|S, M_s\rangle$. The possible values of S are determined by the relation (7.68):

$$|s_1 - s_2| = \left|\frac{1}{2} - \frac{1}{2}\right| = 0 \leq S \leq 1 = \frac{1}{2} + \frac{1}{2} = s_1 + s_2$$

For each value of S, $M_{S,\max} = S$ and $M_{S,\min} = -S$ holds. In the case of this example, we obtain four different realizations of an eigenstate in the coupled basis. This must of course be so, as there are also four states in the product basis. The four states of the coupled basis are:

$$|S, M_s\rangle \in \{|0,0\rangle, |1,-1\rangle, |1,0\rangle, |1,+1\rangle\}$$

The state with $S = 0$ is referred to as the singlet state, while we refer to the three states with $S = 1$ as triplet states. A graphical representation allows a simple confirmation of the form of the upper states in the coupled basis!

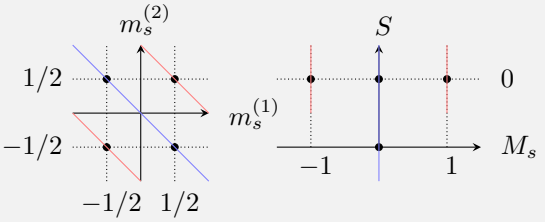


Fig. 46: (Left) Representation of the states of our system in the product basis. (Right) These same states are represented in the coupled basis.

Each state of such a multiplet can be transferred into another state using a ladder operator. Suppose, for example, that we repeatedly apply the lowering operator on the state with the highest quantum number $M_{s,\max}$, we will traverse all realizable states of the multiplet until we reach the state with the lowest quantum number $M_{s,\min}$.

Another possibility to generate $|S, M_s\rangle$ is a rotation of $|m_{s,1}; m_{s,2}\rangle$ into the desired state through the diagonalization of the degenerate subspace. Both options will be discussed in the following.

Diagonalization of the Subspace Consider only the degenerate subspace from (7.77) and diagonalize this. The subspace has the basis $\{|b_i\rangle\} = \{|\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle\}$. The eigenvalues (divided by $\hbar$) will correspond to the possible values of the quantum number S of the coupled basis – the eigenvalues M_s we already know from (7.75), they are in both cases $M_s = 0$. It holds:

$$\hat{\mathbf{S}}^2 \begin{pmatrix} |\uparrow\downarrow\rangle \\ |\downarrow\uparrow\rangle \end{pmatrix} \xrightarrow{\{|b_i\rangle\}} \hbar^2 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} |\uparrow\downarrow\rangle \\ |\downarrow\uparrow\rangle \end{pmatrix} \tag{7.78}$$

The eigenvalues of the matrix from (7.78) are $\lambda_1 = 2\hbar$ and $\lambda_2 = 0$, with which we can calculate the following eigenvectors:

$$\lambda_1 = 2\hbar: \quad |1,0\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \xrightarrow{\{|b_i\rangle\}} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\lambda_1 = 0: \quad |0,0\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \xrightarrow{\{|b_i\rangle\}} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Following the eigenvectors, we can give the transformation matrix T from the product basis to the coupled basis. The elements of this matrix correspond to the Clebsch-Gordan coefficients! It holds:

$$T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{with} \quad \Lambda = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \tag{7.79}$$

Creation Using Ladder Operators According to (7.77), we know that the states $|\uparrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$ of the product basis $\{\hat{S}_{z,1}, \hat{S}_{z,2}\}$ are equivalent to the eigenstates $|1,+1\rangle$ and $|1,-1\rangle$ of the coupled basis $\{\hat{\mathbf{S}}^2, \hat{S}_z\}$. These two states belong to the multiplet with $S = 1$, which contains $2S + 1 = 3$ distinct states, and the third state must evidently be $|1,0\rangle$. However, the representation of $|1,0\rangle$ and also $|0,0\rangle$ in states of the product basis is not as clear as in the cases of $|1,+1\rangle$ and $|1,-1\rangle$, which represent the edges of the triplet.

In the triplet ($S = 1$), we can proceed as follows: We look for the state with the highest quantum number (here $M_S = M_{S,\max} = S = 1$), i.e., $|1,1\rangle = |\uparrow\uparrow\rangle$, and apply the lowering operator from (7.26) to this coupled as well as product state:

$$\left.\begin{aligned} \hat{S}_- |1,1\rangle &= \hbar\sqrt{2}\,|1,-1\rangle \\ \hat{S}_- |\uparrow\uparrow\rangle &= (\hat{S}_-^{(1)} + \hat{S}_-^{(2)})\,|\uparrow\uparrow\rangle = \hbar\,(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \end{aligned}\right\} \implies |1,0\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$$

Thus, $|1,0\rangle$ consists of a linear combination of the mixed states, which is symmetric with respect to the exchange of both particles. Reapplying brings us into the state with the lowest quantum number ($M_S = -1$) in the triplet:

$$\left.\begin{aligned} \hat{S}_- |1,0\rangle &= \hbar\sqrt{2}\,|1,0\rangle \\ \hat{S}_- \tfrac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) &= \tfrac{1}{\sqrt{2}}(\hat{S}_-^{(1)} + \hat{S}_-^{(2)})(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) = \hbar\sqrt{2}\,|\downarrow\downarrow\rangle \end{aligned}\right\} \implies |1,-1\rangle = |\downarrow\downarrow\rangle$$

This gives us the „lowest“ state $|1,-1\rangle$ of the triplet. The single state $|0,0\rangle$ of the singlet ($S = 0$) can be generated by exploiting the orthogonality of the coupled basis states.

By assuming a general linear combination $|0,0\rangle = \alpha\,|\uparrow\downarrow\rangle + \beta\,|\downarrow\uparrow\rangle$, we can determine the coefficients α and β by projection on $|1,1\rangle$, $|1,-1\rangle$, and $|1,0\rangle$. In fact, only $|1,0\rangle$ is relevant (which lies with $|0,0\rangle$ in the degenerate subspace), from which it follows:

$$0 \stackrel{!}{=} \langle 1,0|0,0\rangle = \frac{1}{\sqrt{2}}\left(\langle\uparrow\downarrow| + \langle\downarrow\uparrow|\right)\left(\alpha\,|\uparrow\downarrow\rangle + \beta\,|\downarrow\uparrow\rangle\right) = \frac{1}{\sqrt{2}}\left(\alpha+\beta\right) \implies \beta = -\alpha$$

Since the state $|0,0\rangle$ must be normalized, we choose $\alpha = 1/\sqrt{2}$ and thus also $\beta = -1/\sqrt{2}$. By applying the ladder operators and the orthonormality condition, we have found a complete set of states in the coupled basis and their representation in the product basis:

$$|S,M_s\rangle = \begin{cases} |0,0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \\ |1,0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \\ |1,-1\rangle = |\downarrow\downarrow\rangle \\ |1,+1\rangle = |\uparrow\uparrow\rangle \end{cases}$$

The factors $\pm 1/\sqrt{2}$ again correspond to the sought Clebsch-Gordan coefficients. Thus with both calculation methods we equally find the following transformation:

$$\begin{pmatrix} |1,0\rangle \\ |0,0\rangle \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} |\uparrow\downarrow\rangle \\ |\downarrow\uparrow\rangle \end{pmatrix}$$

For the inverse transformation, the transposed transformation matrix can be used, as we are dealing with a unitary transformation matrix:

$$\begin{pmatrix} |\uparrow\downarrow\rangle \\ |\downarrow\uparrow\rangle \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} |1,0\rangle \\ |0,0\rangle \end{pmatrix}$$

Overall, for the action of $\hat{\mathbf{S}}^2$ on the coupled eigenstates in the matrix representation, we get:

$$\hat{\mathbf{S}}^2 \begin{pmatrix} |1,+1\rangle \\ |1,0\rangle \\ |0,0\rangle \\ |1,-1\rangle \end{pmatrix} \xrightarrow{\{|S,M_s\rangle\}} \hbar^2 \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} |1,+1\rangle \\ |1,0\rangle \\ |0,0\rangle \\ |1,-1\rangle \end{pmatrix} \tag{7.80}$$

As expected, there are no off-diagonal elements left in the coupled basis! At the same time, $\hat{S}_z$ remains diagonal in the matrix representation of the coupled basis, because the transformation was carried out in a degenerate subspace that diagonalizes $\hat{S}_z$:

$$\hat{S}_z \begin{pmatrix} |1,+1\rangle \\ |1,0\rangle \\ |0,0\rangle \\ |1,-1\rangle \end{pmatrix} \xrightarrow{\{|S,M_s\rangle\}} \hbar \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} |1,+1\rangle \\ |1,0\rangle \\ |0,0\rangle \\ |1,-1\rangle \end{pmatrix} \tag{7.81}$$

In practice, Clebsch-Gordan coefficients do not have to be calculated individually for each system of angular momenta. Usually, it is sufficient to look up the correct values in tables. A small excerpt of such tables can be found in the lower three tables: Here, all coefficients for $j_1 \times j_2 = \frac{1}{2} \times \frac{1}{2}$, $1 \times \frac{1}{2}$, and 1×1 are listed. The tables are read as follows: The gray shaded values (bottom right) correspond to the Clebsch-Gordan coefficients, where a shortened representation is common: A value 1/2 thus means $\sqrt{1/2}$! To the left of the gray shaded areas are the values of m_1 and m_2 in the product basis, and above the gray shaded area are the quantum numbers J and m_j of the coupled basis. In all three systems, the „highest“ and „lowest“ states ($m_{j_{\max}}$ and $m_{j_{\min}}$) are the same in the product basis and the coupled basis.

$\mathbf{j_1} \times \mathbf{j_2}$		J
		m_j
m_1	m_2	CG-coefficient

$\mathbf{1/2 \times 1/2}$		1			
		$+1$	1	0	
$+1/2$	$+1/2$	1	0	0	
$+1/2$	$-1/2$	$1/2$	$1/2$	1	
$-1/2$	$+1/2$	$1/2$	$-1/2$	-1	
		$-1/2$	$-1/2$	1	

$\mathbf{1 \times 1/2}$		$3/2$					
		$+3/2$	$3/2$	$1/2$			
$+1$	$+1/2$	1	$+1/2$	$+1/2$			
$+1$	$-1/2$	$1/3$	$2/3$	$3/2$	$1/2$		
0	$+1/2$	$2/3$	$-1/3$	$-1/2$	$-1/2$		
		0	$-1/2$	$2/3$	$1/3$	$3/2$	
		-1	$+1/2$	$1/3$	$-2/3$	$-3/2$	
				-1	$-1/2$	1	

$\mathbf{1 \times 1}$		2						
		$+2$	2	1				
$+1$	$+1$	1	$+1$	$+1$				
$+1$	0	$1/2$	$1/2$	2	1	0		
0	$+1$	$1/2$	$-1/2$	0	0	0		
		$+1$	-1	$1/6$	$1/2$	$1/3$		
		0	0	$2/3$	0	$-1/3$	2	1
		-1	$+1$	$1/6$	$-1/2$	$1/3$	-1	-1
				0	-1	$1/2$	$1/2$	2
				-1	0	$1/2$	$-1/2$	-2
						-1	-1	1

Example: Time Evolution of Particles with $l = 1$ in the Anisotropic Crystal Field

We consider a system of a free particle with angular momentum $l = 1$, which is in an anisotropic crystal field. The Hamilton operator $\hat{H}$ thus includes not only the rotational energy but also an energy component influenced by the external field. For $\hat{H}$ in this example (we abbreviate here $|l, m\rangle \equiv |m\rangle$) the following applies:

$$\hat{H}_{\mathrm{KF}} = \frac{\hat{\mathbf{L}}^2}{2I} + \gamma \left(|-1\rangle \langle +1| + |+1\rangle \langle -1| + |0\rangle \langle 0| \right) \tag{7.82}$$

I is in this case the moment of inertia of the particle, γ is the coupling to the crystal field. We already know the eigenstates of the modulus operator of the angular momentum and its action on these:

$$|1, m\rangle \equiv |m\rangle \in \{|-1\rangle, |0\rangle, |+1\rangle\}$$

Applying $\hat{H}$ to the eigenstates of $\hat{\mathbf{L}}^2$, we obtain in the matrix representation the following result:

$$\hat{H}_{\mathrm{KF}} \begin{pmatrix} |-1\rangle \\ |0\rangle \\ |+1\rangle \end{pmatrix} = \left[\frac{\hbar^2}{2I} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} + \gamma \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \right] \begin{pmatrix} |-1\rangle \\ |0\rangle \\ |+1\rangle \end{pmatrix} \tag{7.83}$$

The expression in the bracket is obviously *not* diagonal but leads to mixtures of $|m\rangle$; the eigenfunctions of $\hat{\mathbf{L}}^2$ are therefore not eigenfunctions of $\hat{H}_{\mathrm{KF}}$. We find these by diagonalizing the matrix H_{KF} – not only do we obtain the eigenfunctions of the Hamilton operator but

also its eigenvalues! It holds:

$$H_{\mathrm{KF}} = \begin{pmatrix} \hbar^2/I & 0 & \gamma \\ 0 & \hbar^2/I + \gamma & 0 \\ \gamma & 0 & \hbar^2/I \end{pmatrix}$$

The eigenvalues follow from $\det|H_{\mathrm{KF}} - \boldsymbol{\lambda}\mathbb{1}| = 0$. In this case we obtain $\lambda_1 = \hbar^2/I - \gamma$ and the doubly degenerate eigenvalue $\lambda_{2,3} = \hbar^2/I + \gamma$. For each eigenvalue, we find the respective eigenfunction $|\lambda_i\rangle$:

$$|\lambda_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \frac{|-1\rangle - |+1\rangle}{\sqrt{2}},\ |\lambda_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = |0\rangle\,,\ |\lambda_3\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \frac{|-1\rangle + |+1\rangle}{\sqrt{2}}$$

Compiling our eigenvectors into the transformation matrix T and the eigenvalues into a diagonal matrix Λ, we obtain:

$$T = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ -1 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \Lambda = \begin{pmatrix} \hbar^2/I - \gamma & 0 & 0 \\ 0 & \hbar^2/I + \gamma & 0 \\ 0 & 0 & \hbar^2/I + \gamma \end{pmatrix} \tag{7.84}$$

We are now interested in the time evolution of any state of this system. Arbitrarily, we choose $|+1\rangle$ here and proceed according to (2.39) – the only time-dependent term comes from the exponential function (or rather the time evolution operator, which we will learn more rigorously in `Quantentheorie II`). We already know $|\psi(t)\rangle$ at time $t = 0$: $|\psi(0)\rangle = |+1\rangle$. For arbitrary times t it holds:

$$|\psi(t)\rangle = e^{i\hat{H}t/\hbar}\,|+1\rangle$$

Without delving further into operator-valued functions: If the description is in the eigenbasis of the Hamilton operator, $\hat{H}$ in the exponential function simply corresponds to the eigenvalue from (7.84). However, we then also have to give $|+1\rangle$ in the eigenbasis of $\hat{H}$ – by subtracting $|\lambda_1\rangle$ from $|\lambda_3\rangle$, we obtain $|+1\rangle = (|\lambda_3\rangle - |\lambda_1\rangle)/\sqrt{2}$. It follows:

$$\begin{aligned} |\psi(t)\rangle &= \frac{1}{\sqrt{2}}\left(e^{i\lambda_3 t/\hbar}\,|\lambda_3\rangle - e^{i\lambda_1 t/\hbar}\,|\lambda_1\rangle\right) = \\ &= \frac{1}{2}\left[e^{i\lambda_3 t/\hbar}\,(|-1\rangle + |+1\rangle) - e^{i\lambda_1 t/\hbar}\,(|-1\rangle - |+1\rangle)\right] = \\ &= \frac{1}{2}\left[\left(e^{i\left(\frac{\hbar}{I}+\frac{\gamma}{\hbar}\right)t} - e^{i\left(\frac{\hbar}{I}-\frac{\gamma}{\hbar}\right)t}\right)|-1\rangle + \left(e^{i\left(\frac{\hbar}{I}+\frac{\gamma}{\hbar}\right)t} + e^{i\left(\frac{\hbar}{I}-\frac{\gamma}{\hbar}\right)t}\right)|+1\rangle\right] = \\ &= e^{i\frac{\hbar}{I}t}\left[i\sin\left(\frac{\gamma}{\hbar}t\right)|-1\rangle + \cos\left(\frac{\gamma}{\hbar}t\right)|+1\rangle\right] \end{aligned} \tag{7.85}$$

The rotation of the phase is determined solely by the size of the moment of inertia I; the coupling to the external field γ, on the other hand, controls in which „direction“ the state $|t\rangle$ points at any time t.

8 Perturbation Theory

Motivation: Approximation Methods and Perturbation Theory in Quantum Theory

The Schrödinger equation can only be solved exactly for very few, highly simplified systems. The potentials that are dealt with within the framework of this script (Coulomb potential, harmonic oscillator) thus represent the exception and not the rule in terms of their exact solvability. As will be shown in this chapter, the Schrödinger equation can be approximately solved for potentials that slightly deviate from these exactly solvable potentials.

The theory that describes such "slightly perturbed" potentials is perturbation theory, which has established itself as a widespread method to compute approximate solutions for analytically not closed solvable quantum mechanical problems.

This chapter deals with the *Rayleigh-Schrödinger perturbation theory*. A central idea of this approximation method is that the Hamiltonian operator $\hat{H}$ can be divided into an unperturbed part $\hat{H}_0$ and a perturbation $\hat{V}$:

$$\hat{H} = \hat{H}_0 + \lambda \hat{V} \tag{8.1}$$

The method is helpful when we can analytically solve the eigenvalue problem of the unperturbed Hamiltonian $\hat{H}_0$, but not the actual problem with the perturbed Hamiltonian $\hat{H}$. We thus start from the following initial situation:

$$\hat{H}_0 \left|\phi_n\right\rangle = \varepsilon_n \left|\phi_n\right\rangle \tag{8.2}$$

If the *order parameter* λ is small enough, the perturbation $\hat{V}$ has only a weak effect, and we can approach the unknown solution for $\hat{H}$ in a series development from the known solution for $\hat{H}_0$. Although the choice of $\hat{H}_0$ and $\hat{V}$ is not always clear, the following conditions must be met to obtain the most accurate results with the least computational effort for perturbative calculations:

- The entire spectrum of eigenfunctions $|\phi_n\rangle$ and eigenenergies ε_n of $\hat{H}_0$ can be calculated *exactly* in the sense of (8.2).
- $\lambda\hat{V}$ is a weak perturbation compared to $\hat{H}_0$, with the order parameter λ scaling the strength of the perturbation.

The last condition will be illustrated in an example. We assume that the eigenenergies are non-degenerate – it follows that for all eigenenergies $\varepsilon_n \neq \varepsilon_{n'}$ $(\forall n \neq n')$. It will later be clear why this condition is an important prerequisite for the functioning of Rayleigh-Schrödinger perturbation theory (in the case of degeneracy, one must use degenerate perturbation theory). Additionally, only *time-independent* Hamiltonians will be examined in this chapter.

Our goal is now to determine the general eigenenergies $\mathcal{E}_n$ and wavefunctions $|\psi_n\rangle$ of the Hamiltonian $\hat{H}$, which fulfill the perturbed Schrödinger equation:

$$\hat{H} \left|\psi_n\right\rangle = \mathcal{E}_n \left|\psi_n\right\rangle \tag{8.3}$$

The index n represents the energetic hierarchy of the state (quantum number). Note: All following derivations apply to a fixed n, such as for the ground state $n = 0$ with $|\psi_0\rangle$ and $\mathcal{E}_0$, or any other (fixed) energy state $|\psi_n\rangle$ and $\mathcal{E}_n$. Let's start with the expansion of the desired eigenfunction $|\psi_n\rangle$ over the (small) order parameter λ:

$$\left|\psi_n\right\rangle = \left|\psi_n^{(0)}\right\rangle + \lambda \left|\psi_n^{(1)}\right\rangle + \lambda^2 \left|\psi_n^{(2)}\right\rangle + \cdots = \sum_{i=0}^{\infty} \lambda^i \left|\psi_n^{(i)}\right\rangle \tag{8.4}$$

We assume that the functions $\{|i\rangle\}$ form an orthonormal system: $\langle\psi_n^{(i)}|\psi_n^{(j)}\rangle = \delta_{ij}$. An analogous series expansion is also established for the desired eigenenergy $\mathcal{E}_n$ of the perturbed Hamiltonian $\hat{H}$:

$$\mathcal{E}_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots = \sum_{i=0}^{\infty} \lambda^i E_n^{(i)} \tag{8.5}$$

We currently know neither the expansion functions $|\psi_n^{(i)}\rangle$ nor the expansion eigenenergies $E_n^{(i)}$. To determine these, we proceed from the Schrödinger equation (8.3), and insert the above series expansions:

$$\left(\hat{H}_0 + \lambda\hat{V}\right)\left(\sum_{i=0}^{\infty} \lambda^i \,|\psi_n^{(i)}\rangle\right) = \left(\sum_{i=0}^{\infty} \lambda^i E_n^{(i)}\right)\left(\sum_{i=0}^{\infty} \lambda^i \,|\psi_n^{(i)}\rangle\right) \tag{8.6}$$

Since we want to expand in powers of λ, it is helpful for clarity to write out the first terms of this equation explicitly:

$$\begin{aligned}&\left(\hat{H}_0 + \lambda\hat{V}\right)\left(|\psi_n^{(0)}\rangle + \lambda\,|\psi_n^{(1)}\rangle + \lambda^2\,|\psi_n^{(2)}\rangle + \ldots\right) = \\ &\left(E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \ldots\right)\left(|\psi_n^{(0)}\rangle + \lambda\,|\psi_n^{(1)}\rangle + \lambda^2\,|\psi_n^{(2)}\rangle + \ldots\right)\end{aligned}$$

We simplify the expression above by multiplying out and reorganizing; for clarity, we group all terms with the same powers of the order parameter λ, including at most the second order λ^2:

$$\begin{aligned}0 = {}&\lambda^0\left(\hat{H}_0\,|\psi_n^{(0)}\rangle - E_0\,|\psi_n^{(0)}\rangle\right) + \\ &+\lambda^1\left(\hat{H}_0\,|\psi_n^{(1)}\rangle + \hat{V}\,|\psi_n^{(0)}\rangle - E_0\,|\psi_n^{(1)}\rangle - E_1\,|\psi_n^{(0)}\rangle\right) + \\ &+\lambda^2\left(\hat{H}_0\,|\psi_n^{(2)}\rangle + \hat{V}\,|\psi_n^{(1)}\rangle - E_0\,|\psi_n^{(2)}\rangle - E_1\,|\psi_n^{(1)}\rangle - E_2\,|\psi_n^{(0)}\rangle\right) + \mathcal{O}(\lambda^3)\end{aligned} \tag{8.7}$$

Each term corresponding to each order λ^i must individually vanish, otherwise the equation (8.7) cannot be satisfied. Solving for each coefficient, we can compute correction terms of higher order step by step.

8.1 Zero-Order Solution

To determine the zero-order solution, the crudest approximate solution, we consistently consider only the zero order λ^0 in (8.7), obtaining:

$$\hat{H}_0\,|\psi_n^{(0)}\rangle = E_n^{(0)}\,|\psi_n^{(0)}\rangle \tag{8.8}$$

(8.8) is determined only by the unperturbed Hamiltonian and thus corresponds to the system without the presence of perturbation. A direct comparison with (8.2) immediately shows:

$$E_n^{(0)} = \varepsilon_n \quad \text{und} \quad |\psi_n^{(0)}\rangle = |\phi_n\rangle \tag{8.9}$$

If only these results are considered in the series expansions (8.4) and (8.5), one recognizes: Unsurprisingly, in zero order, the solutions $|\psi_n\rangle$ and $\mathcal{E}_n$ for the perturbed eigenvalue problem are identical to the solutions $|\phi_n\rangle$ and ε_n of the unperturbed Hamiltonian $\hat{H}_0$! For $\lambda \to 0$, one can write the following solution for $\hat{H}$:

$$\mathcal{E}_n \approx \varepsilon_n \quad \text{and} \quad |\psi_n\rangle \approx |\phi_n\rangle \tag{8.10}$$

8.2 First-Order Solution

Analogously, we now also proceed to determine the first correction term in the series expansions (8.4) and (8.5). Considering the term related to λ^1 from (8.7), we can reformulate it with respect to $|\psi_n^{(0)}\rangle$ and $|\psi_n^{(1)}\rangle$, resulting in:

$$\underbrace{\left(\hat{H}_0 - E_n^{(0)}\right)|\psi_n^{(1)}\rangle}_{\text{(I)}} = \underbrace{\left(E_n^{(1)} - \hat{V}\right)|\psi_n^{(0)}\rangle}_{\text{(II)}} \tag{8.11}$$

By projecting the two relations (I) and (II) from (8.11) onto the bra vector $\langle\psi_n^{(0)}|$, we obtain a closed-form expression for $E_n^{(1)}$, which allows us to calculate the first-order energy correction. According to (8.9), we note that $|\psi_n^{(0)}\rangle = |\phi_n\rangle$, and therefore, as per (8.8), we have the relation $\hat{H}_0\,|\psi_n^{(0)}\rangle = E_n^{(0)}\,|\phi_n\rangle$. We find for the terms (I) and (II):

$$\begin{aligned}
(I) \mid\ & \langle\psi_n^{(0)}|\hat{H}_0 - E_n^{(0)}|\psi_n^{(1)}\rangle = \langle\psi_n^{(0)}|\hat{H}_0|\psi_n^{(1)}\rangle - \langle\psi_n^{(0)}|E_n^{(0)}|\psi_n^{(1)}\rangle = \\
& = E_n^{(0)}\langle\psi_n^{(0)}|\psi_n^{(1)}\rangle - E_n^{(0)}\langle\psi_n^{(0)}|\psi_n^{(1)}\rangle = 0 \\
(II) \mid\ & \langle\psi_n^{(0)}|E_n^{(1)} - \hat{V}|\psi_n^{(0)}\rangle = E_n^{(1)}\langle\psi_n^{(0)}|\psi_n^{(0)}\rangle - \langle\psi_n^{(0)}|\hat{V}|\psi_n^{(0)}\rangle \overset{(8.9)}{=} E_n^{(1)} - \langle\phi_n|\hat{V}|\phi_n\rangle
\end{aligned}$$

Due to the Hermiticity of $\hat{H}_0$, the Hamiltonian operator can act to the left — the advantage of this is that we know the action of $\hat{H}_0$ on $|\psi_n^{(0)}\rangle$, but not on $|\psi_n^{(1)}\rangle$. By equating terms (I) and (II), we thus obtain a simple relation for $E_n^{(1)}$, for which (again using $|\psi_n^{(0)}\rangle = |\phi_n\rangle$):

$$E_n^{(1)} = \langle\phi_n|\hat{V}|\phi_n\rangle \tag{8.12}$$

By considering this energy correction along with the zero-order solution in the series development (8.5), one obtains the solution for $\mathcal{E}_n$ in first order:

$$\mathcal{E}_n \approx \varepsilon_n + \lambda\,\langle\phi_n|\hat{V}|\phi_n\rangle \tag{8.13}$$

To determine the first correction term in the series development (8.4) of $|\psi_n\rangle$, we now project an arbitrary eigenstate of the unperturbed Hamiltonian $\langle\phi_k|$ (thus any eigenstate *except* $\langle\psi_n^{(0)}| = \langle\phi_n|$) from the left onto (8.11):

$$\begin{aligned}
\langle\phi_k|\hat{H}_0 - E_n^{(0)}|\psi_n^{(1)}\rangle &= \langle\phi_k|\hat{H}_0|\psi_n^{(1)}\rangle - \langle\phi_k|\varepsilon_n|\psi_n^{(1)}\rangle = \varepsilon_k\,\langle\phi_k|\psi_n^{(1)}\rangle - \varepsilon_n\,\langle\phi_k|\psi_n^{(1)}\rangle = \\
&= (\varepsilon_k - \varepsilon_n)\,\langle\phi_k|\psi_n^{(1)}\rangle = \langle\phi_k|E_n^{(1)} - \hat{V}|0\rangle = E_n^{(1)}\langle\phi_k|\phi_n\rangle - \langle\phi_k|\hat{V}|\phi_n\rangle = \\
&= -\,\langle\phi_k|\hat{V}|\phi_n\rangle
\end{aligned}$$

Due to the orthogonality of eigenfunctions, the term $\langle\phi_k|\phi_n\rangle$ disappears if $k \neq n$. We have thus found a relation for an (almost) arbitrary expansion coefficient $\langle\phi_k|\psi_n^{(1)}\rangle$ of $|\psi_n^{(1)}\rangle$ in the eigenbasis of the unperturbed Hamiltonian:

$$\langle\phi_k|\psi_n^{(1)}\rangle = \frac{\langle\phi_k|\hat{V}|\phi_n\rangle}{\varepsilon_n - \varepsilon_k} \quad \text{für} \quad k \neq n \tag{8.14}$$

Obviously, in the solution (8.14) we encounter a problem due to the denominator when $k = n$ — our initial assumption that $|\phi_k\rangle$ must *not* be $|\phi_n\rangle = |\psi_n^{(0)}\rangle$ is thus justified in retrospect. Let us now use (8.14) to explicitly compute $|\psi_n^{(1)}\rangle$:

$$\begin{aligned}
|\psi_n^{(1)}\rangle = \mathbb{1}\,|\psi_n^{(1)}\rangle &= \sum_{k=0}^{\infty}|\phi_k\rangle\,\langle\phi_k|\psi_n^{(1)}\rangle = |\phi_n\rangle\,\langle\phi_n|\psi_n^{(1)}\rangle + \sum_{k\neq n}|\phi_k\rangle\,\langle\phi_k|\psi_n^{(1)}\rangle \overset{(8.9)}{=} \\
&= |\phi_n\rangle\,\langle\psi_n^{(0)}|\psi_n^{(1)}\rangle + \sum_{k\neq n}|\phi_k\rangle\,\langle\phi_k|\psi_n^{(1)}\rangle \overset{(8.14)}{=} \sum_{k\neq n}|\phi_k\rangle\,\frac{\langle\phi_k|\hat{V}|\phi_n\rangle}{\varepsilon_n - \varepsilon_k}
\end{aligned}$$

Again, from $|\phi_n\rangle = |\psi_n^{(0)}\rangle$ it follows that $\langle\phi_n|\psi_n^{(1)}\rangle \overset{(8.9)}{=} \langle\psi_n^{(0)}|\psi_n^{(1)}\rangle = 0$ vanishes. The first expansion coefficient for the eigenfunction $|\psi_n\rangle$ of the Hamiltonian $\hat{H}$ thus only encompasses the off-diagonal elements $k \neq n$; the smaller the distance between the neighboring energy levels ε_n and ε_k, the greater the deviation from the unperturbed state $|\phi_k\rangle$. It follows:

$$|\psi_n^{(1)}\rangle = \sum_{k\neq n} \frac{\langle\phi_k|\hat{V}|\phi_n\rangle}{\varepsilon_n - \varepsilon_k} |\phi_k\rangle \tag{8.15}$$

By considering this correction along with the zero-order solution in the series development (8.4), one obtains the following solution for $|\psi_n\rangle$ in first order:

$$|\psi_n\rangle \approx |\phi_n\rangle + \lambda \sum_{k\neq n} \frac{\langle\phi_k|\hat{V}|\phi_n\rangle}{\varepsilon_n - \varepsilon_k} |\phi_k\rangle \tag{8.16}$$

8.3 Second-Order Solution

Lastly, to also determine the second correction terms in the series developments (8.4) and (8.5), we again use the expression related to λ^2 from (8.7). We obtain:

$$\underbrace{\left(\hat{H}_0 - E_n^{(0)}\right)|\psi_n^{(2)}\rangle}_{\text{(I)}} = \underbrace{\left(E_n^{(1)} - \hat{V}\right)|\psi_n^{(1)}\rangle + E_n^{(2)}|\psi_n^{(0)}\rangle}_{\text{(II)}} \tag{8.17}$$

Analogous to the first order correction, we again project (8.17) onto the bra vector $\langle\psi_n^{(0)}|$:

$$\begin{aligned}
\text{(I)} \mid\ & \langle\psi_n^{(0)}|\hat{H}_0 - E_n^{(0)}|\psi_n^{(2)}\rangle = \langle\psi_n^{(0)}|\hat{H}_0|\psi_n^{(2)}\rangle - \langle\psi_n^{(0)}|E_n^{(0)}|\psi_n^{(2)}\rangle = \\
& = E_n^{(0)}\langle\psi_n^{(0)}|\psi_n^{(2)}\rangle - E_n^{(0)}\langle\psi_n^{(0)}|\psi_n^{(2)}\rangle = 0 \\
\text{(II)} \mid\ & \langle\psi_n^{(0)}|E_n^{(1)} - \hat{V}|\psi_n^{(1)}\rangle + \langle\psi_n^{(0)}|E_n^{(2)}|\psi_n^{(0)}\rangle = \\
& = E_n^{(1)}\langle\psi_n^{(0)}|\psi_n^{(1)}\rangle - \langle\psi_n^{(0)}|\hat{V}|\psi_n^{(1)}\rangle + E_n^{(2)}\langle\psi_n^{(0)}|\psi_n^{(0)}\rangle = E_n^{(2)} - \langle\psi_n^{(0)}|\hat{V}|\psi_n^{(1)}\rangle
\end{aligned}$$

We thus find that the second order energy correction depends on $|\psi_n^{(1)}\rangle$. By substituting the aforementioned expression, we obtain:

$$E_n^{(2)} = \langle\psi_n^{(0)}|\hat{V}|\psi_n^{(1)}\rangle \overset{(8.9)}{=} \sum_{k\neq n} \langle\phi_n|\hat{V}|\phi_k\rangle \frac{\langle\phi_k|\hat{V}|\phi_n\rangle}{\varepsilon_n - \varepsilon_k}$$

Thereby, we find a valid expression for the second energy correction. We notice that unlike the first energy correction (8.12), only the off-diagonal elements $k \neq n$ are considered here:

$$E_n^{(2)} = \sum_{k\neq n} \frac{|\langle\phi_k|\hat{V}|\phi_n\rangle|^2}{\varepsilon_n - \varepsilon_k} \tag{8.18}$$

Including this second-order energy correction along with the first-order correction and the zero-order solution in the series expansion (8.5), one obtains the complete solution for $\mathcal{E}_n$ in second order:

$$\mathcal{E}_n \approx \varepsilon_n + \lambda\langle\phi_n|\hat{V}|\phi_n\rangle + \lambda^2 \sum_{k\neq n} \frac{|\langle\phi_k|\hat{V}|\phi_n\rangle|^2}{\varepsilon_n - \varepsilon_k} \tag{8.19}$$

To determine the second correction term in the series expansion of $|\psi_n\rangle$, we again project an arbitrary eigenstate of the unperturbed Hamiltonian $\langle\phi_k|$ (assuming again $|\phi_k\rangle \neq |\phi_n\rangle$) from the left onto (8.17):

$$\begin{aligned}\langle\phi_k|\hat{H}_0 - E_n^{(0)}|\psi_n^{(2)}\rangle &= \langle\phi_k|\hat{H}_0|\psi_n^{(2)}\rangle - \langle\phi_k|\varepsilon_n|\psi_n^{(2)}\rangle = \varepsilon_k \langle\phi_k|\psi_n^{(2)}\rangle - \varepsilon_n \langle\phi_k|\psi_n^{(2)}\rangle = \\ &= (\varepsilon_k - \varepsilon_n)\langle\phi_k|\psi_n^{(2)}\rangle = \langle\phi_k|E_n^{(1)} - \hat{V}|\psi_n^{(1)}\rangle + \langle\phi_k|E_n^{(2)}|\phi_n\rangle = \\ &= E_n^{(1)}\langle\phi_k|\psi_n^{(1)}\rangle - \langle\phi_k|\hat{V}|\psi_n^{(1)}\rangle + E_n^{(2)}\langle\phi_k|\phi_n\rangle\end{aligned}$$

Once more, we find an expression for an expansion coefficient $\langle\phi_k|\psi_n^{(2)}\rangle$ in the eigenbasis of the unperturbed system; however, this time, the first order correction terms $E_n^{(1)}$ from (8.12) and $|\psi_n^{(1)}\rangle$ from (8.15) appear. We write:

$$\langle\phi_k|\psi_n^{(2)}\rangle = \frac{E_n^{(1)}\langle\phi_k|\psi_n^{(1)}\rangle - \langle\phi_k|\hat{V}|\psi_n^{(1)}\rangle}{\varepsilon_n - \varepsilon_k} \overset{(8.12)}{=} \frac{\langle\phi_k|\hat{V}|\psi_n^{(1)}\rangle - \langle\phi_n|\hat{V}|\phi_n\rangle\langle\phi_k|\psi_n^{(1)}\rangle}{\varepsilon_n - \varepsilon_k} \tag{8.20}$$

As before, the initial assumption $k \neq n$ prevents the energy denominator from diverging. Explicitly, we find, analogously to the procedure for $|\psi_n^{(1)}\rangle$, now for $|\psi_n^{(2)}\rangle$:

$$\begin{aligned}|\psi_n^{(2)}\rangle = \mathbb{1}\,|\psi_n^{(2)}\rangle &= |\phi_n\rangle\langle\phi_n|\psi_n^{(2)}\rangle + \sum_{k\neq n}|\phi_k\rangle\langle\phi_k|\psi_n^{(2)}\rangle \overset{(8.9)}{=} \\ &= |\phi_n\rangle\langle\psi_n^{(0)}|\psi_n^{(2)}\rangle + \sum_{k\neq n}|\phi_k\rangle\langle\phi_k|\psi_n^{(2)}\rangle \overset{(8.14)}{=} \\ &= \sum_{k\neq n}|\phi_k\rangle\frac{\langle\phi_k|\hat{V}|\psi_n^{(1)}\rangle - \langle\phi_n|\hat{V}|\phi_n\rangle\langle\phi_k|\psi_n^{(1)}\rangle}{\varepsilon_n - \varepsilon_k}\end{aligned}$$

Specifically, the expression for $|\psi_n^{(2)}\rangle$ results in the following:

$$|\psi_n^{(2)}\rangle = \sum_{k\neq n}\frac{\langle\phi_k|\hat{V}|\psi_n^{(1)}\rangle - \langle\phi_n|\hat{V}|\phi_n\rangle\langle\phi_k|\psi_n^{(1)}\rangle}{\varepsilon_n - \varepsilon_k}|\phi_k\rangle \tag{8.21}$$

For $|\psi_n^{(1)}\rangle$, in applying the formula, expression (8.15) must still be inserted. This results in the following approximation solution for $|\psi_n\rangle$ in second order:

$$|\psi_n\rangle \approx |\phi_n\rangle + \lambda|\psi_n^{(1)}\rangle + \lambda^2\sum_{k\neq n}\frac{\langle\phi_k|\hat{V}|\psi_n^{(1)}\rangle - \langle\phi_n|\hat{V}|\phi_n\rangle\langle\phi_k|\psi_n^{(1)}\rangle}{\varepsilon_n - \varepsilon_k}|\phi_k\rangle \tag{8.22}$$

Example: Perturbed 2×2 Matrix

We consider the Hamiltonian operator $\hat{H}$ of an unspecified two-dimensional system in a complete basis $\{|\phi_i\rangle\}$:

$$\hat{H} \xrightarrow{\{|\phi_i\rangle\}} \begin{pmatrix} A_1 & W \\ W^* & A_2 \end{pmatrix} = \underbrace{\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}}_{H_0} + \underbrace{\begin{pmatrix} 0 & W \\ W^* & 0 \end{pmatrix}}_{V} \tag{8.23}$$

It is supposed to hold that $A_1 < A_2$ and $W \in \mathbb{C}$! By decomposing the matrix into diagonal and off-diagonal elements, it is evident that the two eigenvalues of the unperturbed system $\hat{H}_0$ are given by $\varepsilon_1 = A_1$ and $\varepsilon_2 = A_2$.

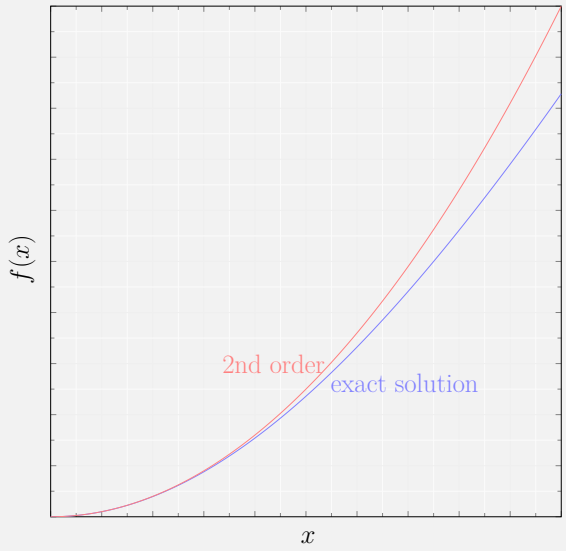


Fig. 47: For small $|W|$ or large distances between the energy levels $A_1 \gg A_2$, the second-order correction and the exact solution agree very well.

According to (8.8), this is simultaneously the estimate for the eigenenergies of the perturbed system in zero order, which we want to write here as $E_1^{(0)}$ and $E_2^{(0)}$. According to (8.8), the corresponding eigenvectors of the unperturbed system, $|\phi_1\rangle$ and $|\phi_2\rangle$, are simultaneously the estimation of the eigenvectors of the perturbed system in zero order, which we want to write here as $|\psi_1^{(0)}\rangle$ and $|\psi_2^{(0)}\rangle$. It follows:

$$\begin{aligned} |\psi_1^{(0)}\rangle = |\phi_1\rangle \xrightarrow{\{|\phi_i\rangle\}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{zu} \quad \varepsilon_1 = A_1 \\ |\psi_2^{(0)}\rangle = |\phi_2\rangle \xrightarrow{\{|\phi_i\rangle\}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{zu} \quad \varepsilon_2 = A_2 \end{aligned} \tag{8.24}$$

Here, we want to apply perturbation theory up to the second order, and thus can accordingly infer from (8.5): $E_i \approx E_i^{(0)} + E_i^{(1)} + E_i^{(2)}$ (the order parameter λ is evidently equal to 1 in our example according to (8.23)). But what do the eigenvalues E_1 and E_2 of the entire system $\hat{H}$ look like when only a small perturbation is present? For the first-order correction $E_i^{(1)}$, we obtain:

$$\begin{aligned} E_1^{(1)} \overset{(8.12)}{=} \langle\phi_1|\hat{V}|\phi_1\rangle \overset{(8.24)}{=} \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & W \\ W^* & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0 \\ E_2^{(1)} \overset{(8.12)}{=} \langle\phi_2|\hat{V}|\phi_2\rangle \overset{(8.24)}{=} \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & W \\ W^* & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0 \end{aligned} \tag{8.25}$$

The first-order correction vanishes! We will directly see that this is not the case for the second-order correction – the following result is found:

$$\begin{aligned} E_1^{(2)} \overset{(8.18)}{=} \frac{|\langle\phi_2|\hat{V}|\phi_1\rangle|^2}{\varepsilon_1 - \varepsilon_2} \overset{(8.24)}{=} \frac{1}{A_1 - A_2} \left| \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & W \\ W^* & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right|^2 = \frac{|W|^2}{A_1 - A_2} \\ E_2^{(2)} \overset{(8.18)}{=} \frac{|\langle\phi_1|\hat{V}|\phi_2\rangle|^2}{\varepsilon_2 - \varepsilon_1} \overset{(8.24)}{=} \frac{1}{A_2 - A_1} \left| \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & W \\ W^* & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right|^2 = \frac{|W|^2}{A_2 - A_1} \end{aligned} \tag{8.26}$$

By summing the results (8.24), (8.25), and (8.26), we obtain the energy eigenvalues of $\hat{H}$ according to second-order Rayleigh-Schrödinger perturbation theory:

$$E_1 \approx A_1 - \frac{|W|^2}{A_2 - A_1} \quad \text{and} \quad E_2 \approx A_2 + \frac{|W|^2}{A_2 - A_1} \tag{8.27}$$

Exact Solution The simple form of the perturbed Hamiltonian $\hat{H}$ indeed also allows for an exact solution by directly solving the eigenvalue problem $\det|\hat{H} - \mathbb{1}E| = 0$:

$$0 = \det|H - \mathbb{1}E| = (A_1 - E)(A_2 - E) - |W|^2 = E^2 - E(A_1 + A_2) + A_1 A_2 - |W|^2$$

Solving this quadratic equation leads us to the following expression, which we can expand up to the second order using $f(x) = \sqrt{1 + x^2} \approx 1 + x^2/2$:

$$\begin{aligned} E_{1,2} &= \frac{A_1 + A_2}{2} \pm \frac{A_1 - A_2}{2}\sqrt{1 + \left(\frac{2|W|}{A_1 - A_2}\right)^2} \approx \\ &\approx \frac{A_1 + A_2}{2} \pm \frac{A_1 - A_2}{2}\left(1 + \frac{2|W|^2}{(A_1 - A_2)^2}\right) \end{aligned} \tag{8.28}$$

The exact solution developed up to the second order ((8.28) (depending on whether one evaluates it for E_1 or E_2) is completely identical with the result (8.27) of the perturbation calculation! In the figure below, the approximation for $f(x)$ with $x = 2|W|/(A_1 - A_2)$ is graphically illustrated.

9 Concepts of Quantum Theory

In this chapter, we deal with advanced concepts, or rather, we try to understand very fundamental principles of quantum theory. We also delve into specific technical realizations of quantum mechanical phenomena, such as the quantum computer or quantum cryptography.

9.1 Series of Stern-Gerlach Apparatuses

In Chapter 7, we got to know the Stern-Gerlach apparatus. In the following, we want to illustrate some of the fundamental properties of quantum mechanics using multiple Stern-Gerlach apparatuses arranged in sequence.

Simple Stern-Gerlach Apparatus In Figure 48, the following experimental setup is sketched: A stream of uncharged spin-$\frac{1}{2}$ particles moving in the x direction passes through a Stern-Gerlach apparatus with an arbitrary magnetic field angle ϑ (measured in the yz-plane, relative to the positive z direction). We assume that the particles emitted from the source (on the left) are polarized in the z direction and are in an initial state $|\psi\rangle = |\uparrow\rangle$. The particles leaving the Stern-Gerlach apparatus on the lower path and thus in the state $|\downarrow_\vartheta\rangle$ (with the energy eigenvalue $\lambda_\downarrow = -\hbar/2$) are blocked. The particles that leave the Stern-Gerlach apparatus on the upper path and thus are in the state $|\uparrow_\vartheta\rangle$ (with $\lambda_\uparrow = +\hbar/2$) can pass and are detected. What is the probability that a particle hits our detector?

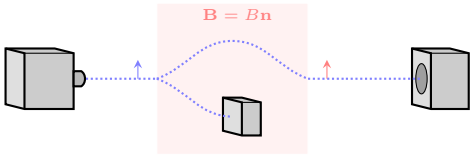


Fig. 48: The number of particles originally prepared in the state $|\uparrow\rangle$ that leave the Stern-Gerlach apparatus, set at an arbitrary angle ϑ, in the state $|\uparrow_\vartheta\rangle$ (upper path) should be determined. The lower path (state $|\downarrow_\vartheta\rangle$) is blocked.

The particles leaving the Stern-Gerlach apparatus are in the following state according to (7.55) (we assume that $\varphi = 0$):

$$|\uparrow_\vartheta\rangle \overset{(7.55)}{=} \cos(\vartheta/2)\,|\uparrow\rangle + \sin(\vartheta/2)\,|\downarrow\rangle \quad \text{for} \quad \lambda_\uparrow = +\hbar/2 \tag{9.1}$$

The probability that the measurement of the particle with initial state $|\uparrow\rangle$ subsequently results in the state $|\uparrow_\vartheta\rangle$ is:

$$\begin{aligned} P(\uparrow,\uparrow_\vartheta) = |\,\langle\uparrow\,|\,\uparrow_\vartheta\rangle\,|^2 &\overset{(9.1)}{=} |\,\langle\uparrow|\,(\cos(\vartheta/2)\,|\uparrow\rangle + \sin(\vartheta/2)\,|\downarrow\rangle)\,|^2 = \\ &= |\cos(\vartheta/2)\langle\uparrow\,|\,\uparrow\rangle + \sin(\vartheta/2)\langle\uparrow\,|\,\downarrow\rangle|^2 = \cos^2(\vartheta/2) \end{aligned}$$

Naturally, due to particle conservation, with $P(\uparrow,\downarrow_\vartheta) = \sin^2(\vartheta/2)$ (as the probability of obtaining the measurement value $-\hbar/2$), the relation $P(\uparrow,\uparrow_\vartheta) + P(\uparrow,\downarrow_\vartheta) = \cos^2(\vartheta/2) + \sin^2(\vartheta/2) = 1$ must be fulfilled. As a reminder: The concept of probability here should be understood such that we can specify the following expected value $\langle N'\rangle$ in a multitude of equivalent experiments conducted with individual particles (i.e., the source generates a beam of a total of N particles):

$$\langle N'\rangle = N\cos^2(\vartheta/2) \tag{9.2}$$

Double Stern-Gerlach Apparatus Let's now consider the more complicated case where two Stern-Gerlach apparatuses with measurement angles ϑ and ϑ', respectively, are placed in sequence. Again, it should be the case that the original particle beam is polarized in the z direction ($|\psi\rangle = |\uparrow\rangle$). From both apparatuses, *only* the states $|\uparrow_\vartheta\rangle$ and $|\uparrow_{\vartheta'}\rangle$, corresponding to the $+\hbar/2$ eigenvalues, are allowed through (see Figure 49). We have:

$$\begin{aligned} |\uparrow_\vartheta\rangle &= \cos(\vartheta/2)\,|\uparrow\rangle + \sin(\vartheta/2)\,|\downarrow\rangle \quad \text{for} \quad \lambda_\uparrow = +\hbar/2 \\ |\uparrow_{\vartheta'}\rangle &= \cos(\vartheta'/2)\,|\uparrow\rangle + \sin(\vartheta'/2)\,|\downarrow\rangle \quad \text{for} \quad \lambda'_\uparrow = +\hbar/2 \end{aligned} \tag{9.3}$$

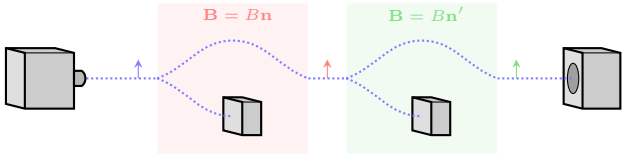


Fig. 49: Two Stern-Gerlach apparatuses connected in series with different magnetic field orientations ϑ and ϑ'. In both apparatuses, only the states corresponding to the eigenvalues $+\hbar/2$ are allowed through, and the states corresponding to the eigenvalues $-\hbar/2$ are each blocked.

The probability that the state $|\psi\rangle = |\uparrow\rangle$ transitions to the state $|\uparrow_\vartheta\rangle$ after the first Stern-Gerlach apparatus is (as already calculated) given by:

$$P(\uparrow, \uparrow_\vartheta) = \langle \hat{P}_{\uparrow\vartheta} \rangle = |\,\langle\uparrow \mid \uparrow_\vartheta\rangle\,|^2 = \cos^2(\vartheta/2)$$

The projector $\hat{P}_{\uparrow\vartheta}$ ensures that every particle is in the state $|\uparrow_\vartheta\rangle$ after passing the first Stern-Gerlach apparatus (the other particle beam is blocked). We now want to calculate the probability that the particle is in the state $|\uparrow_{\vartheta'}\rangle$ after the second apparatus:

$$\begin{aligned} P'(\uparrow_\vartheta, \uparrow_{\vartheta'}) &= |\,\langle\uparrow_\vartheta \mid \uparrow_{\vartheta'}\rangle\,|^2 \overset{(9.3)}{=} \\ &= |\,(\cos(\vartheta/2)\,\langle\uparrow| + \sin(\vartheta/2)\,\langle\downarrow|)\,(\cos(\vartheta'/2)\,|\uparrow\rangle + \sin(\vartheta'/2)\,|\downarrow\rangle)\,|^2 = \\ &= |\,\cos(\vartheta/2)\cos(\vartheta'/2)\langle\uparrow \mid \uparrow\rangle + \sin(\vartheta/2)\sin(\vartheta'/2)\langle\downarrow \mid \downarrow\rangle + \\ &\quad + \cos(\vartheta/2)\sin(\vartheta'/2)\langle\uparrow \mid \downarrow\rangle + \sin(\vartheta/2)\cos(\vartheta'/2)\langle\downarrow \mid \uparrow\rangle\,|^2 = \\ &= [\cos(\vartheta/2)\cos(\vartheta'/2) + \sin(\vartheta/2)\sin(\vartheta'/2)]^2 = \\ &= \cos^2((\vartheta' - \vartheta)/2) \end{aligned} \tag{9.4}$$

For general angles ϑ and ϑ', we thus expect that with initially N particles, ultimately $\langle N'' \rangle$ particles in the state $|\uparrow_{\vartheta'}\rangle$ will hit the detector:

$$\langle N'' \rangle = N \cdot P(\uparrow, \uparrow_\vartheta) P'(\uparrow_\vartheta, \uparrow_{\vartheta'}) \overset{(9.2,9.5)}{=} N \cos^2\left(\frac{\vartheta}{2}\right) \cos^2\left(\frac{\vartheta' - \vartheta}{2}\right) \tag{9.5}$$

Analogous to a series of optical polarization filters, we can set the Stern-Gerlach apparatuses so that, for example, at the end, we measure *no* particles anymore. In the case that the first Stern-Gerlach apparatus allows the incoming beam completely through ($\vartheta = 0$) and the second apparatus is aligned exactly orthogonal to it ($\vartheta' = \pi$), such an extinction would occur. However, if the first Stern-Gerlach apparatus is rotated to $\vartheta = \pi/2$ (thus letting only half of the incoming beam through), while the second apparatus remains unchanged ($\vartheta' = \pi$), the total probability $P = 0.25$ for transmission results. Interestingly, an additional filter can thus lead to greater transmission!

We can understand the result as the successive realization of a final state; the "path" to the final state is then divided into "partial paths": The probability of measuring the final state corresponds to the product of the probabilities of the individual part steps.

9.2 Axioms of Quantum Theory

Let's now attempt to formulate some of the rules of quantum mechanics for calculating measurable predictions in terms of axioms:

1.) **Information Content**: The state of a physical system is fully characterized by a state vector $|\psi\rangle$. All information about a system is thus contained in the state vector. However, this applies only to *pure states*; for *mixed states*, the entire information about the system is contained in the density operator $\hat{\rho} = |\psi\rangle \langle\psi|$ (see `Quantum Theory II`).

2.) **Observable**: A physical observable A, i.e., a real measurable quantity, always corresponds to a hermitian operator $\hat{A} = \hat{A}^\dagger$. Such operators act on state vectors $\hat{A}|\psi\rangle$ to access the information contained therein.

3.) **Measurement Value**: The eigenvalues a_n of an operator $\hat{A}$ are the *only* possible measurement values of the observable A. The eigenvalue equation $\hat{A}|\phi_n\rangle = a_n|\phi_n\rangle$ follows, where we call $|\phi_n\rangle$ an eigenstate of $\hat{A}$ with eigenvalue a_n.

4.) **Measurement Probability**: Given an arbitrary state $|\psi\rangle$, the probability of obtaining the measurement value $\hat{a}_n$ of the observable A can be given via the projection onto the corresponding eigenstate $p_n = |\langle\phi_n|\psi\rangle|^2 = \langle\psi|\hat{P}_n|\psi\rangle$. The state $|\psi\rangle$ is mapped to a state $|\phi_n\rangle$ using the projection operator $\hat{P}_n = |\phi_n\rangle\langle\phi_n|$. A measurement leads to the "collapse of the wave function" onto $|\phi_n\rangle$.

5.) **Time Evolution**: The time evolution of a state $|\psi(t)\rangle$ is determined by the Schrödinger equation $i\hbar|\dot{\psi}(t)\rangle = \hat{H}|\psi(t)\rangle$.

6.) **Classical Analogue**: Operators with classical analogues are determined by canonical quantization by replacing the canonical coordinates r_i and p_i with complementary operators that satisfy the commutation relation $[\hat{r}_i, \hat{p}_j] = i\hbar\delta_{ij}$. Operators whose classical analogues are functions of r_i and p_i are symmetrized in the order of the non-commuting operators.

Calculation Rules for Probabilities If we assume a completely prepared system, the following calculation rules for probabilities can be noted (according to Dietrich Grau, *Übungsaufgaben zur Quantentheorie*, V5.52, 2020):

(1) **Final States**: The probability for a "final state" is given by the absolute square of a complex number; namely, the *probability amplitude* of the final state. A final state refers to the distinguishable, alternative outcomes of the experiment under consideration.

(2) **Alternative Realizations**: If there are several, within the experiment, fundamentally indistinguishable alternative ways ("virtual paths") to achieve a final state, the probability amplitude for reaching that final state is the sum of the probability amplitudes for the relevant virtual paths (*interference*).

(3) **Successive Alternative Realizations**: If one of the virtual paths consists of several "sub-paths", the probability amplitude for the entire virtual path is the product of the probability amplitudes for the sub-paths.

(4) **Events**: If multiple final states are grouped into a "event", the probability for this event is given by the sum of the probabilities for the respective final states.

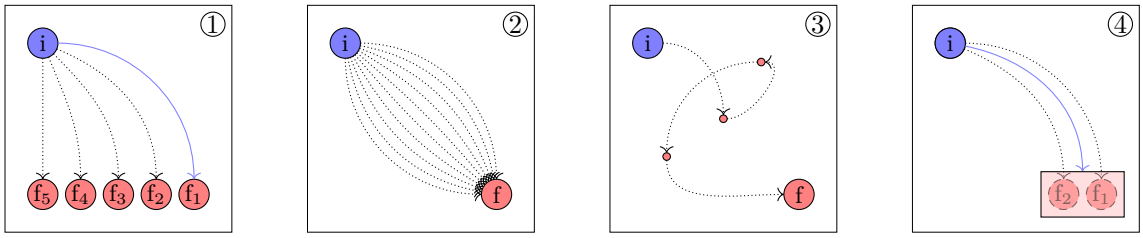


Fig. 50: Schematic representation for calculating probabilities of experiment outcomes: (1) different final states, (2) alternative realizations, (3) successive realization, and (4) events from multiple final states.

9.3 The Measurement Process

The measurement process plays a central role in quantum mechanics, since we not only intervene explicitly in a quantum system but also understand the measurement process as a means of preparing the system. In the following, we will discuss different terms for quantum states $|\psi\rangle$ (which we will revisit in the `Quantum Theory II` in the context of the density operator) and also touch on the most frequently "killed" animal in the history of physics – Schrödinger's cat.

9.3.1 Separable and Entangled States

Consider two systems $\mathcal{H}_A$ and $\mathcal{H}_B$, which interact with each other and can be described together in $\mathcal{H}_{A+B}$. For the individual systems, we find the corresponding observables $\hat{A}$ and $\hat{B}$, as well as a complete orthonormal basis $\{|a_i\rangle\}$ and $\{|b_i\rangle\}$, respectively. These bases are however *not* the eigenbases of the operators $\hat{A}$ and $\hat{B}$; it should only be that $\hat{A}$ acts in $\mathcal{H}_A$, while $\hat{B}$ acts exclusively in $\mathcal{H}_B$. If a state $|\psi\rangle$ from $\mathcal{H}_{A+B}$ can be written as follows, we call this state *separable*:

$$|\psi\rangle = \left(\sum_i \alpha_i |a_i\rangle\right) \otimes \left(\sum_j \beta_j |b_j\rangle\right) = |\psi_a\rangle \otimes |\psi_b\rangle \quad \text{with} \quad \alpha_i, \beta_j \in \mathbb{C} \tag{9.6}$$

For the coefficients of such a *product state* of the linear superpositions of the states $|\psi_a\rangle$ and $|\psi_b\rangle$, it holds that $\sum_i |\alpha_i|^2 = \sum_j |\beta_j|^2 = 1$. If such a representation as a sum of product states ($|a_i\rangle \otimes |b_j\rangle \equiv |a_i b_j\rangle$) is not possible, we call the state *entangled*. Such a state is given in full generality as follows:

$$|\phi\rangle = \sum_{ij} \gamma_{ij} |a_i\rangle \otimes |b_j\rangle \quad \text{with} \quad \gamma_j \in \mathbb{C} \tag{9.7}$$

The state $|\phi\rangle$ is *not separable*. The separability criterion is only met if all γ_{ij} can be represented as $\gamma_{ij} = \alpha_i \beta_j$, making an explicit separation between $|\psi_a\rangle$ and $|\psi_b\rangle$ possible. A clear separation of γ_{ij} into α_i and β_j (where $i, j = 1, \ldots, N$) is in principle only possible by solving a N^2-large system of equations. Whether this solution exists can be verified by writing all coefficients γ_{ij} in the form of a matrix and then checking if $\text{rank}(\gamma_{ij}) = 1$ holds. If instead $\text{rank}(\gamma_{ij}) > 1$, then the state is entangled.
It follows that in the case of separability, the matrix $\gamma_{ij} = \alpha_i \beta_j$ corresponds to the outer product of the vectors, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_N)^\intercal$ and $\boldsymbol{\beta} = (\beta_1, \beta_2, \cdots, \beta_N)^\intercal$, and a matrix formed by $\boldsymbol{\alpha} \otimes \boldsymbol{\beta}$ must always have a rank of 1.
This method is not limited to two-state systems and can be generalized to arbitrarily large systems ($N_\mathcal{H} > 2$).

Example: Separability of a Two-Particle System

We consider a system of two interacting electrons $i = 1, 2$ with spins $\{|\uparrow\rangle_i, |\downarrow\rangle_i\}$. To decide whether two different states are separable or entangled, we examine the rank of the matrix formed from the coefficients γ_{ij}. First, we generally compare a separable state in a representation like (9.6) with a representation like (9.7):

$$\begin{aligned} |\psi\rangle &= (\alpha_\uparrow |\uparrow\rangle_1 + \alpha_\downarrow |\downarrow\rangle_1) \otimes (\beta_\uparrow |\uparrow\rangle_2 + \beta_\downarrow |\downarrow\rangle_2) = \\ &= \alpha_\uparrow\beta_\uparrow |\uparrow\uparrow\rangle + \alpha_\downarrow\beta_\downarrow |\downarrow\downarrow\rangle + \alpha_\uparrow\beta_\downarrow |\uparrow\downarrow\rangle + \alpha_\downarrow\beta_\uparrow |\downarrow\uparrow\rangle \\ |\phi\rangle &= \gamma_{\uparrow\uparrow} |\uparrow\uparrow\rangle + \gamma_{\downarrow\downarrow} |\downarrow\downarrow\rangle + \gamma_{\uparrow\downarrow} |\uparrow\downarrow\rangle + \gamma_{\downarrow\uparrow} |\downarrow\uparrow\rangle \end{aligned}$$

A comparison of $|\psi\rangle$ and $|\phi\rangle$ should allow a clear assignment of the coefficients γ_{ij} with α_i and β_j in the form $\gamma_{ij} = \alpha_i\beta_j$ for a separable state. Summarized as a matrix, we can better understand the decomposition into the outer product:

$$\begin{pmatrix} \gamma_{\uparrow\uparrow} & \gamma_{\uparrow\downarrow} \\ \gamma_{\downarrow\uparrow} & \gamma_{\downarrow\downarrow} \end{pmatrix} = \begin{pmatrix} \alpha_\uparrow\beta_\uparrow & \alpha_\uparrow\beta_\downarrow \\ \alpha_\downarrow\beta_\uparrow & \alpha_\downarrow\beta_\downarrow \end{pmatrix} = \begin{pmatrix} \alpha_\uparrow \\ \alpha_\downarrow \end{pmatrix} \begin{pmatrix} \beta_\uparrow & \beta_\downarrow \end{pmatrix} \tag{9.8}$$

At this point, it also becomes clear why a matrix formed by $\boldsymbol{\alpha} \otimes \boldsymbol{\beta}$ must have a rank of 1: The first and second columns are linearly dependent on each other. The first column corresponds to the second column multiplied by a factor $\beta_\uparrow/\beta_\downarrow$.

Separable State As a simple example, consider the following state:

$$|\psi_1\rangle = |\uparrow\rangle_1 \otimes |\uparrow\rangle_2 \equiv |\uparrow\uparrow\rangle$$

In the context of (9.6), $|\psi_1\rangle = 1\,|\uparrow\uparrow\rangle + 0\,|\downarrow\downarrow\rangle + 0\,|\uparrow\downarrow\rangle + 0\,|\downarrow\uparrow\rangle$ applies. The coefficients matrix according to (9.7) has rank one and can therefore be written as an outer product:

$$\text{Rank} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 1 \implies \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix}$$

The state is thus separable, only the coefficients $\alpha_\uparrow = \beta_\uparrow = 1$ and as a direct consequence $\gamma_{\uparrow\uparrow} = 1$ are non-zero.

Entangled State Now consider the following state:

$$|\psi_2\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$$

The coefficients matrix according to (9.7) has rank two and therefore cannot be written as the outer product of two coefficient vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$:

$$\text{Rank} \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 \end{pmatrix} = 2 \implies \text{entangled state}$$

We infer that the state $|\psi_2\rangle$ is not separable but entangled.

Expectation Values for Separable States We now want to determine the expectation value of $\hat{A}$ for a general separable state according to (9.6) in the system $\mathcal{H}_{A+B}$, where both $\alpha_i \neq 0$ and $\beta_j \neq 0$ should hold. We need to note that $\hat{A}$ acts *only* on states from $\mathcal{H}_A$! The construction of

$\langle \hat{A} \rangle$ leads us to:

$$\begin{aligned}
\langle \psi | \hat{A} | \psi \rangle &= \Big(\sum_{ij} \alpha_i^* \beta_j^* \langle a_i | \otimes \langle b_j | \Big) \hat{A} \Big(\sum_{kl} \alpha_k \beta_l | a_k \rangle \otimes | b_l \rangle \Big) = \\
&= \sum_{jl} \beta_j^* \beta_l \langle b_j | b_l \rangle \sum_{ik} \alpha_i^* \alpha_k \langle a_i | \hat{A} | a_k \rangle = \\
&= \sum_{jl} \beta_j^* \beta_l \delta_{jl} \Big(\sum_i \alpha_i^* \alpha_i \langle a_i | \hat{A} | a_i \rangle + \sum_k \sum_{i \neq k} \alpha_i^* \alpha_k \langle a_i | \hat{A} | a_k \rangle \Big) = \Big| \sum_j |\beta_j|^2 = 1 \\
&= \sum_i |\alpha_i|^2 \langle a_i | \hat{A} | a_i \rangle + \sum_k \sum_{i \neq k} \alpha_i^* \alpha_k \langle a_i | \hat{A} | a_k \rangle
\end{aligned} \tag{9.9}$$

We do not know the explicit action of $\hat{A}$ on any arbitrary state $|a_i\rangle$ in this general example, so $\langle a_i | \hat{A} | a_k \rangle$ must remain the final result here. However, we know that $\hat{A}$ does not act on $|b_j\rangle$, which allowed us to pull the scalar product $\langle b_j | b_l \rangle$ to the front. The calculation of the expectation value in (9.9) leads us to diagonal terms ($i = k$), which can be interpreted as classical probabilities, as well as off-diagonal terms ($i \neq k$), which we call *interference terms*. Due to these terms, we speak of a *coherent superposition* – the coherence of the states allows the formation of interferences in analogy to optics.

Expectation Values for Entangled States However, we have already recognized that not every state is separable. Therefore, let's calculate the expectation value $\langle \hat{A} \rangle$ again for a general entangled state as described in (9.7):

$$\begin{aligned}
\langle \phi | \hat{A} | \phi \rangle &= \Big(\sum_{ij} \gamma_{ij}^* \langle a_i b_j | \Big) \hat{A} \Big(\sum_{kl} \gamma_{kl} | a_k b_l \rangle \Big) = \\
&= \sum_{ij} \sum_{kl} \gamma_{ij}^* \gamma_{kl} \langle b_j | b_l \rangle \langle a_i | \hat{A} | a_k \rangle = \\
&= \sum_{ij} \sum_{kl} \gamma_{ij}^* \gamma_{kl} \delta_{jl} \langle a_i | \hat{A} | a_k \rangle = \\
&= \sum_{ij} \sum_{k} \gamma_{kj} \gamma_{ij}^* \langle a_i | \hat{A} | a_k \rangle = \Big| \gamma_{kj} = \langle a_k b_j | \phi \rangle \,,\ \gamma_{ij}^* = \langle \phi | a_i b_j \rangle \\
&= \sum_{ij} \sum_{k} \langle a_k b_j | \phi \rangle \langle \phi | a_i b_j \rangle \langle a_i | \hat{A} | a_k \rangle = \\
&= \sum_{ij} \sum_{k} \langle a_k | \langle b_j | \phi \rangle \langle \phi | b_j \rangle | a_i \rangle \langle a_i | \hat{A} | a_k \rangle = \Big| \sum_i |a_i\rangle \langle a_i| = \mathbb{1} \\
&= \sum_{j} \sum_{k} \langle a_k | \langle b_j | \phi \rangle \langle \phi | b_j \rangle \hat{A} | a_k \rangle = \\
&= \mathrm{Tr}_A \Big\{ \mathrm{Tr}_B \{ | \phi \rangle \langle \phi | \} \hat{A} \Big\}
\end{aligned} \tag{9.10}$$

No interference term appears here, i.e., it is an *incoherent superposition* or a *statistical mixture*. We will learn about the formalism of the density operator $\hat{\rho}$ in `Quantum Theory II`, which can sensibly describe statistical mixtures.

Example: Coherence of a Two-Particle System

Let's discuss the coherence behavior of a two-particle system using the states $|\psi_1\rangle$ and $|\psi_2\rangle$ more precisely. Let $|\psi_1\rangle$ be the most general two-particle product state with the realizable states $\{|\uparrow_a\rangle, |\downarrow_a\rangle\}$ for the first particle and $\{|\uparrow_b\rangle, |\downarrow_b\rangle\}$ for the second particle. We find the following wavefunction according to (9.6):

$$|\psi_1\rangle = (\alpha_1 |\uparrow_a\rangle + \alpha_2 |\downarrow_a\rangle) \otimes (\beta_1 |\uparrow_b\rangle + \beta_2 |\downarrow_b\rangle)$$

We calculate the expectation value of the observable $\hat{A}$ of the system $\mathcal{H}_A$. However, it must be noted that $\hat{A}$ can only act on the first particle. It follows for the quantity $\langle\hat{A}\rangle$:

$$\begin{aligned}
\langle\hat{A}\rangle &= \langle\psi_1|\hat{A}|\psi_1\rangle = \\
&= [(\alpha_1^* \langle\uparrow_a| + \alpha_2^* \langle\downarrow_a|) \otimes (\beta_1^* \langle\uparrow_b| + \beta_2^* \langle\downarrow_b|)] \hat{A} [(\alpha_1 |\uparrow_a\rangle + \alpha_2 |\downarrow_a\rangle) \otimes (\beta_1 |\uparrow_b\rangle + \beta_2 |\downarrow_b\rangle)] = \\
&= (\alpha_1^* \langle\uparrow_a| + \alpha_2^* \langle\downarrow_a|) \hat{A} (\alpha_1 |\uparrow_a\rangle + \alpha_2 |\downarrow_a\rangle) \cdot (\beta_1^* \langle\uparrow_b| + \beta_2^* \langle\downarrow_b|) (\beta_1 |\uparrow_b\rangle + \beta_2 |\downarrow_b\rangle) = \\
&= \Big(\alpha_1^*\alpha_1 \langle\uparrow_a |\hat{A}| \uparrow_a\rangle + \alpha_2^*\alpha_2 \langle\downarrow_a |\hat{A}| \downarrow_a\rangle + \alpha_1^*\alpha_2 \langle\uparrow_a |\hat{A}| \downarrow_a\rangle + \alpha_2^*\alpha_1 \langle\downarrow_a |\hat{A}| \uparrow_a\rangle\Big) \cdot \\
&\quad (\beta_1^*\beta_1 \langle\uparrow_b | \uparrow_b\rangle + \beta_2^*\beta_2 \langle\downarrow_b | \downarrow_b\rangle + \beta_1^*\beta_2 \langle\uparrow_b | \downarrow_b\rangle + \beta_2^*\beta_1 \langle\downarrow_b | \uparrow_b\rangle) = \\
&= \Big(|\alpha_1|^2 \langle\uparrow_a |\hat{A}| \uparrow_a\rangle + |\alpha_2|^2 \langle\downarrow_a |\hat{A}| \downarrow_a\rangle + \alpha_1^*\alpha_2 \langle\uparrow_a |\hat{A}| \downarrow_a\rangle + \alpha_2^*\alpha_1 \langle\downarrow_a |\hat{A}| \uparrow_a\rangle\Big) \cdot \\
&\quad \Big(|\beta_1|^2 + |\beta_2|^2\Big) = \\
&= |\alpha_1|^2 \langle\uparrow_a |\hat{A}| \uparrow_a\rangle + |\alpha_2|^2 \langle\downarrow_a |\hat{A}| \downarrow_a\rangle + \underbrace{\alpha_1^*\alpha_2 \langle\uparrow_a |\hat{A}| \downarrow_a\rangle + \alpha_2^*\alpha_1 \langle\downarrow_a |\hat{A}| \uparrow_a\rangle}_{\text{Interference terms}}
\end{aligned}$$

The result corresponds exactly to the previously derived equation (9.9), and we recognize that the product state generates interference terms. But how does this look in the case of an entangled state $|\psi_2\rangle$? We define the following wavefunction:

$$|\psi_2\rangle = \alpha |\uparrow_a\rangle \otimes |\downarrow_b\rangle + \beta |\downarrow_a\rangle \otimes |\uparrow_b\rangle \equiv \alpha |\uparrow\downarrow\rangle + \beta |\downarrow\uparrow\rangle$$

The rank of the coefficients matrix is 2; the state is therefore not separable. We calculate the expectation value of the same observable $\hat{A}$ as before, which acts only on the first particle.

$$\begin{aligned}
\langle\hat{A}\rangle &= \langle\psi_2|\hat{A}|\psi_2\rangle = \\
&= (\alpha^* \langle\uparrow\downarrow| + \beta^* \langle\downarrow\uparrow|) \hat{A} (\alpha |\uparrow\downarrow\rangle + \beta |\downarrow\uparrow\rangle) = \\
&= \alpha^*\alpha \langle\uparrow_a |\hat{A}| \uparrow_a\rangle \langle\downarrow_b | \downarrow_b\rangle + \beta^*\beta \langle\downarrow_a |\hat{A}| \downarrow_a\rangle \langle\uparrow_b | \uparrow_b\rangle + \\
&\quad + \alpha^*\beta \langle\uparrow_a |\hat{A}| \downarrow_a\rangle \langle\downarrow_b | \uparrow_b\rangle + \beta^*\alpha \langle\downarrow_a |\hat{A}| \uparrow_a\rangle \langle\uparrow_b | \downarrow_b\rangle = \\
&= |\alpha|^2 \langle\uparrow_a |\hat{A}| \uparrow_a\rangle + |\beta|^2 \langle\downarrow_b |\hat{A}| \downarrow_b\rangle
\end{aligned}$$

Obviously, there are no off-diagonal elements; a state completely entangled with the system $\mathcal{H}_B$ cannot interfere in the system $\mathcal{H}_A$.

9.3.2 Schrödinger's Cat and the Limit of Quantum Mechanics

Motivation: An Explanation for the Behavior of Macroscopic Objects.

The question arises why in everyday life with macroscopic objects, we cannot observe the same phenomena as with microscopic quantum particles. For example, electrons sent through two closely spaced slits form the known interference pattern of the double-slit experiment (which can be explained through the interference terms described above), while such an effect has never been observed with tennis balls sent through two slits.

If, however, quantum physics is supposed to be a fundamental, universal theory, then it must also be able to describe the behavior of macroscopic objects such as tennis balls in the limiting case – after all, these too are ultimately made of atoms that obey the laws of quantum physics. But how come that the conglomeration of atoms forming two tennis balls apparently never manifests in a product state? An answer to this is provided by the effect of decoherence.

The transition from a product state, which exhibits interference terms during measurement, to an entangled state, which no longer has interference terms during a measurement, is called *decoherence*. Typically, decoherence occurs when a quantum mechanical system interacts with its environment. It leads to an entanglement with the particles of the environment (which are not measured), resulting in the loss of interference properties in the particles of the object to be measured. This means that in everyday macroscopic objects – which inevitably interact with the environment – a measurement can be described using classical probabilities. This limiting factor of quantum states also plays a role in the paradox of *Schrödinger's Cat*. Erwin Schrödinger wrote in 1935 in his essay "The Present Situation in Quantum Mechanics":

> "One can even set up quite ridiculous cases. A cat is penned up in a steel chamber, along with the following diabolical device (which must be secured against direct interference by the cat): In a Geiger counter there is a tiny bit of radioactive substance, so small that perhaps in the course of an hour one of the atoms decays, but also, with equal probability, perhaps none; if it happens, the counter tube discharges and through a relay releases a hammer that shatters a small flask of hydrocyanic acid. If one has left this entire system to itself for an hour, one would say that the cat still lives if meanwhile no atom has decayed. The first atomic decay would have poisoned her. The ψ-function of the entire system would express this by having in it the living and the dead cat (pardon the expression) mixed or smeared out in equal parts. Typical instances of this kind are actually found quite frequently in the valid interpretation of quantum mechanics. These rather naive examples have the disadvantage of allowing an incorrect interpretation which leaps unnoticed across the precarious edge of direct observation. This prevents us from naively accepting a 'blurred model' as a reflection of reality. In itself, it would not contain anything unclear or contradictory. There is a difference between a blurred or out-of-focus photograph and a photograph of clouds and wafts of mist."

The answer to the problem posed by Schrödinger is provided by the phenomenon of decoherence: In macroscopic objects (such as a cat), which consist of a very large number of quantum particles, it becomes increasingly difficult to impossible to shield these from their environment. Interactions inevitably occur between the particles of the cat and the environment. This inevitably leads to a state of decoherence where the observable quantum properties practically disappear. A superposition of a live and dead cat is therefore indeed meaningless because the coherence condition is no longer fulfilled from the outset due to decoherence.

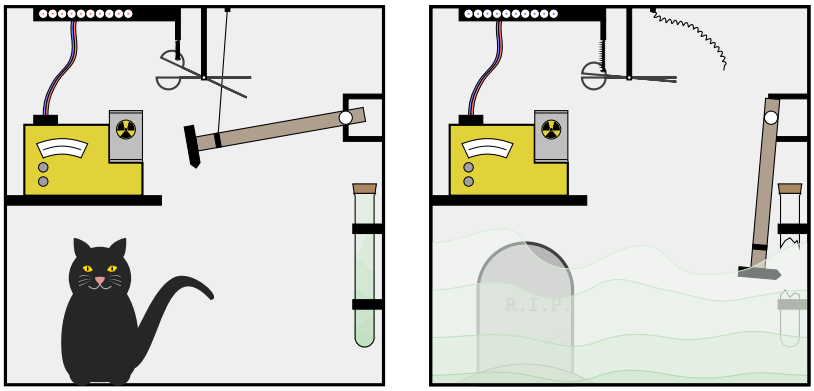


Fig. 51: Schrödinger's hell machine: Only when opening the box can the observer verify whether the cat inside is alive (left) or dead (right).

9.4 Bell's Inequalities and Hidden Variables

Motivation: Completeness of Quantum Physics

Let us imagine the following experiment: A particle source between the physicists Alice and Bob generates two entangled spin-$\frac{1}{2}$ particles, which fly apart in opposite directions, one towards Alice and the other towards Bob. These entangled spin-$\frac{1}{2}$ particles are produced through the spontaneous decay of spin-0 particles – due to the conservation of angular momentum, the spin of the spin-$\frac{1}{2}$ particles must always be opposite. That means, if Alice and Bob measure along the same axis, they will always find an opposite spin, regardless of the measurement direction they have agreed upon. It holds that:

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow\rangle_A |\downarrow\rangle_B - |\downarrow\rangle_A |\uparrow\rangle_B \right) \tag{9.11}$$

This representation is done in the eigenbasis of $\hat{S}_z$; it is immediately apparent that the spin of the particles is anti-correlated in the z-direction. However, it turns out that this is also the case when performing a basis transformation along any other measurement axis – see for example (9.38), where this state was converted into the x-basis. Note: The minus sign in (9.11) is important because the same state with a plus sign is no longer anti-correlated in every direction!

Quantum mechanics now tells us the following: Before the measurement, the spin of both particles is completely indeterminate, no matter in which direction one measures it. In a much more fundamental sense, the spin direction before measurement is not existent at all. If Alice measures the spin of the first particle on the left side of the laboratory in any direction $\mathbf{n}$, she will randomly obtain either the result $|\uparrow_\mathbf{n}\rangle_A$ or $|\downarrow_\mathbf{n}\rangle_A$. Only through this measurement does she define reality and fix the spin of "her" particle to the measurement result. Simultaneously, her measurement causes the second particle, which has by then arrived on the right side of the laboratory, to instantly adopt the opposite state. If, for example, Alice measured $|\uparrow_\mathbf{n}\rangle_A$, Bob can be sure to measure the state $|\downarrow_\mathbf{n}\rangle_B$ – even if his measurement occurs so quickly after Alice's measurement that the information could not have traveled from Alice to Bob at the speed of light. It appears as if Alice's measurement influenced the spin of Bob's particle faster than the speed of light! This seemed to contradict relativity theory to Einstein, and he therefore rejected this notion. He called the described effect "spooky action at a distance".

In 1935, Albert Einstein published, together with Boris Podolsky and Nathan Rosen („EPR“), a much-noted article. In it, the hypothesis is proposed that quantum physics is *incomplete*. Essentially, it concerns the measurement of complementary observables which, according to the Heisenberg uncertainty principle, cannot be measured simultaneously, or not even simultaneously exist as real. EPR took position and momentum as an example; here we want to follow the argument based on the spin measurement on two orthogonal axes, according to the revised version of the EPR experiment by David Bohm.

We again assume the entangled state (9.11). Suppose Alice measures the spin of her particle as $|\uparrow\rangle_A$ in the z-direction, and Bob measures the spin of his (entangled) particle as $|\downarrow_x\rangle_B$ in the x-direction. But what would have happened if Alice had also measured her particle in the x-direction, like Bob? EPR argue: Had Alice also decided on the x-direction, then she would have measured $|\uparrow_x\rangle_A$, because the particles are anti-correlated in every measurement direction. Thus, according to EPR, Alice knows the spin in the z-direction (a measurement she performed), *but she also apparently knows the spin in the x-direction* because she can infer from Bob's measurement how her measurement would have turned out in the x-direction.
This would mean that Alice and Bob circumvented the uncertainty relation, because the operators S_z and S_x do not commute and a simultaneous spin measurement in these orthogonal directions should therefore not be possible according to quantum physics. With this thought experiment, as EPR argue, it would be proven that the spin is very much part of physical reality in both the z-direction as well as in the x-direction.

EPR assume in their argumentation that a valid physical theory must fulfill the following assumptions:

- **Reality assumption**: If, without disturbing a system, a quantity can be predicted with certainty, then there exists an element of physical reality that can be associated with that quantity.
- **Locality assumption**: Based on two systems; if both systems *do not* interact with each other, a measurement on one system does not change the state of the other system.

Since Alice must make the decision about the measurement axis only shortly before measuring the first particle, this measurement can, assuming *locality*, have no disturbing influence on the reality of the second particle. Nevertheless, with her measurement, she can always deduce the spin of the second particle along the chosen measurement axis. From this, EPR conclude that both the spin in the z-direction and the spin in the x-direction must be part of physical reality, and therefore, according to EPR, quantum mechanics is incomplete.

For a long time, these problems were considered unanswerable by physics and seemed more an issue for metaphysics. But then, in 1964, the physicist John Stewart Bell published an article that made experimental verification of this problem possible.

9.4.1 Hidden Variables and Spooky Action at a Distance

Einstein, Podolsky, and Rosen argue in their work that the apparent paradoxes of quantum mechanics could be solved with so-called *hidden variables*. This is to mean the following: Alice and Bob conduct their spin measurements along different measurement axes, namely along the z-axis ($\vartheta = 0$, basis $\{|\uparrow\rangle, |\downarrow\rangle\}$), at a 45° angle ($\vartheta = \pi/4$, basis $\{|\nearrow\rangle, |\swarrow\rangle\}$), and furthermore along the x-axis ($\vartheta = \pi/2$, basis $\{|\rightarrow\rangle, |\leftarrow\rangle\}$).

Each particle already carries with it at its creation in so-called hidden variables $\boldsymbol{\lambda}$ the information within itself of how the measurement outcome will be along each measurement axis. A specific particle could for example have the hidden variables $\boldsymbol{\lambda} = (\downarrow, \nearrow, \rightarrow)$; thereby it would be simultaneously determined that one would get the result $|\uparrow\rangle$ when measuring in the z-direction, the result $|\nearrow\rangle$ when measuring at a 45°-direction, and the result $|\rightarrow\rangle$ when measuring in the x-direction. Also in the entanglement of two particles, there would then be nothing mysterious anymore. The two entangled spin-$\frac{1}{2}$ particles are indeed created at a common point and simply carry since their creation the opposite hidden variables within themselves. If Alice's particle contains the information $\boldsymbol{\lambda}_A = (\downarrow, \nearrow, \rightarrow)$, then Bob's entangled twin particle has initially the opposite information $\boldsymbol{\lambda}_B = (\uparrow, \swarrow, \leftarrow)$. For the correlation of measurement results between Alice and Bob, "spooky action at a distance" is no longer necessary.

9.4.2 Bell's Inequality

We can now consider one of the particles in our experiment (for example Alice's particle) and enter all possible state combinations of hidden variables in Table 1. Alice and Bob each measure with the bases $\{|\uparrow\rangle, |\downarrow\rangle\}$ ($\vartheta = 0$), $\{|\nearrow\rangle, |\swarrow\rangle\}$ ($\vartheta = \pi/4$), and $\{|\rightarrow\rangle, |\leftarrow\rangle\}$ ($\vartheta = \pi/2$). Furthermore, we can arrange the different state combinations in a set-theoretical Venn Diagram – see Figure 52. In this diagram, three circles (A, B, and C) each symbolize the three sets.

- Set A includes all states of the hidden variables where Alice, when measuring in the z-direction, gets the state $|\uparrow\rangle$.
- Set B includes all states of the hidden variables where Alice, when measuring at a 45°-direction, gets the state $|\nearrow\rangle$.
- Set C includes all states of the hidden variables where Alice, when measuring in the x-direction, gets the state $|\rightarrow\rangle$.

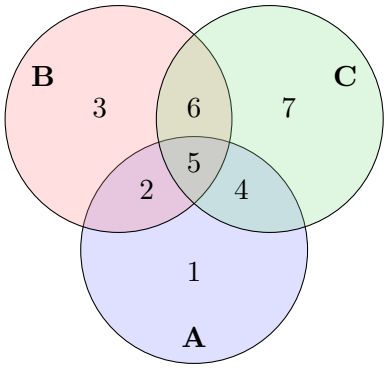


Fig. 52: Venn Diagram for a better understanding of Bell's Inequality. It is assumed that the measurement results are predetermined by hidden variables. In set A are all particles whose measurement in the z-direction yields $|\uparrow\rangle$, in set B are all particles whose measurement in a 45°-direction yields $|\nearrow\rangle$, and in set C are all particles whose measurement in the x-direction yields $|\rightarrow\rangle$.

A particle in the state with the hidden variables $\boldsymbol{\lambda}_A = (\uparrow, \swarrow, \leftarrow)$ belongs, for example, only to set A, but not to set B and C. It thus lies in area 1 of the Venn Diagram. A particle in the state $\boldsymbol{\lambda}_A = (\uparrow, \nearrow, \leftarrow)$, on the other hand, belongs to both set A and set B (but not to set C), and thus lies in the overlap area 2 of the Venn Diagram.
If Alice had the possibility of measuring the spin of every particle in all three directions, then with the help of Table 1 she could uniquely assign every measurement result to an area within

the Venn Diagram in Figure 52. But exactly this simultaneous measurement is forbidden by quantum physics. Alice can merely do the following: She can measure the spin of her particle along one axis, and Bob can measure the spin of the entangled particle along another axis.

hidden variables			set			area
			A	B	C	in the Diagram
$\uparrow$	$\nearrow$	$\rightarrow$	×	×	×	5
$\uparrow$	$\nearrow$	$\leftarrow$	×	×		2
$\uparrow$	$\swarrow$	$\rightarrow$	×		×	4
$\uparrow$	$\swarrow$	$\leftarrow$	×			1
$\downarrow$	$\nearrow$	$\rightarrow$		×	×	6
$\downarrow$	$\nearrow$	$\leftarrow$		×		3
$\downarrow$	$\swarrow$	$\rightarrow$			×	7
$\downarrow$	$\swarrow$	$\leftarrow$				–

Table 1: Possible states of the assumed hidden variables in Alice's particle, and their assignment in the Venn Diagram in Figure 52.

Alice measures, for example, along the z-axis and gets the result $|\uparrow\rangle$. So, she assumes that the first entry of her "hidden variable" looks like this: $\boldsymbol{\lambda}_A = (\uparrow, *, *)$. The second and third entries she does not know after her measurement, which we symbolize with a $*$. But if Bob now also conducts a measurement on his entangled particle, for instance in the direction $\vartheta = \pi/4$, and gets the result $|\nearrow\rangle$, then Alice deduces that she would have measured oppositely $|\swarrow\rangle$ if she had chosen the same measurement direction as Bob. She thus knows at least two of the three "hidden variables" of her particle: $\boldsymbol{\lambda}_A = (\uparrow, \swarrow, *)$.

This can be useful! For instance, all states $\boldsymbol{\lambda}_A = (\uparrow, \swarrow, *)$ are contained in the subset "A without B" (in other words, the intersection of A and the inverse set of B; formally: $A \wedge \bar{B}$) in the Venn Diagram, which covers areas 1 and 4. If we restrict ourselves to the case where Alice only measures along $\vartheta = 0$ and $\vartheta = \pi/4$ and Bob measures along $\vartheta = \pi/4$ and $\vartheta = \pi/2$, we are left with the non-trivial areas $B \wedge \bar{C}$ and $A \wedge \bar{C}$, for which we can make similar considerations.

subset	Alice		Bob		Alice's hidden variables	area in the Diagram
	0°	45°	45°	90°		
$A \wedge \bar{B}$	$\lvert\uparrow\rangle_A$		$\lvert\nearrow\rangle_B$		$(\uparrow, \swarrow, *)$	$1 \cup 4$
$B \wedge \bar{C}$		$\lvert\nearrow\rangle_A$		$\lvert\rightarrow\rangle_B$	$(*, \nearrow, \leftarrow)$	$2 \cup 3$
$A \wedge \bar{C}$	$\lvert\uparrow\rangle_A$			$\lvert\rightarrow\rangle_B$	$(\uparrow, *, \leftarrow)$	$1 \cup 2$

Table 2: Joint measurements by Alice and Bob allow the determination of 2 out of 3 (hypothetical) hidden variables of Alice's particle and thus the assignment to certain subareas in the Venn Diagram.

Alice and Bob can conduct a certain number of measurements (for example 1000 measurements) along the axes 0° (Alice) and 45° (Bob), then the same number of measurements along the axes 45° (Alice) and 90° (Bob), and finally the same number of measurements along the axes 0° (Alice) and 90° (Bob). Whenever the measurements agree with the result from Table 2 (i.e., "spin-up" for both measurements in the respective measurement basis), it is noted as a hit for the respective subset. In the end, Alice is able to specify the number of hits N for the respective subsets.

If the hidden variables are real, then each particle, for which Alice has noted a hit, is in principle (in objective reality) either to be assigned to areas 1, 2, 3, or 4 in the Venn Diagram, and there

is an (objective) hit count $N(1)$, $N(2)$, $N(3)$, and $N(4)$ for each area. Since these numbers can only be positive, it must trivially hold:

$$N(1) + N(2) + N(3) + N(4) \geq N(1) + N(2) \tag{9.12}$$

Alice doesn't know the individual values, but based on Table 2, she can easily conclude:

$$\begin{aligned} N(A \wedge \bar{B}) &= N(1) + N(4) \\ N(B \wedge \bar{C}) &= N(2) + N(3) \\ N(A \wedge \bar{C}) &= N(1) + N(2) \end{aligned}$$

Plugging this into (9.12), you obtain Bell's inequality:

$$N(A \wedge \bar{B}) + N(B \wedge \bar{C}) \geq N(A \wedge \bar{C}) \tag{9.13}$$

We can now actually compare this inequality with the predictions of quantum theory, and also with experiments using Table 2! It turns out that with an entangled state like in (9.11), the probability that Alice and Bob measure "up-spin" in their respective measurement basis depends on the difference angle $\Delta\vartheta$ of the two measurement bases as follows:

$$P(\uparrow_{\vartheta 1};\uparrow_{\vartheta 2}) = \frac{1}{2}\sin^2\left(\frac{\vartheta_2 - \vartheta_1}{2}\right) = \frac{1}{2}\sin^2\left(\frac{\Delta\vartheta}{2}\right) \tag{9.14}$$

Example: Measurement Probability of Alice and Bob

The probability that Alice and Bob both measure "spins-up" along arbitrary measurement axes ϑ_1 and ϑ_2 is to be derived. We define the first measurement direction (of Alice) as the z-axis, i.e., $\vartheta = 0$. We are allowed to do this because the entangled state (9.11) has no preferred direction: No matter in which direction you measure, the particles are always anti-correlated. If you express the Bell state (9.11) in any basis $\{|\uparrow_\vartheta\rangle, |\downarrow_\vartheta\rangle\}$, the result is again $|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow_\vartheta\rangle_A |\downarrow_\vartheta\rangle_B - |\downarrow_\vartheta\rangle_A |\uparrow_\vartheta\rangle_B)$. Bob's rotated measurement direction then corresponds to the angle $\Delta\vartheta$ of the original problem.
Let's first calculate the probability P_1 that Alice measures the state $|\uparrow\rangle_A$:

$$\begin{aligned} P_1 &= \langle\psi|\hat{P}_{\uparrow A}\mathbb{1}_B|\psi\rangle \overset{(9.11)}{=} \\ &= \frac{1}{2}\left(\langle\uparrow|_A\langle\downarrow|_B - \langle\downarrow|_A\langle\uparrow|_B\right)\left(|\uparrow\rangle_A\langle\uparrow|_A\right)\left(|\uparrow\rangle_A|\downarrow\rangle_B - |\downarrow\rangle_A|\uparrow\rangle_B\right) = \\ &= \frac{1}{2}\left(\langle\uparrow|_A\langle\downarrow|_B - \langle\downarrow|_A\langle\uparrow|_B\right)\left(|\uparrow\rangle_A\langle\uparrow|\uparrow\rangle_A|\downarrow\rangle_B + |\uparrow\rangle_A\langle\uparrow|\downarrow\rangle_A|\uparrow\rangle_B\right) = \\ &= \frac{1}{2}\left(\langle\uparrow|_A\langle\downarrow|_B - \langle\downarrow|_A\langle\uparrow|_B\right)|\uparrow\rangle_A|\downarrow\rangle_B = \\ &= \frac{1}{2}\left(\langle\uparrow|\uparrow\rangle_A\langle\downarrow|\downarrow\rangle_B - \langle\downarrow|\uparrow\rangle_A\langle\uparrow|\downarrow\rangle_B\right) = \frac{1}{2} \end{aligned} \tag{9.15}$$

It's not a surprise as the spin before the measurement is indefinite in every direction – Alice therefore measures the state $|\uparrow\rangle$ with a probability of 50%. After the measurement of Alice, the state $|\psi\rangle$ collapses and becomes the state $|\tilde{\psi}\rangle$.

$$\begin{aligned} |\tilde{\psi}\rangle = \hat{P}_{\uparrow A}\mathbb{1}_B|\psi\rangle &= \frac{1}{\sqrt{2}}|\uparrow\rangle_A\langle\uparrow|_A\left(|\uparrow_A\rangle|\downarrow\rangle_B - |\downarrow\rangle_A|\uparrow\rangle_B\right) = \\ &= \frac{1}{\sqrt{2}}|\uparrow\rangle_A\langle\uparrow|\uparrow\rangle_A|\downarrow\rangle_B - |\uparrow\rangle_A\langle\uparrow|\downarrow\rangle_A|\uparrow\rangle_B = \frac{1}{\sqrt{2}}|\uparrow\rangle_A|\downarrow\rangle_B \end{aligned}$$

The new state must still be normalized, however; we thus get for $|\tilde{\psi}\rangle$:

$$|\tilde{\psi}\rangle = |\uparrow\rangle_A |\downarrow\rangle_B \tag{9.16}$$

After Alice has measured the state $|\uparrow\rangle_A$ at her particle, the state of Bob's particle is thus fixed at $|\downarrow\rangle_B$ (along the same measurement direction). That is precisely the effect that EINSTEIN called "spooky action at a distance"! Now we only need to calculate with what probability P_2 Bob measures the state $|\uparrow_\vartheta\rangle$ with the $|\tilde{\psi}\rangle$ state after Alice's measurement, according to (9.16):

$$\begin{aligned}\hat{P}_{\uparrow\vartheta} = |\uparrow_\vartheta\rangle\langle\uparrow_\vartheta| &\overset{(7.55)}{=} \left(\cos\left(\frac{\vartheta}{2}\right)|\uparrow\rangle + \sin\left(\frac{\vartheta}{2}\right)|\downarrow\rangle\right)\left(\cos\left(\frac{\vartheta}{2}\right)\langle\uparrow| + \sin\left(\frac{\vartheta}{2}\right)\langle\downarrow|\right) = \\ &= \cos^2\left(\frac{\vartheta}{2}\right)|\uparrow\rangle\langle\uparrow| + \sin^2\left(\frac{\vartheta}{2}\right)|\downarrow\rangle\langle\downarrow| + \sin\left(\frac{\vartheta}{2}\right)\cos\left(\frac{\vartheta}{2}\right)\left(|\uparrow\rangle\langle\downarrow| + |\downarrow\rangle\langle\uparrow|\right)\end{aligned} \tag{9.17}$$

This allows us to calculate the state $\mathbb{1}_A\hat{P}_{\uparrow\vartheta B}|\tilde{\psi}\rangle$ after Bob's measurement (under the condition that Alice has measured $|\uparrow\rangle_A$ previously):

$$\begin{aligned}|\Psi\rangle = \mathbb{1}_A\hat{P}_{\uparrow\vartheta B}|\tilde{\psi}\rangle &\overset{(9.16)}{=} \hat{P}_{\uparrow\vartheta B}|\uparrow\rangle_A|\downarrow\rangle_B \overset{(9.17)}{=} \\ &= \left[\cos^2\left(\frac{\vartheta}{2}\right)|\uparrow\rangle_B\langle\uparrow|_B + \sin^2\left(\frac{\vartheta}{2}\right)|\downarrow\rangle_B\langle\downarrow|_B\right]|\uparrow\rangle_A|\downarrow\rangle_B + \\ &\quad + \left[\sin\left(\frac{\vartheta}{2}\right)\cos\left(\frac{\vartheta}{2}\right)\left(|\uparrow\rangle_B\langle\downarrow|_B + |\downarrow\rangle_B\langle\uparrow|_B\right)\right]|\uparrow\rangle_A|\downarrow\rangle_B = \\ &= \cos^2\left(\frac{\vartheta}{2}\right)|\uparrow\rangle_A|\uparrow\rangle_B\langle\uparrow|\downarrow\rangle_B + \sin^2\left(\frac{\vartheta}{2}\right)|\uparrow\rangle_A|\downarrow\rangle_B\langle\downarrow|\downarrow\rangle_B \\ &\quad + \sin\left(\frac{\vartheta}{2}\right)\cos\left(\frac{\vartheta}{2}\right)\left(|\uparrow\rangle_A|\uparrow\rangle_B\langle\downarrow|\downarrow\rangle_B + |\uparrow\rangle_A|\downarrow\rangle_B\langle\uparrow|\downarrow\rangle_B\right) \\ &= \sin^2\left(\frac{\vartheta}{2}\right)|\uparrow\rangle_A|\downarrow\rangle_B + \sin\left(\frac{\vartheta}{2}\right)\cos\left(\frac{\vartheta}{2}\right)|\uparrow\rangle_A|\uparrow\rangle_B\end{aligned} \tag{9.18}$$

One sees that also Bob's measurement changes the state to $|\Psi\rangle$. Now we can calculate the probability P_2 that Bob measures his particle in the state $|\uparrow_\theta\rangle_B$, provided that Alice had previously measured at her particle $|\uparrow\rangle_A$:

$$\begin{aligned}P_2 = \langle\tilde{\psi}|\Psi\rangle &\overset{(9.16,9.18)}{=} \\ &= \langle\uparrow|_A\langle\downarrow|_B\left(\sin^2\left(\frac{\vartheta}{2}\right)|\uparrow\rangle_A|\downarrow\rangle_B + \sin\left(\frac{\vartheta}{2}\right)\cos\left(\frac{\vartheta}{2}\right)|\uparrow\rangle_A|\uparrow\rangle_B\right) \\ &= \sin^2\left(\frac{\vartheta}{2}\right)\langle\uparrow|\uparrow\rangle_A\langle\downarrow|\downarrow\rangle_B + \sin\left(\frac{\vartheta}{2}\right)\cos\left(\frac{\vartheta}{2}\right)\langle\uparrow|\uparrow\rangle_A\langle\downarrow|\downarrow\rangle_B = \sin^2\left(\frac{\vartheta}{2}\right)\end{aligned} \tag{9.19}$$

The joint probability $P(\uparrow;\uparrow_\vartheta)$ is now the product of P_1 and P_2:

$$P(\uparrow;\uparrow_\vartheta) = P_1P_2 \overset{(9.15,9.19)}{=} \frac{1}{2}\sin^2\left(\frac{\vartheta}{2}\right)$$

As already argued in the beginning, this is the same probability $P(\uparrow_{\vartheta 1},\uparrow_{\vartheta 2})$, that Alice measures $|\uparrow_{\vartheta 1}\rangle$ and Bob $|\uparrow_{\vartheta 2}\rangle$:

$$P(\uparrow_{\vartheta 1};\uparrow_{\vartheta 2}) = \frac{1}{2}\sin^2\left(\frac{\Delta\vartheta}{2}\right) \tag{9.20}$$

If we still follow the orientation according to Table 2, it becomes apparent that:

- for determining $N(A \wedge \bar{B})$, Alice had to measure in the 0° direction, and Bob in the 45°

direction. Therefore, in this case, $\Delta\vartheta = 45° = \pi/4$.

- for determining $N(B \wedge \bar{C})$, Alice had to measure in the 45° direction, and Bob in the 90° direction. Here, $\Delta\vartheta = 45° = \pi/4$.
- for determining $N(A \wedge \bar{C})$, Alice had to measure in the 0° direction, and Bob in the 90° direction. Here, $\Delta\vartheta = 90° = \pi/2$.

With this, we can compare the Bell inequality (9.13) with the quantum mechanically predicted measurement result:

$$\frac{1}{2}\sin^2\left(\frac{\pi/4}{2}\right) + \frac{1}{2}\sin^2\left(\frac{\pi/4}{2}\right) \geq \frac{1}{2}\sin^2\left(\frac{\pi/2}{2}\right) \implies 0.1464 \geq 0.25$$

This is obviously wrong! Thus, quantum theory clearly cannot be reconciled with the idea of hidden variables (at least not if they act locally only). In fact, this prediction of quantum theory was first verified in 1972 by Stuart Freedman and John Clauser, and thereafter by Alain Aspect and many other experimenters in increasingly sophisticated experiments. It indeed seems that Einstein was wrong in this case, and that in quantum physics, some sort of measurement (the "observation") creates reality: Before the measurement, Alice's particle has no definite spin in any direction, and after the measurement, the particle indeed has a definite spin along the measurement direction but not in the orthogonal direction.

Although a measurement on one particle in an entangled state also collapses the other particle, no information can be transmitted between the particles in this way. Thus, no information is transmitted faster than the speed of light in keeping with relativity theory.

Deepening: Formal Derivation of Bell's Inequalities

We define a continuous hidden variable λ (previously $\boldsymbol{\lambda}$ could only take three positions), which should satisfy the following relation with an appropriate distribution function $\rho(\lambda)$:

$$\int d\lambda\, \rho(\lambda) = 1 \tag{9.21}$$

We now define – with reference to the hidden variable λ – the expected correlation for different measurement directions $\mathbf{m}$ and $\mathbf{n}$ (as well as $\mathbf{p}$). Here, the correlation of the measurements in any directions $E(\mathbf{m}, \mathbf{n})$ is to be understood as the product of the Pauli matrices $\hat{\sigma}_1(\mathbf{n})$ and $\hat{\sigma}_2(\mathbf{m})$ (or spin operators):

$$E(\mathbf{m}, \mathbf{n}) = \int d\lambda\, \rho(\lambda)\, \langle \hat{\sigma}_1(\mathbf{m}) \hat{\sigma}_2(\mathbf{n}) \rangle_\lambda$$

Due to the *locality* assumed by EPR, it should also be valid that the spin measurements are independent of each other and (after production) there is no interaction between the particles:

$$\langle \hat{\sigma}_1(\mathbf{m}) \hat{\sigma}_2(\mathbf{n}) \rangle_\lambda = \langle \hat{\sigma}_1(\mathbf{m}) \rangle_\lambda \langle \hat{\sigma}_2(\mathbf{n}) \rangle_\lambda$$

Following the assumption of *reality*, it follows that the measurement of the spin gives the observable quantity $\langle \hat{\sigma}_i \rangle (\mathbf{n}) \pm 1$ – and that for each measurement direction $\mathbf{n}$ (by using the Pauli matrices, the usual factor $\hbar/2$ is omitted here). The spins of both particles are always oriented oppositely due to our problem setting; it follows through $E(\mathbf{m}, \mathbf{m}) = -1$:

$$\langle \hat{\sigma}_1(\mathbf{m}) \rangle_\lambda = - \langle \hat{\sigma}_2(\mathbf{m}) \rangle_\lambda$$

We can now summarize this for the expected correlation of the two measurements:

$$E(\mathbf{m},\mathbf{n}) = -\int d\lambda\, \rho(\lambda)\, \langle\hat{\sigma}_1(\mathbf{m})\rangle_\lambda\, \langle\hat{\sigma}_1(\mathbf{n})\rangle_\lambda \tag{9.22}$$

Calculating the difference between two expected correlations, with $\mathbf{n}$ now shifted to $\mathbf{p}$, we can furthermore write with $\langle\hat{\sigma}_1(\mathbf{m})\rangle_\lambda^2 = (\pm 1)^2 = 1$:

$$\begin{aligned} E(\mathbf{m},\mathbf{n}) - E(\mathbf{m},\mathbf{p}) &= \int d\lambda\, \rho(\lambda)\, (\langle\hat{\sigma}_1(\mathbf{m})\rangle_\lambda\, \langle\hat{\sigma}_1(\mathbf{p})\rangle_\lambda - \langle\hat{\sigma}_1(\mathbf{m})\rangle_\lambda\, \langle\hat{\sigma}_1(\mathbf{n})\rangle_\lambda) = \\ &= \int d\lambda\, \rho(\lambda)\, \Big(\langle\hat{\sigma}_1(\mathbf{m})\rangle_\lambda\, \langle\hat{\sigma}_1(\mathbf{n})\rangle_\lambda^2\, \langle\hat{\sigma}_1(\mathbf{p})\rangle_\lambda - \langle\hat{\sigma}_1(\mathbf{m})\rangle_\lambda\, \langle\hat{\sigma}_1(\mathbf{n})\rangle_\lambda\Big) = \\ &= \int d\lambda\, \rho(\lambda)\, \langle\hat{\sigma}_1(\mathbf{m})\rangle_\lambda\, \langle\hat{\sigma}_1(\mathbf{n})\rangle_\lambda\, \Big(\langle\hat{\sigma}_1(\mathbf{n})\rangle_\lambda\, \langle\hat{\sigma}_1(\mathbf{p})\rangle_\lambda - 1\Big) \end{aligned}$$

However, we are only interested in the absolute difference of the correlations; because the product of two expectation values $|\,\langle\hat{\sigma}_1(\mathbf{m})\rangle_\lambda\, \langle\hat{\sigma}_1(\mathbf{n})\rangle_\lambda\,| \leq 1$ must be (at most we get ± 1 as an expectation value), we can perform the following estimate:

$$\begin{aligned} |E(\mathbf{m},\mathbf{n}) - E(\mathbf{m},\mathbf{p})| &\leq \int d\lambda\, \rho(\lambda)\, (1 - \langle\hat{\sigma}_1(\mathbf{n})\rangle_\lambda\, \langle\hat{\sigma}_1(\mathbf{p})\rangle_\lambda) = \\ &= \int d\lambda\, \rho(\lambda) - \int d\lambda\, \rho(\lambda)\, \langle\hat{\sigma}_1(\mathbf{n})\rangle_\lambda\, \langle\hat{\sigma}_1(\mathbf{p})\rangle_\lambda \overset{(9.22)}{=} 1 + E(\mathbf{n},\mathbf{p}) \end{aligned}$$

Simple rearrangement thus leads us to the *Bell inequality*, which must be satisfied for a locally-realistic theory *with* hidden variables:

$$1 + E(\mathbf{n},\mathbf{p}) - |E(\mathbf{m},\mathbf{n}) - E(\mathbf{m},\mathbf{p})| \geq 0 \tag{9.23}$$

Correlation between two operators For the expectation value of the correlation between the spin operators $\hat{\sigma}_1(\mathbf{m})$ and $\hat{\sigma}_2(\mathbf{n})$, we find the following expression:

$$E(\mathbf{m},\mathbf{n}) = \langle\psi|\hat{\sigma}_1(\mathbf{m})\hat{\sigma}_2(\mathbf{n})|\psi\rangle = -\mathbf{m}\mathbf{n} = -\cos(\Delta\varphi) \tag{9.24}$$

We show this explicitly using the state $|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - \downarrow\uparrow)$ from (9.11) and the Pauli matrices for arbitrary measurement directions according to (7.54). The evaluation takes place in the matrix representation of $\hat{\sigma}_i$, hence for example $\langle\uparrow|\hat{\sigma}|\uparrow\rangle_1 = \cos(\vartheta_1)$:

$$\begin{aligned} E(\mathbf{m},\mathbf{n}) &= \langle\psi|\hat{\sigma}_1(\mathbf{m})\hat{\sigma}_2(\mathbf{n})|\psi\rangle \overset{(9.11)}{=} \\ &= \frac{1}{2}\,(\langle\uparrow\downarrow| - \langle\downarrow\uparrow|)\,\hat{\sigma}_1\hat{\sigma}_2\,(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) = \;\big|\text{ e.g. } \langle\uparrow|\hat{\sigma}|\uparrow\rangle_1 = \langle\uparrow|\hat{\sigma}|\uparrow\rangle_2 \\ &= \langle\uparrow|\hat{\sigma}|\uparrow\rangle_1\, \langle\downarrow|\hat{\sigma}|\downarrow\rangle_2 - \langle\uparrow|\hat{\sigma}|\downarrow\rangle_1\, \langle\downarrow|\hat{\sigma}|\uparrow\rangle_2 = \\ &= \cos(\varphi_1)(-\cos(\varphi_2)) - \sin(\varphi_1)\sin(\varphi_2) = -\cos(\varphi_1 - \varphi_2) = -\cos(\Delta\vartheta) \end{aligned}$$

Now, if we explicitly substitute $E(\mathbf{m},\mathbf{n})$ from (9.24) into Bell's inequality (9.23), the following results:

$$1 - \mathbf{n}\mathbf{p} - |\mathbf{m}\mathbf{p} - \mathbf{m}\mathbf{n}| \geq 0$$

We understand $\mathbf{m}$, $\mathbf{n}$ and $\mathbf{p}$ as direction vectors, where ϕ_1 is the angle between $\mathbf{m}$ and $\mathbf{n}$ and ϕ_2 is the angle between $\mathbf{m}$ and $\mathbf{p}$. Thus, for Bell's inequality we find:

$$1 - \cos(\phi_2 - \phi_1) - |\cos(\phi_2) - \cos(\phi_1)| \geq 0 \tag{9.25}$$

This inequality is *not* satisfied for all angles ϕ_1 and ϕ_2 and thus violates (9.23)!

9.5 Technological Applications

In the following section, some technological applications such as random generators, quantum computers, and quantum encryption will be discussed.

9.5.1 Random Number Generators

For computers, generating random numbers is a great challenge. Most often, algorithms only generate so-called pseudorandom numbers. The sequences of numbers generated by such algorithms seem random but are not truly random since they are generated by a deterministic algorithm. If a specific random number algorithm is initialized with the same starting value ("seed") every time, the same sequence of pseudorandom numbers will result each time.

However, through quantum theory, it is now possible for us to actually generate sequences of true random numbers based on the superposition of a two-state system. Let's look at Figure 48. A source emits particles, which have been polarized in the z-direction and are in the state $|\uparrow\rangle$. These electrons, prepared in such a way, subsequently move through a Stern-Gerlach apparatus, whose magnetic field is rotated by $\Delta\vartheta = \pi/2$ relative to the prepared polarization direction. The Stern-Gerlach apparatus is thus *de facto* a measuring device that measures the spin in the x-direction. This causes the electrons prepared in the z-direction to be in a superposition of the states $|\uparrow_x\rangle$ and $|\downarrow_x\rangle$ from the perspective of the measurement basis of the Stern-Gerlach apparatus:

$$|\uparrow\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_x\rangle + |\downarrow_x\rangle\right)$$

The apparatus allows only the transmission of the $|\uparrow_x\rangle$ state; and that with a probability of $P(\uparrow,\uparrow_x) = 0.5$. Thus, with the same probability, the electron does not pass the Stern-Gerlach apparatus! If we now assign the transmission of the electron the value $|\uparrow_x\rangle = |0\rangle$ and the opposite state $|\downarrow_x\rangle = |1\rangle$, we generate $|0\rangle$ and $|1\rangle$ with an evenly distributed probability. The randomness is inherent here – we cannot predict whether the initial state $|\uparrow\rangle$ will collapse into $|0\rangle$ or $|1\rangle$. Thus, an absolutely random sequence of zeros and ones can be generated, which can be interpreted by a computer program as a perfect sequence of random numbers.

9.5.2 Quantum Computers

While a classical computer works with the classical bits 0 and 1, a quantum computer uses *qubits* (or *Q-bits*); that is, a system with two distinct quantum states $|0\rangle$ and $|1\rangle$. Apart from the notation, this does not seem to be a big difference at first glance. However, the quantum mechanical states $|0\rangle$ and $|1\rangle$ are capable of forming a superposition:

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$$

Furthermore, we can, depending on our computer system, entangle states. Analogue to the classical computer, arbitrary numbers are to be represented in binary. In Table 3, the coding of numbers 0 to 7 using three classical bits and three entangled qubits is shown.
Of course, we can also bring the entangled states into superposition (although, unfortunately, in practice this becomes increasingly difficult with the degree of entanglement). For n-entangled states we find:

$$\begin{aligned}|\psi\rangle &= a_0\,|00\ldots0\rangle + a_1\,|00\ldots1\rangle + \cdots + a_{2^n-1}\,|11\ldots1\rangle = \\ &= a_0\,|0\rangle + a_1\,|1\rangle + \cdots + a_{2^n-1}\,|2^n-1\rangle\end{aligned} \tag{9.26}$$

A single state thus contains the information of $2^n - 1$ complex numbers in the form of the expansion coefficients. In the case of $n = 3$ qubits, as in Table 3, there are only 7 coefficients,

	class.	entangled states
0	000	$\|0\rangle \otimes \|0\rangle \otimes \|0\rangle = \|000\rangle \equiv \|0\rangle$
1	001	$\|0\rangle \otimes \|0\rangle \otimes \|1\rangle = \|001\rangle \equiv \|1\rangle$
2	010	$\|0\rangle \otimes \|1\rangle \otimes \|0\rangle = \|010\rangle \equiv \|2\rangle$
3	011	$\|0\rangle \otimes \|1\rangle \otimes \|1\rangle = \|011\rangle \equiv \|3\rangle$
4	100	$\|1\rangle \otimes \|0\rangle \otimes \|0\rangle = \|100\rangle \equiv \|4\rangle$
5	101	$\|1\rangle \otimes \|0\rangle \otimes \|1\rangle = \|101\rangle \equiv \|5\rangle$
6	110	$\|1\rangle \otimes \|1\rangle \otimes \|0\rangle = \|110\rangle \equiv \|6\rangle$
7	111	$\|1\rangle \otimes \|1\rangle \otimes \|1\rangle = \|111\rangle \equiv \|7\rangle$

Table 3: Example for the binary representation of numbers with 3 classical bits or three entangled qubits.

but for $n = 100$ qubits, there would already be 10^{30} coefficients! If we now perform a calculation, that is, let an operator act on $|\psi\rangle$, the linearity of quantum mechanics follows:

$$\hat{U}\,|\psi\rangle = a_0\hat{U}\,|0\rangle + a_1\hat{U}\,|1\rangle + \cdots + a_{2^n-1}\hat{U}\,|2^n-1\rangle \tag{9.27}$$

The effect of $\hat{U}$ on the superposition state thus corresponds to the *parallel* effect on the basis states; we thus perform $2^n - 1$ calculations simultaneously! An operation does not destroy the superposition and is also reversible – only when we measure the state does the superposition collapse into a single basis state $|i\rangle$. To better imagine the effect of any $\hat{U}$ in the simple case of the superposition of two states, one can use the Bloch sphere. We have shown in Figure 40 that a superposition of $|0\rangle$ and $|1\rangle$ (or $|\uparrow\rangle$ and $|\downarrow\rangle$) can be represented as a vector in the context of this space. When $\hat{U}$ acts on $|\psi\rangle$, it results in a rotation of $|\psi\rangle$ and a change in the expansion coefficients. In the case of a superposition of more than two basis states, the illustrative concept of the Bloch sphere is unfortunately no longer applicable.

9.5.3 "No-cloning" Theorem

Motivation: Cloning of Quantum States

As we already know, the measurement of a general quantum state destroys that state. For example, if we measure the spin in the z direction of the state $|\psi\rangle = \alpha\,|\uparrow\rangle - \beta\,|\downarrow\rangle$ using $\hat{S}_z$, the result of the individual measurement will either be $|\uparrow\rangle$ or $|\downarrow\rangle$, with probabilities $|\alpha|^2$ and $|\beta|^2$, respectively. However, after the measurement, the original state is irrevocably destroyed and depending on the (random) measurement result, transitions to the state $|\psi\rangle' = |\uparrow\rangle$ or $|\psi\rangle' = |\downarrow\rangle$. Therefore, multiple measurements are not possible, and there is no way to determine the values of α and β from this single measurement.

But perhaps there is a unitary clone operator $\hat{U}_\mathrm{K}$, which is somehow able to transfer the (Source) state $|\psi_S\rangle$ onto another (Target) state $|\psi_T\rangle$, without changing the (Source) state? In this section we will explore why this is not possible, why such a clone operator violates fundamental laws of quantum physics and would also allow for superluminal communication.

Is it thus possible to copy a general, arbitrary (Source) state $|\psi_S\rangle \in \mathcal{H}_S$ onto a (Target) state $|0_T\rangle \in \mathcal{H}_T$ through a unitary operation $\hat{U}_\mathrm{K}$; or in other words: Is it possible to clone a quantum state? To answer this question, we first define such a (hypothetical) clone operator through the following relation:

$$\hat{U}_\mathrm{K}\left(|\psi_S\rangle \otimes |0_T\rangle\right) = |\psi_S\rangle \otimes |\psi_T\rangle \tag{9.28}$$

$\hat{U}_\mathrm{K}$ causes the state of the (Source) system $\mathcal{H}_S$ to now also be established in the (Target) system $\mathcal{H}_T$. The clone operator should be able to overwrite $|0_T\rangle$ *with any arbitrary* state. We therefore

consider two arbitrary, different states $|\phi\rangle$ and $|\psi\rangle$:

$$\begin{aligned}\hat{U}_{\mathrm{K}}\left(|\phi_S\rangle\,|0_T\rangle\right) &= |\phi_S\rangle\,|\phi_T\rangle \\ \hat{U}_{\mathrm{K}}\left(|\psi_S\rangle\,|0_T\rangle\right) &= |\psi_S\rangle\,|\psi_T\rangle\end{aligned} \tag{9.29}$$

To determine if the assumed clone operator $\hat{U}_{\mathbf{K}}$ can truly clone any arbitrary state, we consider the expression $\langle\phi|\psi\rangle$. We also assume that $|0\rangle$ is normalized, hence $\langle 0|0\rangle = 1$. Utilizing the unitarity of $\hat{U}_{\mathrm{K}}$, it holds:

$$\begin{aligned}\langle\phi|\psi\rangle &= \langle\phi_S|\psi_S\rangle\,\langle 0_T|0_T\rangle = \langle\phi_S|\,\langle 0_T|\psi_S\rangle\,|0_T\rangle = \langle\phi_S|\,\langle 0_T|\hat{U}_{\mathrm{K}}^{\dagger}\hat{U}_{\mathrm{K}}|\psi_S\rangle\,|0_T\rangle \overset{(9.29)}{=} \\ &= \langle\phi_S|\,\langle\phi_T|\psi_S\rangle\,|\psi_T\rangle = \langle\phi_S|\psi_T\rangle\,\langle\phi_T|\psi_S\rangle = \langle\phi|\psi\rangle^2\end{aligned}$$

Thus, we have found that the clone operator only works for the state $|\phi\rangle$, and also for the state $|\psi\rangle$ under the following condition:

$$\langle\phi|\psi\rangle = \langle\phi|\psi\rangle^2$$

However, this means that $|\phi\rangle$ and $|\psi\rangle$ are not arbitrary, independent states: First consider the case where $\langle\phi|\psi\rangle = 1$. Then, it must be $|\phi\rangle = e^{i\varphi}\,|\psi\rangle$. Since the global phase does not matter, this means in other words, that the states $|\phi\rangle$ and $|\psi\rangle$ are identical. Conversely, if $\langle\phi|\psi\rangle = 0$, then the state $|\phi\rangle$ is orthonormal to $|\psi\rangle$.

Therefore, we have shown: If there is a clone operator $\hat{U}_{\mathrm{K}}$ that works for any arbitrary state $|\phi\rangle$, it only works (trivially) for identical, and otherwise only for orthogonal states. It does not work for any other states. Thus, it is clear that a clone operator for arbitrary states cannot exist.

Example: Cloning of a General Qubit

Let's consider a general qubit, represented by the superposition of the states $|0\rangle$ and $|1\rangle$ with the complex coefficients α and β as:

$$|\psi_S\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \tag{9.30}$$

We let the clone operator $\hat{U}_{\mathrm{K}}$ act on $|\psi_S\rangle\,|0_T\rangle$, to transfer the state $|\psi_S\rangle$ onto $|0_T\rangle$. Following (9.29):

$$\begin{aligned}\hat{U}_{\mathrm{K}}\left(|\psi_S\rangle\,|0_T\rangle\right) &= |\psi_S\rangle\,|\psi_T\rangle \overset{(9.30)}{=} \\ &= \left(\alpha\,|0_S\rangle + \beta\,|1_S\rangle\right)\left(\alpha\,|0_T\rangle + \beta\,|1_T\rangle\right) = \\ &= \alpha^2\,|0_S\rangle\,|0_T\rangle + \alpha\beta\,|0_S\rangle\,|1_T\rangle + \beta\alpha\,|1_S\rangle\,|0_T\rangle + \beta^2\,|1_S\rangle\,|1_T\rangle \equiv \\ &\equiv \alpha^2\,|00\rangle + \alpha\beta\,|01\rangle + \beta\alpha\,|10\rangle + \beta^2\,|11\rangle\end{aligned} \tag{9.31}$$

Due to the linearity of the superposition we can also write:

$$\begin{aligned}\hat{U}_{\mathrm{K}}\left(|\psi_S\rangle\,|0_T\rangle\right) &= \hat{U}_{\mathrm{K}}\left[\left(\alpha\,|0_S\rangle + \beta\,|1_S\rangle\right)|0_T\rangle\right] = \\ &= \alpha\hat{U}_{\mathrm{K}}\,|0_S\rangle\,|0_T\rangle + \beta\hat{U}_{\mathrm{K}}\,|1_S\rangle\,|0_T\rangle \overset{(9.29)}{=} \\ &= \alpha\,|0_S\rangle\,|0_T\rangle + \beta\,|1_S\rangle\,|1_T\rangle \equiv \alpha\,|00\rangle + \beta\,|11\rangle\end{aligned} \tag{9.32}$$

Comparing (9.31) with (9.32), we recognize that $\alpha^2 = \alpha$, $\beta^2 = \beta$ and $\alpha\beta = \beta\alpha = 0$ must be satisfied. This is, however, only fulfilled if $\alpha = 1$ and $\beta = 0$ (or vice versa) holds. It is therefore evident that the clone operator $\hat{U}_{\mathrm{K}}$ can only copy the base states $|0\rangle$ and $|1\rangle$, but not general superpositions as in (9.30).

The fact that a clone operator cannot exist may seem at first glance very unfortunate for certain applications. For instance, it is an inherent problem of the quantum computer that it is not possible to simply copy an arbitrary quantum state, composed of a superposition of many qubits. However, if one considers the consequences of a potential clone operator more closely, one realizes that they would set aside both the foundations of quantum physics itself and the special theory of relativity.

Superluminal Communication with the Clone Operator If a clone operator existed, the following superluminal communication protocol could be applied: Alice and Bob wish to communicate with each other superluminally in binary code. Exactly in the middle between them is a particle source that emits a steady stream of entangled spin-$\frac{1}{2}$ particles in opposite directions, so that Alice always receives a particle when the corresponding, entangled particle arrives at Bob. The entangled particles have the following state:

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow\rangle_A |\downarrow\rangle_B - |\downarrow\rangle_A |\uparrow\rangle_B \right) \tag{9.33}$$

This means, if Alice measures in the z basis, there is a 50% probability that she will find her particle in state $|\uparrow\rangle_A$, but also a 50% probability that she will determine the state $|\downarrow\rangle_A$. Due to the entanglement, we can be sure: Every time Alice measures $|\uparrow\rangle_A$, Bob will measure $|\downarrow\rangle_B$, and vice versa.
A crucial point in our argument is the use of two orthogonal measurement bases. One measurement basis is simply the spin in the z direction $\{|\uparrow\rangle, |\downarrow\rangle\}$, while the other measurement basis is the spin in the x direction $\{|\uparrow_x\rangle, |\downarrow_x\rangle\} \equiv \{|\rightarrow\rangle, |\leftarrow\rangle\}$. We already know from (7.55) how to express $|\rightarrow\rangle$ and $|\leftarrow\rangle$ in the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$:

$$|\rightarrow\rangle \equiv |\uparrow_x\rangle \overset{(7.55)}{=} \cos(\pi/4) |\uparrow\rangle + \sin(\pi/4) |\downarrow\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow\rangle + |\downarrow\rangle \right) \tag{9.34}$$

$$|\leftarrow\rangle \equiv |\downarrow_x\rangle \overset{(7.55)}{=} \sin(\pi/4) |\uparrow\rangle - \cos(\pi/4) |\downarrow\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow\rangle - |\downarrow\rangle \right) \tag{9.35}$$

If we add (9.34) and (9.35), or subtract (9.35) from (9.34), we get $|\uparrow\rangle$ and $|\downarrow\rangle$ in the basis $\{|\rightarrow\rangle, |\leftarrow\rangle\}$:

$$|\rightarrow\rangle + |\leftarrow\rangle = \sqrt{2} |\uparrow\rangle \implies |\uparrow\rangle = \frac{1}{\sqrt{2}} \left(|\rightarrow\rangle + |\leftarrow\rangle \right) \tag{9.36}$$

$$|\rightarrow\rangle - |\leftarrow\rangle = \sqrt{2} |\downarrow\rangle \implies |\downarrow\rangle = \frac{1}{\sqrt{2}} \left(|\rightarrow\rangle - |\leftarrow\rangle \right) \tag{9.37}$$

We can thus use (9.36) and (9.37) to express the entangled state $|\psi\rangle$ from (9.33) in the basis $\{|\rightarrow\rangle, |\leftarrow\rangle\}$ and show that it retains an equivalent form in the new basis too. It follows:

$$
\begin{aligned}
|\psi\rangle &= \frac{1}{\sqrt{2}} \quad (|\uparrow\rangle_A |\downarrow\rangle_B - |\downarrow\rangle_A |\uparrow\rangle_B) \overset{(9.36,9.37)}{=} \\
&= \frac{1}{\sqrt{2}} \quad \left[\frac{1}{2}(|\rightarrow\rangle_A + |\leftarrow\rangle_A)(|\rightarrow\rangle_B - |\leftarrow\rangle_B) - \frac{1}{2}(|\rightarrow\rangle_A - |\leftarrow\rangle_A)(|\rightarrow\rangle_B + |\leftarrow\rangle_B)\right] = \\
&= \frac{1}{\sqrt{2}} \quad \left[\frac{1}{2}(|\rightarrow\rangle_A |\rightarrow\rangle_B - |\leftarrow\rangle_A |\leftarrow\rangle_B + |\leftarrow\rangle_A |\rightarrow\rangle_B - |\rightarrow\rangle_A |\leftarrow\rangle_B) - \right. \\
&\qquad \left. - \frac{1}{2}(|\rightarrow\rangle_A |\rightarrow\rangle_B - |\leftarrow\rangle_A |\leftarrow\rangle_B - |\leftarrow\rangle_A |\rightarrow\rangle_B + |\rightarrow\rangle_A |\leftarrow\rangle_B)\right] = \\
&= \frac{1}{\sqrt{2}} \quad \left[\frac{1}{2}(|\rightarrow\rangle_A |\rightarrow\rangle_B - |\leftarrow\rangle_A |\leftarrow\rangle_B + |\leftarrow\rangle_A |\rightarrow\rangle_B - |\rightarrow\rangle_A |\leftarrow\rangle_B) + \right. \\
&\qquad \left. + \frac{1}{2}(-|\rightarrow\rangle_A |\rightarrow\rangle_B + |\leftarrow\rangle_A |\leftarrow\rangle_B + |\leftarrow\rangle_A |\rightarrow\rangle_B - |\rightarrow\rangle_A |\leftarrow\rangle_B)\right] = \\
&= \frac{1}{sqrt2}(|\leftarrow\rangle_A |\rightarrow\rangle_B - |\rightarrow\rangle_A |\leftarrow\rangle_B) \qquad (9.38)
\end{aligned}
$$

We thus recognize: If Alice measures her particle in the basis $\{|\rightarrow\rangle, |\leftarrow\rangle\}$, she will also find with a 50% probability the state $|\rightarrow\rangle_A$ and can be sure that the (distant) particle at Bob is also instantaneously in the anti-correlated state $|\leftarrow\rangle_B$. With a 50% probability, however, Alice measures the state $|\rightarrow\rangle_A$, and then knows that the particle at Bob is likewise immediately in the state $|\leftarrow\rangle_B$.

Everything is therefore prepared for the superluminal communication protocol: Whenever Alice wants to send a 0, she measures her particle in the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$. Her particle will then be either in the state $|\uparrow\rangle_A$ or $|\downarrow\rangle_A$. The entangled particle at Bob will be in the opposite state $|\downarrow\rangle_B$ or $|\uparrow\rangle_B$, as can be seen from (9.33). If Bob can only conduct *one* measurement on his particle in the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$, it does not benefit him, since he will detect one or the other state with a 50% probability. But if Bob had a clone operator, he could replicate the state of his particle and conduct very many measurements. He would then observe (when measured in the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$ that the measurement on all cloned particles always yields uniformly $|\downarrow\rangle_B$ or uniformly $|\uparrow\rangle_B$. From this, he could infer that Alice coded a 0; instantaneously, even if she were many million kilometers away.

If Alice, however, wants to code a 1, then she performs the measurement on her particle in the basis $\{|\rightarrow\rangle, |\leftarrow\rangle\}$. Her particle will then be either in the state $|\rightarrow\rangle_A$ or $|\leftarrow\rangle_A$. The entangled particle at Bob will be in the opposite state $|\leftarrow\rangle_B$ or $|\rightarrow\rangle_B$, as can be seen from (9.38). Let us assume that Bob continues to take measurements of his particle in the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$. If his particle is in the state $|\rightarrow\rangle_B$, he has according to (9.34) a 50-50 probability of measuring $|\uparrow\rangle$ or $|\downarrow\rangle$. The same applies if his particle is in the state $|\leftarrow\rangle_B$. If he can only conduct one measurement, he is unable to determine this probability. But with the help of the clone operator, he could again repeat this measurement on the cloned particles as often as he likes! He would then observe that the measurement in the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$ on the cloned states yields completely random results, and could infer from this that Alice coded a 1.

9.5.4 Quantum Cryptography

Motivation: Key exchange in cryptography.

When two communication partners (as usual Alice and Bob) want to communicate secretly using an encryption algorithm, they must first agree on a key and exchange it between them. No matter how good the actual encryption algorithm is: there is always a risk that an unauthorized third person, Eve (after "Eavesdropper"), intercepts this key or learns of it through another way. Thus, any encryption strategy, no matter how good otherwise, is at risk.

This is a seemingly unsolvable problem in cryptography. How are Alice and Bob supposed to know that their key has indeed remained secret? It turns out that quantum physics offers an elegant solution to this problem.

We want to discuss a particular form of "quantum key distribution", namely the Bennett-Brassard protocol (BB84). Let's assume that Alice wants to send a secret message to Bob. An encryption should be used, the key of which must be known to both Alice and Bob. To establish a key, Alice could, for example, generate a stream of two entangled spin-$\frac{1}{2}$ particles each, which are (as already described in the previous chapter) in the following state:

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow\rangle_A |\downarrow\rangle_B - |\downarrow\rangle_A |\uparrow\rangle_B \right) \tag{9.39}$$

Alice could then perform spin measurements on "her" particle in the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$ and send the other, entangled particles to Bob, who measures them in the same basis. There is a 50% probability that Alice will determine the state $|\uparrow\rangle_A$. She then notes down a 1, as agreed. The entangled particle at Bob's side is then automatically in the state $|\downarrow\rangle_B$; for Bob, $|\downarrow\rangle_B$ means 1. Conversely, if Alice determines the state $|\downarrow\rangle_A$, Bob will measure $|\uparrow\rangle_B$, and both can note down a 0. This way, Alice and Bob can generate an arbitrarily long, identical random sequence of zeros and ones that represent their shared key.
However, the key thus exchanged is by no means secure against potential eavesdropping, since Eve could observe the communication channel and also perform a spin measurement in the z-basis on the particle that Alice sends to Bob. This would allow her to obtain the key as well. Eve's measurement does not change the state of the particle, since she measures in the same basis as Alice and Bob. Bob cannot, therefore, notice that the key has been "intercepted".

This problem can, however, be elegantly solved: Alice randomly decides for each measurement whether she measures the spin state in the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$ or in an orthonormal direction $\{|\leftarrow\rangle, |\rightarrow\rangle\}$. For $|\rightarrow\rangle \equiv |\uparrow_x\rangle$ and $|\leftarrow\rangle \equiv |\downarrow_x\rangle$ we have already derived in the previous section (9.38) that the entangled state $|\psi\rangle$ can also be written as follows:

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|\leftarrow\rangle_A |\rightarrow\rangle_B - |\rightarrow\rangle_A |\leftarrow\rangle_B \right) \tag{9.40}$$

This means, if Alice randomly determines the state $|\rightarrow\rangle_A$, Bob will find the state $|\leftarrow\rangle_B$ (provided that both measure in the basis $\{|\leftarrow\rangle, |\rightarrow\rangle\}$), and both can note down a 1. In the opposite case, both write down a 0.
The crucial point is that Alice does not tell Bob whether she measured in the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$ or $\{|\leftarrow\rangle, |\rightarrow\rangle\}$. Bob must, therefore, decide (randomly) in which of these two bases he performs measurements himself for each measurement. In cases where Alice and Bob happen to measure in the same basis, both will have noted the same random bit. In all cases where they have measured in different bases, the measurement results (and thus the noted bits) are completely

uncorrelated.

At the end of the particle transmission, Alice tells Bob in which bases she measured her spin states. Bob can thus discard all measurements in which he measured in a different basis than Alice, resulting in a shared random key at the end!

Alice	Basis	↕	↕	↔	↕	↔	↕	↔	↔	↔	↕	↔	↕
	Measurement	1	0	0	1	1	0	1	1	1	0	0	1
Bob	Basis	↕	↔	↔	↕	↔	↔	↕	↕	↔	↕	↕	↕
	Measurement	1	1	0	1	1	1	0	1	0	0	0	1
	Key	1	-	0	1	-	-	-	-	0	0	-	1

Table 4: Alice and Bob measure the spin of a particle in two (orthogonal) bases; the key corresponds only to the measured states that were measured in the same bases (highlighted in green). However, if Eve eavesdrops, she may measure in a different basis and thus change the state (highlighted in red). This is manifested when comparing parts of the bit sequence of Alice and Bob: Despite the same measurement basis, they have noted a different digit!

But what happens if Eve wants to eavesdrop? She now has the same problem as Bob because she also doesn't know in which basis Alice measures her particle spins. Therefore, Eve also has to randomly choose a basis for each individual measurement. If she chooses the wrong basis, while Alice and Bob both measure in the same basis, her measurement alters the state of the particle. Alice and Bob should ideally measure the opposite spin of their particle, but with a 50% probability, Bob will now measure the same state as Alice due to Eve's interference. This is a consequence of the "no-cloning" theorem from (9.29), which prohibits copying a state.

To check whether someone has observed the key exchange (i.e., changed the state of the particles through measurement), Alice and Bob compare a small subsequence of their bit sequence. If there is a discrepancy (we recall that Bob keeps only the measurements that took place in the same basis as Alice's), both know they were eavesdropped on and will not use the key anymore.

10 Appendix

In the following sections, aspects that were not discussed in detail in the script will be examined more closely. Besides the transition from classical to quantum mechanics, the mathematical background of the differential equations encountered in the previous chapters will be investigated here, specifically the Hermite, Legendre, and Laguerre differential equations and the used Rodrigues formulas for solutions.

10.1 Transition from Classical to Quantum Mechanics

One assumes a set of generalized coordinates with N independent degrees of freedom:

$$\mathbf{q} = (q_1, \ldots, q_N) \quad \text{and} \quad \mathbf{p} = (p_1, \ldots, p_N) \tag{10.1}$$

$H(\mathbf{q}, \mathbf{p}, t)$ as the Hamiltonian function shall depend functionally on these generalized coordinates. The Hamiltonian equations of motion consist of $2N$ differential equations for $2N$ unknown functions $\mathbf{q}(t)$ and $\mathbf{p}(t)$ and are determined by:

$$\dot{q}_k = \frac{\partial H}{\partial p_k} \quad \text{and} \quad \dot{p}_k = -\frac{\partial H}{\partial q_k} \quad \text{with} \quad k = 1, \ldots, N \tag{10.2}$$

If a transformation of the phase space onto itself is now performed, then $(\mathbf{q}, \mathbf{p}) \to (\mathbf{q}', \mathbf{p}')$ or $H(\mathbf{q}, \mathbf{p}, t) \to \tilde{H}(\mathbf{q}', \mathbf{p}', t)$. It is called a canonical transformation if the Hamiltonian equations of motion remain invariant under such a transformation.

$$\dot{q}'_k = \frac{\partial \tilde{H}}{\partial p'_k} \quad \text{and} \quad \dot{p}'_k = -\frac{\partial \tilde{H}}{\partial q'_k}$$

The transformed Hamiltonian function $\tilde{H}$ can differ from the original Hamiltonian function H by the time derivative of a generating function $S = S(\mathbf{q}, \mathbf{p}', t)$, as this vanishes during variation of action:

$$\tilde{H}(\mathbf{q}', \mathbf{p}', t) = H(\mathbf{q}, \mathbf{p}, t) + \frac{\partial S}{\partial t} \tag{10.3}$$

S as a function of the original position coordinate and the transformed momentum coordinate is chosen such that the following relations hold:

$$p_k = \frac{\partial}{\partial q_k} S(q_k, p'_k, t) \quad \text{and} \quad q'_k = \frac{\partial}{\partial p'_k} S(q_k, p'_k, t) \tag{10.4}$$

An attempt is now made to use canonical transformations to produce a Hamiltonian function $\tilde{H}(\mathbf{q}', \mathbf{p}', t)$ that vanishes for all times:

$$\tilde{H}(\mathbf{q}', \mathbf{p}', t) = H(\mathbf{q}, \mathbf{p}, t) + \frac{\partial S}{\partial t} = 0 \tag{10.5}$$

If that holds, the Hamiltonian equations of motion can be used, and the following simplified results for transformed position and momentum are obtained:

$$\dot{q}'_k = \frac{\partial \tilde{H}}{\partial p'_k} = 0 \implies q'_k = \text{const.}$$

$$\dot{p}'_k = -\frac{\partial \tilde{H}}{\partial q'_k} = 0 \longrightarrow p'_k = \text{const.}$$

q'_k and $p'_k = \alpha$ are thus both constants of motion and are determined by the initial condition of the system. With the obtained momentum, $S(\mathbf{q}, \mathbf{p}', t) = S(\mathbf{q}, \boldsymbol{\alpha}, t)$. The total derivative of this generating function yields:

$$\frac{d}{dt} S(\mathbf{q}, \boldsymbol{\alpha}, t) = \frac{\partial S(\mathbf{q}, \boldsymbol{\alpha}, t)}{\partial \mathbf{q}} d\mathbf{q} + \frac{\partial S(\mathbf{q}, \boldsymbol{\alpha}, t)}{\partial t} = \mathbf{p} d\mathbf{q} - H(\mathbf{q}, \mathbf{p}, t) = L(\mathbf{q}, \dot{\mathbf{q}}, t)$$

Integrating both sides over time, it is recognized that the generating function S corresponds to the action:

$$S(\mathbf{q}, \boldsymbol{\alpha}, t) = \int dt\, L(\mathbf{q}, \dot{\mathbf{q}}, t) \tag{10.6}$$

If a particle is considered in three-dimensional space, the momentum from (10.4) can be written as:

$$\mathbf{p} = \boldsymbol{\nabla} S$$

According to (10.3), in classical mechanics, for the Hamiltonian function $H(\mathbf{r}, \mathbf{p}, t) = H(\mathbf{r}, \boldsymbol{\nabla} S, t)$ and the action $S(\mathbf{r}, \boldsymbol{\alpha}, t)$. The classical Hamilton-Jacobi equation (10.5) can then be written as

$$H(\mathbf{r}, \mathbf{p}, t) + \frac{\partial S}{\partial t} = \frac{1}{2m}(\boldsymbol{\nabla} S)^2 + V(\mathbf{r}) + \frac{\partial S}{\partial t} = 0 \tag{10.7}$$

Now we move on to quantum theory: Bringing all terms to one side, the three-dimensional Schrödinger equation (2.8) in quantum mechanics can be written as follows:

$$-\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2\psi(\mathbf{r}, t) + V(\mathbf{r})\psi(\mathbf{r}, t) - i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r}, t) = 0$$

The wave function of a particle can be described as a superposition of plane matter waves, as shown here:

$$\psi(\mathbf{r}, t) = Ae^{i(\mathbf{kr}-\omega t)} = A\exp\left(\frac{iS}{\hbar}\right) \tag{10.8}$$

For the action S, known in wave optics as "eikonal", holds:

$$S = \mathbf{p} \cdot \mathbf{r} - Et \tag{10.9}$$

The wave function from (10.8) can now be inserted into the Schrödinger equation:

$$\left[-\frac{-\hbar^2}{2m}\boldsymbol{\nabla}^2 + V(\mathbf{r})\right]Ae^{\frac{iS}{\hbar}} = i\hbar\frac{\partial}{\partial t}Ae^{\frac{iS}{\hbar}}$$

The constant factors A cancel out. The time derivative on the right-hand side of the equation can be carried out as:

$$i\hbar\frac{\partial}{\partial t}e^{\frac{iS}{\hbar}} = i\hbar\left(\frac{i}{\hbar}\frac{\partial S}{\partial t}\right)e^{\frac{iS}{\hbar}} = -\dot{S}e^{\frac{iS}{\hbar}}$$

On the left side, only the derivative term is considered, as $\psi(\mathbf{r}, t)$ does not change due to the potential:

$$-\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2 Ae^{\frac{iS}{\hbar}} = -\frac{\hbar^2}{2m}\boldsymbol{\nabla}\left[\frac{i}{\hbar}\boldsymbol{\nabla} Se^{\frac{iS}{\hbar}}\right] = -\frac{i\hbar}{2m}\left[\boldsymbol{\nabla}^2 S + \frac{i}{\hbar}(\boldsymbol{\nabla} S)^2\right]e^{\frac{iS}{\hbar}}$$

The exponential functions cancel out, and we obtain:

$$\frac{1}{2m}(\boldsymbol{\nabla} S)^2 - \frac{i\hbar}{2m}\Delta S + V(\mathbf{r}) = -\dot{S}$$

The result can be brought to the form of the classical Hamilton-Jacobi equation (10.7):

$$\frac{1}{2m}(\boldsymbol{\nabla} S)^2 + V(\mathbf{r}) + \frac{\partial S}{\partial t} = \frac{i\hbar}{2m}\Delta S \tag{10.10}$$

As can be seen, we obtain an additional term $\frac{i\hbar}{2m}\Delta S$ on the right side, which is referred to as the *quantum correction*. Letting $\hbar$ approach zero ($\hbar \to 0$) transitions from quantum physics to classical physics, and we obtain the classical Hamilton-Jacobi equation (10.7) again.

10.2 Hermite Differential Equation and Polynomials

We were able in the chapter on the harmonic oscillator to reduce the Schrödinger equation in the oscillator potential to the Hermite differential equation in (4.37); however, we did not solve it explicitly but instead used the Rodrigues formula for Hermite polynomials in (4.38) to anticipate the solution directly. We will discuss one possible solution approach here in more detail and examine the properties of Hermite polynomials more closely.

10.2.1 Derivation of the Rodrigues Formula for Hermite Polynomials

To derive the Rodrigues formula, we first start with a simple approach: $u(y) = e^{-y^2}$. Differentiating this expression with respect to y gives us a differential equation for the known solution:

$$u' = -2yu \Longrightarrow u' + 2yu = 0$$

Differentiating this differential equation $(n+1)$ times leads us to an increasingly complex expression:

$$u^{(n+2)} + 2(yu)^{(n+1)} = 0$$

To explicitly perform the differentiation of the product term $(yu)^{(n+1)}$, we apply the Leibniz formula – we obtain a sum, which terminates already at the second term, because y only appears linearly in $(yu)^{(n+1)}$:

$$(y \cdot u)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} y^{(k)} u^{(n-k)} = yu^{(n)} + (n+1)u^{(n-1)}$$

Substituting $(n+1)$ instead of n in the Leibniz formula, we can directly insert into the above expression. Furthermore, we substitute $\xi(y) = (-1)^n u^{(n)}$ (the term $(-1)^n$ is only inserted *ad hoc* here to obtain the desired solution structure):

$$u^{(n+2)} + 2yu^{(n+1)} + (n+1)u^{(n)} = 0 \Longrightarrow \xi'' + 2y\xi'' + 2(n+1)\xi = 0$$

We carry out another substitution $\xi = \zeta e^{-y^2}$, where we directly insert all derivatives and cancel the Gaussian function:

$$\left[\zeta'' - 4y\zeta' + 4y^2\zeta - 2\zeta\right] + 2y\left[\zeta' - 2y\zeta\right] + 2(n+1)\zeta = 0 \Longrightarrow \zeta'' - 2y\zeta' + 2n\zeta = 0$$

Now we compare the obtained differential equation with the required Hermitian differential equation:

$$h''(y) - 2yh'(y) + 2nh(y) = 0 \tag{10.11}$$

The two expressions match! Thus, $\zeta(y) = h(y) = H_n(y)$, and by substituting all substitutions back into $\zeta(y)$, we obtain the *Rodrigues formula*:

$$H_n(y) = (-1)^n e^{y^2} \frac{d^n}{dy^n} e^{-y^2} \tag{10.12}$$

With this, all Hermite polynomials can be calculated very simply; the first five $H_n(y)$ are formally given below:

$$\begin{aligned} H_0(y) &= 1 \\ H_1(y) &= 2y \\ H_2(y) &= 4y^2 - 2 \\ H_3(y) &= 8y^3 - 12y \\ H_4(y) &= 16y^4 - 48y^2 + 12 \end{aligned} \tag{10.13}$$

In Figure 53, the polynomials from (10.13) are plotted: considering the nodal rule of wave functions, this behavior can already be recognized here.

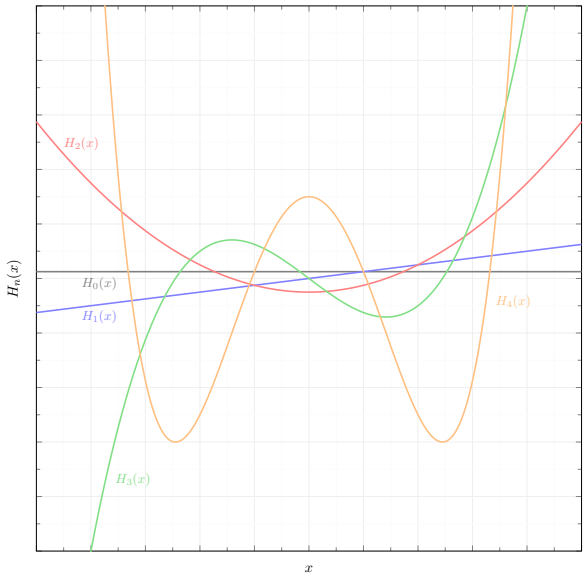


Fig. 53: The first five Hermite polynomials.

10.2.2 Recursive Representation of Hermite's Differential Equation

To represent Hermite's differential equation as in (10.11), we can also find a recursive form. We start from the Rodrigues formula (10.12): the index of $H_n(y)$ represents both the energy quantum number of the harmonic oscillator and the degree of differentiation in the formula. If we differentiate (10.12) again, we get:

$$\begin{aligned}\frac{d}{dy}H_n(y) &\overset{(10.12)}{=} \frac{d}{dy}\left[(-1)^n e^{y^2}\frac{d^n}{dy^n}e^{-y^2}\right] = \\ &= (-1)^n\left[2ye^{y^2}\frac{d^n}{dy^n}e^{-y^2} + e^{y^2}\frac{d^{n+1}}{dy^{n+1}}e^{-y^2}\right] = \\ &= 2yH_n(y) - H_{n+1}(y)\end{aligned}$$

The minus sign results from the need to adjust the prefactor $(-1)^n$ to the higher degree of the polynomial $H_{n+1}(y)$. We can thus write:

$$H_{n+1}(y) = \left[2y - \frac{d}{dy}\right]H_n(y) \tag{10.14}$$

Starting from $H_0(y) = 1$, this recursive representation derives all Hermite polynomials from (10.13). By substituting the preceding $H_n(y)$ into $H_{n+1}(y)$, we obtain:

$$H_n(y) = \left[2y - \frac{d}{dy}\right]^n \cdot 1 \tag{10.15}$$

The term $\cdot 1$ indicates the first Hermite polynomial. Regarding the algebraic solution for the Schrödinger equation of the harmonic oscillator, we can make another modification to (10.14):

$$H_{n+1}(y) = e^{\frac{1}{2}y^2}\left[y - \frac{d}{dy}\right]^n e^{-\frac{1}{2}y^2}H_n(y) \tag{10.16}$$

By explicitly setting $H_n(y)$ as the Hermite polynomial (again $H_0(y) = 1$), we see that another explicit expression arises from the recursive form:

$$H_n(y) = e^{\frac{1}{2}y^2}\left[y - \frac{d}{dy}\right]^n e^{-\frac{1}{2}y^2} \tag{10.17}$$

10.2.3 Symmetry

It can be easily shown that the Hermite polynomials, depending on their degree, are either symmetric or antisymmetric (this was already required at the beginning of the chapter due to the symmetry of the oscillator potential). By substituting $y \to -y$, we get:

$$H_n(-y) \overset{(10.12)}{=} (-1)^n e^{(-y)^2} \frac{d^n}{d(-y)^n} e^{-(-y)^2} = (-1)^n e^{y^2} (-1)^n \frac{d^n}{dy^n} e^{-y^2} = (-1)^n H_n(y)$$

10.2.4 Orthogonality

The Hermite polynomials are orthogonal in a weighted Hilbert space $L_2(\mathbb{R}, \exp\{-y^2\}dy)$; this means that scalar products can only be formed with the corresponding Gaussian weight function. We require that $n < m$ holds and thus obtain:

$$\begin{aligned}
\langle H_n(y) H_m(y) \rangle &= \int_{\mathbb{R}} dy\, e^{-x^2} H_n(y) H_m(y) = \\
&= (-1)^m \int_{\mathbb{R}} dy\, e^{-x^2} e^{+x^2} H_n(y) \frac{d^m}{dy^m} \left(e^{-x^2} \right) = \\
&= (-1)^m \left[H_n(y) \frac{d^{m-1}}{dy^{m-1}} \left(e^{-x^2} \right) \Big|_{\mathbb{R}} - \int_{\mathbb{R}} dy\, \frac{dH_n(y)}{dy} \frac{d^{m-1}}{dy^{m-1}} \left(e^{-x^2} \right) \right] = \\
&= (-1)^m (-1)^{n+1} \int_{\mathbb{R}} dy\, \frac{d^{n+1} H_n(y)}{dy^{n+1}} \frac{d^{m-n-1}}{dy^{m-n-1}} \left(e^{-x^2} \right) = 0
\end{aligned}$$

The first term of the partial integration vanishes because the Gaussian function retains its form even after $m-1$ derivations and falls off at the boundaries faster than any polynomial. We perform the partial integration $(n+1)$ times, but in the last step, we derive a polynomial of degree n. The term vanishes, and the orthogonality is proven.

10.2.5 Normalizability

It holds that $n = m$, and we directly perform all calculation steps of the upper derivative:

$$\begin{aligned}
\langle H_n(y) H_n(y) \rangle &= \int_{\mathbb{R}} dy\, e^{-x^2} H_n(y) H_n(y) = (-1)^n (-1)^n \int_{\mathbb{R}} dy\, \frac{d^n H_n(y)}{dy^n} e^{-x^2} = \\
&= \frac{d^n H_n(y)}{dy^n} \int_{\mathbb{R}} dy\, e^{-x^2} = 2^n n! \sqrt{\pi}
\end{aligned} \tag{10.18}$$

Considering (10.15), we recognize that $H_n(y) \propto (2y)^n$. The n-th derivative is therefore independent of y and can be taken out of the integral; it holds $H_n^{(n)}(y) = 2^n n!$. The integration over the Gaussian function is a standard integral and yields the term $\sqrt{\pi}$.

10.3 Legendre Differential Equation and Polynomials

At this point, we concern ourselves with deriving the Rodrigues formula for the Legendre differential equation, considering not only the ordinary case but also the associated Legendre differential equation. To maintain clarity and generality, we keep the substitution $u = \cos(\vartheta)$. The Legendre differential equation has the following form:

$$\begin{aligned}
0 &= \frac{d}{du} \left[(1-u^2) f' \right] + \left[l(l+1) - \frac{m^2}{1-u^2} \right] f = \\
&= (1-u^2) f'' - 2u f' + \left[l(l+1) - \frac{m^2}{1-u^2} \right] f
\end{aligned}$$

10.3.1 Legendre Polynomials

First, we solve the Legendre differential equation for the case $m = 0$. We start with an initial guess and will cleverly manipulate this guess to arrive at the desired equation. We define:

$$f \equiv f(u) = (1 - u^2)^n$$

The guess $(1 - u^2)$ can be justified by the prefactor in the second derivative from (5.119); the exponent is used to allow multiple differentiations. We differentiate $f(u)$ once and obtain a first differential equation:

$$f' = -2nu(1 - u^2)^{n-1} \implies (1 - u^2)f' + 2nuf = 0$$

Differentiating again brings us to a form that already shows similarities with the required equation (5.119). We finally differentiate the second-order differential equation n more times:

$$\begin{aligned} 0 &= (1 - u^2)f'' + 2(n - 1)uf' + 2nf = \\ &= [(1 - u^2)f'']^{(n)} + [2(n - 1)uf']^{(n)} + 2nf^{(n)} \end{aligned}$$

Using the Leibniz formula, we can further simplify the non-trivial terms; the polynomial term also limits the number of possible differentiations:

$$\begin{aligned} [(1 - u^2)f'']^{(n)} &= \sum_{k=0}^{n} \binom{n}{k} (1 - u^2)^{(k)} f^{(n+2-k)} = \\ &= (1 - u^2)f^{(n+2)} - 2nuf^{(n+1)} - n(n - 1)f^{(n)} \\ [2(n - 1)uf']^{(n)} &= 2(n - 1) \sum_{k=0}^{n} \binom{n}{k} u^{(k)} f^{(n+1-k)} = \\ &= 2(n - 1)uf^{(n+1)} + 2n(n - 1)f^{(n)} \end{aligned}$$

By inserting our found expressions into the processed guess, we recognize that it now almost corresponds to the Legendre differential equation (5.119):

$$(1 - u^2)f^{(n+2)} - 2uf^{(n+1)} + n(n + 1)f^{(n)} = 0$$

Only the n-th derivatives still differ from the required form. We substitute one last time:

$$P_n \equiv P_n(u) = \frac{d^n}{du^n} f(u)$$

Only now do our guess and (5.119) fully agree – using the solution, we have reconstructed the searched-for differential equation:

$$(1 - u^2)P_n'' - 2uP_n' + n(n + 1)P_n' = 0 \tag{10.19}$$

We call our solution functions *Legendre polynomials*. These are generated by the following Rodrigues formula:

$$P_n(u) = \frac{d^n}{du^n}(1 - u^2)^n \tag{10.20}$$

10.3.2 Associated Legendre Polynomials

So far, we have only considered the special case $m = 0$; when we drop this restriction and allow $m \in \mathbb{Z}$, we can derive the complete Legendre differential equation from (5.119). In reference to the solution of the ordinary Legendre differential equation, we replace $n \to l$ in (10.20).

We start with the Ansatz from our previous result in (10.20). By inserting the solution into the full associated Legendre differential equation and differentiating m times, we obtain:

$$\begin{aligned} 0 &= (1-u^2)P_l'' - 2uP_l' + l(l+1)P_l = \quad \mid \partial_x^m \\ &= [(1-u^2)P_l'']^{(m)} - [2uP_l']^{(m)} + l(l+1)P_l^{(m)} \end{aligned}$$

We again apply the Leibniz formula to differentiate the multiplicative terms sensibly:

$$\begin{aligned} [(1-u^2)P_l'']^{(m)} &= \sum_{k=0}^{m} \binom{m}{k} (1-u^2)^{(k)} P_l^{(m+2-k)} = \\ &= (1-u^2)P_l^{(m+2)} - 2muP_l^{(m+1)} - m(m-1)P_l^{(m)} \\ [2uf']^{(m)} &= 2\sum_{k=0}^{m} \binom{m}{k} u^{(k)} P_l^{(m+1-k)} = 2uP_l^{(m+1)} + 2mP_l^{(m)} \end{aligned}$$

The results thus obtained are inserted into the Ansatz of the differential equation, resulting in:

$$(1-u^2)P_l^{(m+2)} - 2(m+1)uP_l^{(m+1)} + [l(l+1) - m(m+1)]P_l^{(m)} = 0$$

The prefactor of the first derivative contains a term $(m+1)$, which does not match with (5.119)! To remove this, we multiply the entire equation by $(1-u^2)^m$. This allows us to pull together the previously separate terms $P_l^{(m+2)}$ and $P_l^{(m+1)}$, enabling a more compact notation:

$$\frac{d}{du}\left[(1-u^2)^{m+1}P_l^{(m+1)}\right] + [l(l+1) - m(m+1)](1-u^2)^m P_l^{(m)}$$

We substitute again. The Ansatz is by no means trivial and seems "reasonable" at first glance, but it will lead us to the desired result. We assert:

$$\xi \equiv \xi(u) = (1-u^2)^{m/2} P_l$$

However, for us, the first derivative of $P_l^{(m+1)}$ is also relevant, leading to a more or less complex expression:

$$P_l^{(m+1)} = \frac{d}{du}(1-u^2)^{-m/2}\xi = mu(1-u^2)^{-m/2-1}\xi + (1-u^2)^{-m/2}\xi'$$

To evaluate the entire term in the square bracket, we substitute the expression found above:

$$\begin{aligned} \frac{d}{du}\left[(1-u^2)^{m+1}P_l^{(m+1)}\right] =& \frac{d}{du}\left[mu(1-u^2)^{\frac{m}{2}}\xi + (1-u^2)^{\frac{m+2}{2}}\xi'\right] = \\ =& m(1-u^2)^{\frac{m}{2}}\xi - m^2u^2(1-u^2)^{\frac{m-1}{2}}\xi + mu(1-u^2)^{\frac{m}{2}}\xi' - \\ & - (m+2)u(1-u^2)^{\frac{m}{2}}\xi' + (1-u^2)^{\frac{m+2}{2}}\xi'' = \\ =& (1-u^2)^{\frac{m}{2}}\Big\{\left[m - m^2u^2(1-u^2)^{-1}\right]\xi + \\ & + [mu - (m+2)u]\,\xi' + (1-u^2)\xi''\Big\} \\ =& (1-u^2)^{\frac{m}{2}}\Big\{\left[m - mu^2 - m^2u^2\right](1-u^2)^{-1}\xi - \\ & - 2u\xi' + (1-u^2)\xi''\Big\} \end{aligned}$$

We can insert the results obtained into our original differential equation and finally obtain an expression that almost matches the result.

$$(1-u^2)^{\frac{m}{2}}\left\{(1-u^2)\xi'' - 2u\xi' + l(l+1)\xi + \left[\frac{m - mu^2 - m^2u^2}{1-u^2} - m(m+1)\right]\xi\right\} = 0$$

Simplifying the fractional term and canceling out $(1-u^2)^{\frac{m}{2}}$ finally results in the associated Legendre differential equation (5.119):

$$(1-u^2)\xi'' - 2u\xi' + \left[l(l+1) - \frac{m^2}{1-u^2}\right]\xi = 0 \tag{10.21}$$

By performing all our substitutions backwards, the associated Legendre polynomials can now be explicitly given in the Rodrigues representation:

$$P_l^m(u) = (1-u^2)^{\frac{m}{2}} \frac{d^{l+m}}{du^{l+m}}(1-u^2)^n \tag{10.22}$$

It should be noted that in literature, an additional factor $1/(2^l l!)$ appears – this serves to ensure that the first Legendre polynomial $P_0^0 = 1$. We will find this factor in the next section in the form of normalization.

l	m	$P_l^m(x)$
0	0	1
1	-1	$\sqrt{1-x^2}/2$
	0	x
	+1	$-\sqrt{1-x^2}$
2	-2	$3(1-x^2)/24$
	-1	$x\sqrt{1-x^2}/2$
	0	$(3x^2-1)/2$
	+1	$-3x\sqrt{1-x^2}$
	+2	$3(1-x^2)$

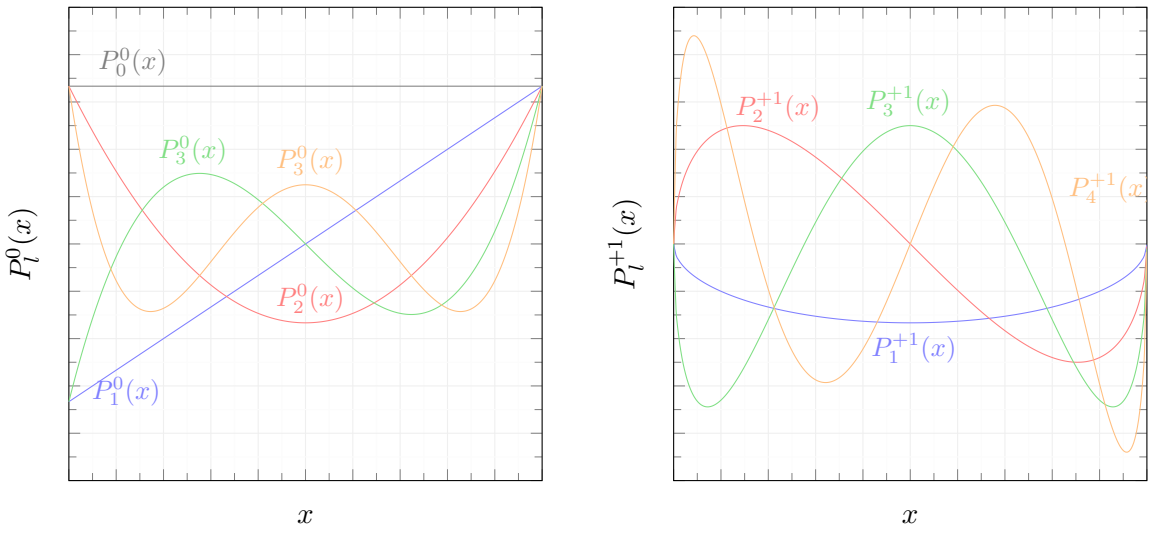


Fig. 54: (left) Associated Legendre polynomials for $m = 0$ for $l = 0, \ldots, 4$. (right) Associated Legendre polynomials for $m = +1$ for $l = 1, \ldots, 4$.

10.3.3 Orthogonality and Normalization

To form an orthonormal system, the spherical harmonics (and thus the associated Legendre polynomials) must be orthogonal and normalized. We will investigate this behavior using P_k^m

and P_l^m. However, first, we perform an auxiliary calculation and evaluate the following integral via a recurrence relation:

$$\begin{aligned}
\mathbb{I}_n &= \int_{-1}^{+1} dx\,(1-x^2)^n = \int_{-1}^{+1} dx\,1\cdot(1-x^2)^n = \\
&= \int_{-1}^{+1} dx\,\left\{\frac{d}{dx}\left[x(1-x^2)^n\right] - x\frac{d}{dx}(1-x^2)^n\right\} = \\
&= \left[x(1-x^2)^n\right]_{-1}^{+1} + 2n\int_{-1}^{+1} dx\,x^2(1-x^2)^{n-1} = \\
&= 2n\int_{-1}^{+1} dx\,[1-(1-x^2)](1-x^2)^{n-1} = \\
&= 2n\left[\int_{-1}^{+1} dx\,(1-x^2)^{n-1} - \int_{-1}^{+1} dx\,(1-x^2)^n\right] = 2n\left[\mathbb{I}_{n-1} - \mathbb{I}_n\right]
\end{aligned}$$

We reinstate into the result found an additional recurrence relation for $\mathbb{I}_{n-1}$, obtaining iteratively:

$$\begin{aligned}
\mathbb{I}_n &= \frac{2n}{2n+1}\mathbb{I}_{n-1} = \frac{(2n)(2n-2)}{(2n+1)(2n-1)}\mathbb{I}_{n-2} = \frac{2^n n!}{(2n+1)!!}\mathbb{I}_0 = \\
&= 2^n n!\frac{2^n n!}{(2n+1)!}\cdot 2 = 2^{2n+1}\frac{n!n!}{(2n+1)!}
\end{aligned} \tag{10.23}$$

We will use this expression later. For the sake of avoiding excessive notation, we substitute $v = (1-u^2)$ and calculate the product between the polynomials P_k^m and P_l^m:

$$\begin{aligned}
\frac{1}{N^2} &= \int_{-1}^{1} du\,P_k^m P_l^m = \int_{-1}^{1} du\,v^m\left(\partial_u^{k+m}v^k\right)\left(\partial_u^{l+m}v^l\right) = \\
&= v^m\left(\partial_u^{k+m}v^k\right)\left(\partial_u^{l+m-1}v^l\right)\Big|_{-1}^{1} - \int_{-1}^{1} du\,\left(\partial_u^{l+m-1}v^l\right)\partial_u\left[v^m\left(\partial_u^{k+m}v^k\right)\right] = \\
&= (-1)^1\int_{-1}^{1} du\,\left(\partial_u^{l+m-1}v^l\right)\partial_u\left[v^m\left(\partial_u^{k+m}v^k\right)\right] = \quad |\times(l+m) \\
&= (-1)^{l+m}\int_{-1}^{1} du\,v^l\,\partial_u^{l+m}\left[v^m\left(\partial_u^{k+m}v^k\right)\right] = (-1)^{l+m}\int_{-1}^{1} du\,v^l\mathbb{D}(k;l)
\end{aligned}$$

We perform the partial integration $(l+m)$ times to create a polynomial completely free from derivation. The expression in front of the integral disappears on each iteration because $v = (1-u^2)$ vanishes at $u = +1$ and $u = -1$! We calculate the final derivation term $\partial_u^{l+m}[\ldots]$ separately:

$$\begin{aligned}
\mathbb{D}(k;l) &= \partial_u^{l+m}\left[(1-u^2)^m\left(\partial_u^{k+m}(1-u^2)^k\right)\right] \propto \quad |\,k=l \\
&\propto \partial_u^{l+m}\left[(-1)^m u^{2m}\left(\partial_u^{l+m}(-1)^l u^{2l}\right)\right] = \\
&= (-1)^{l+m}\partial_u^{l+m}\left[u^{2m}\frac{(2k)!}{(l-m)!}u^{l-m}\right] = \\
&= (-1)^{l+m}\frac{(2k)!}{(l-m)!}\partial_u^{l+m}\left[u^{l+m}\right] = \\
&= (-1)^{l+m}(2l)!\frac{(l+m)!}{(l-m)!}
\end{aligned} \tag{10.24}$$

We require that $l \geq k$ must hold (in the reverse case, we would have needed to integrate partially differently)! The entire expression in the square bracket has a polynomial degree $\deg[\ldots] = k+m$, thus it is less than the degree of derivation. At this point, we can demand that $k = l$ holds, and P_k^m and P_l^m are orthogonal to each other; if the polynomial and derivation

degrees are equal, we only need to consider the term with the largest exponent, as all other terms vanish due to derivation.
Inserting (10.23) and (10.24) into the original integral we obtain a rounded expression for the normalization of the associated Legendre polynomials:

$$\begin{aligned}\frac{1}{N^2} &= \int_{-1}^{1} du\, P_l^m P_l^m = (-1)^{l+m} \int_{-1}^{1} du\, v^l \partial_u^{l+m} \left[v^m \left(\partial_u^{l+m} v^l \right) \right] \overset{(10.24)}{=} \\ &= (-1)^{l+m}(-1)^{l+m}(2l)! \frac{(l+m)!}{(l-m)!} \int_{-1}^{1} du\, (1-u^2)^l \overset{(10.23)}{=} \\ &= (-1)^{2(l+m)}(2l)! \frac{(l+m)!}{(l-m)!} 2^{2l+1} \frac{l!l!}{(2l+1)!} = \\ &= \left[(-1)^{(l+m)} 2^l l! \right]^2 \frac{2}{2l+1} \frac{(l+m)!}{(l-m)!} \end{aligned} \tag{10.25}$$

Returning to the spherical harmonics from (5.114); in addition to the normalization of the Legendre polynomials (10.25), we also have to account for the normalization of the φ-dependent part in (5.117). It follows:

$$\frac{1}{N^2} = \int_0^{2\pi} d\varphi\, e^{-im\varphi} e^{+im\varphi} = 2\pi \implies N = \frac{1}{\sqrt{2\pi}} \tag{10.26}$$

Together, (10.25) and (10.26) now provide the complete normalization of the spherical harmonics:

$$N = \frac{(-1)^{(l+m)}}{2^l l!} \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} \tag{10.27}$$

10.4 Laguerre Differential Equation and Polynomials

Analogous to the Hermite and Legendre differential equations, we aim to investigate the Laguerre differential equation and its solutions here. With the quantities $\beta \geq 0$ and $w \in \mathbb{R}$, the equation is:

$$xy'' + (\beta + 1 - x)y' + wy = 0 \tag{10.28}$$

To encounter the Rodrigues formula of the Laguerre differential equation, we will again begin with an Ansatz; we conduct our derivation here directly for the general case $\beta \geq 0$, while $\beta > 0$ is referred to as the *associated Laguerre differential equation.*

10.4.1 Rodrigues Formula of the Laguerre Differential Equation

Similar to the other cases, we begin from the "solution" and derive it through successive differentiation back to our desired differential equation. We set:

$$y(x) \equiv y = e^{-x}\, x^{w+\beta}$$

Let's start our derivation by differentiating the upper equation once and transforming it:

$$y' = -e^{-x}\, x^{w+\beta} + (w+\beta) e^{-x}\, x^{w+\beta-1} = \left(\frac{w+\beta}{x} - 1 \right) y \implies xy' + (x - w - \beta)y = 0$$

We obtain a first differential equation that, however, does not yet resemble the required form from (10.28). We differentiate again:

$$xy'' + (x - w - \beta + 1)y' + y = 0$$

Each of these terms will be differentiated another w times; to evaluate the product terms, we apply the Leibniz formula:

$$\frac{d^w}{x^w}xy'' = \sum_{k=0}^{w}\binom{w}{k}x^{(k)}y^{(w+2-k)} = xy^{(w+2)} + wy^{(w+1)}$$
$$\frac{d^w}{x^w}(x-w-\beta+1)y' = (x-w-\beta+1)y^{(w+1)} + ny^{(w)}$$

Inserting the results of our auxiliary calculation into the differential equation, we can again substitute: $k_w = y^{(w)}$. This yields the following new equation:

$$xy^{(w)} + (x-\beta+1)y^{(w+1)} + (w)y^{(w)} = k_w'' + (x-\beta+1)k_w' + (w+1)k_w = 0$$

While the second derivative term in this form already aligns, the prefactors of the first and zeroth derivatives still present problems. Therefore, we substitute again:

$$k_w(x) \equiv k_w = x^\beta e^{-x} L_w^\beta$$

The first and second derivatives of k_w are somewhat laborious to calculate, hence, without delving into technical details, we ahead the result:

$$k_w' = \left[\frac{\beta}{x}L_w^\beta - L_w^\beta + L_w^{\beta\prime}\right]x^\beta e^{-x}$$
$$k_w'' = \left[L_w^{\beta\prime\prime} + 2L_w^{\beta\prime}\left(\frac{\alpha}{x}-1\right) + L_w^\beta\left(\frac{\alpha(\alpha-1)}{x^2} - \frac{2\alpha}{x}+1\right)\right]x^\beta e^{-x}$$

Inserting these results into the obtained differential equation and canceling, we obtain:

$$xL_w^{\beta\prime\prime} + (\beta+1-x)L_w^{\beta\prime} + wL_w^\beta = 0$$

This corresponds exactly to the Laguerre differential equation! Thus, we have found a solution formula with which all solutions L_w^β, (so-called "associated Laguerre polynomials"), can be constructed. This formula is known as the Rodrigues formula. By reversing all substitutions, we obtain (without any particular normalization):

$$L_w^\beta(x) \equiv L_w^\beta = x^{-\beta}e^x\frac{d^w}{dx^w}\left(x^{w+\beta}e^{-x}\right) \tag{10.29}$$

A commonly used method for the simple generation of associated Laguerre polynomials L_w^β is the following summation formula:

$$L_w^\beta(x) = \sum_{m=0}^{w}(-1)^m\frac{(w+\beta)!}{(w-m)!(\beta+m)!m!}x^m \tag{10.30}$$

This summation formula corresponds to the Rodrigues formula (10.29), provided it is supplemented with an additional prefactor $1/w!$:

$$L_w^\beta(x) = \frac{1}{w!}x^{-\beta}e^x\frac{d^w}{dx^w}\left(x^{w+\beta}e^{-x}\right) \tag{10.31}$$

10.4.2 Orthogonality and Normalizability

The Laguerre polynomials are only normalizable when a weighting function $g(x) = x^\alpha e^{-x}$ is used. Substituting our Rodrigues formula; we will replace $\xi_\tau(x) \equiv \xi_\tau = x^{\tau+\beta}e^{-x}$ to shorten the following expressions. This yields:

$$\begin{aligned}
\frac{1}{N^2} &= \int_0^\infty dx\, x^\alpha e^{-x} L_v^\beta L_w^\beta = \int_0^\infty dx\, x^\alpha e^{-x} x^{-\beta} e^x \xi_v^{(v)} x^{-\beta} e^x \xi_w^{(w)} = \\
&= \int_0^\infty dx\, x^{-\beta} e^x \xi_v^{(v)} \xi_w^{(w)} = \\
&= x^{-\beta} e^x \xi_v^{(v)} \xi_w^{(w-1)} \Big|_0^\infty - \int_0^\infty dx\, \xi_w^{(w-1)} \left(x^{-\beta} e^x \xi_v^{(v)}\right)^{(1)} = \;|\times w \\
&= (-1)^w \int_0^\infty dx\, \xi_w \left(x^{-\beta} e^x \xi_v^{(v)}\right)^{(w)} = (-1)^w \int_0^\infty dx\, x^{w+\beta} e^{-x}\, \mathbb{D}(w;v)
\end{aligned}$$

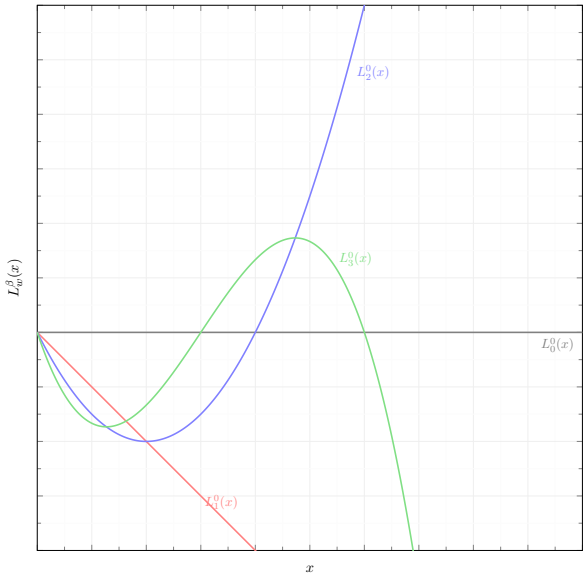


Fig. 55: A selection of Laguerre polynomials each at $\beta = 0$.

$\mathbb{D}(w;v)$ corresponds to a constant, as shown in the lower auxiliary calculation; the remaining integral can be equated with the Γ function: $\Gamma(w+\beta+1)$. For the nested derivation, we finally obtain, assuming $w > v$:

$$\begin{aligned}
\mathbb{D}(w;v) &= \partial_x^w \left[x^{-\beta} e^x \partial_x^v \left(x^{v+\beta} e^{-x}\right)\right] = \\
&= \partial_x^w \left[x^{-\beta} e^x \sum_{k=0}^{v} \binom{v}{k} \left(x^{v+\beta}\right)^{(k)} \left(e^{-x}\right)^{(v-k)}\right] = \\
&= (-1)^v \partial_x^w \left[x^{-\beta} e^x e^{-x} \sum_{k=0}^{v} \binom{v}{k} \left(x^{v+\beta}\right)^{(k)} (-1)^{-k}\right] = \\
&= (-1)^v \sum_{k=0}^{v} \binom{v}{k} (-1)^k \partial_x^w \left[x^{-\beta} \frac{(v+\beta)!}{(v+\beta-k)!} x^{v+\beta-k}\right] = \\
&= (-1)^v \sum_{k=0}^{v} \binom{v}{k} (-1)^k \frac{(v+\beta)!}{(v+\beta-k)!} \partial_x^w x^{v-k} = \;|v = w \\
&= (-1)^v \binom{v}{0} (-1)^0 \frac{(v+\beta)!}{(v+\beta-0)!} \partial_x^w x^w = (-1)^v w!
\end{aligned}$$

Evaluating the above integral gives us the following normalization: $N = (w!\,\Gamma(w+\beta+1))^{-1/2}$. Altogether, we can now represent the normalized Rodrigues formula of the Laguerre polynomials as:

$$\tilde{L}_w^\beta(x) = \frac{1}{\sqrt{w!\,\Gamma(w+\beta+1)}} \frac{d^w}{dx^w} x^{-\beta} e^x \left(x^{w+\beta} e^{-x} \right) \tag{10.32}$$

11 Useful Relations

In the following table, a selection of useful eigenvalue relations of operators from this script is given; found are, in addition to the notation used for the operator, the eigenstate as well as the actual eigenvalue equation:

$\hat{A}$	Operator	$\lvert a\rangle$	Eigenstate	Effect
Common Operators				
$\hat{\mathbf{x}}$	Position	$\lvert\mathbf{x}\rangle$	Position	$\hat{\mathbf{x}}\lvert\mathbf{x}\rangle = \mathbf{x}\lvert\mathbf{x}\rangle$
$\hat{\mathbf{p}}$	Momentum	$\lvert\mathbf{p}\rangle$	Momentum	$\hat{\mathbf{p}}\lvert\mathbf{p}\rangle = \mathbf{p}\lvert\mathbf{p}\rangle$
$\hat{H}$	Hamiltonian	$\lvert\psi_n\rangle$	Energy	$\hat{H}\lvert\psi_n\rangle = E_n\lvert\psi_n\rangle$
$\hat{\Pi}$	Parity	$\lvert\varphi_\pi\rangle$	Parity	$\hat{\Pi}\lvert\varphi_\pi\rangle = p_\pi\lvert\varphi_\pi\rangle$
H. Oscillator				
$\hat{N}$	Number State	$\lvert n\rangle$	H. Oscillator	$\hat{N}\lvert n\rangle = n\lvert n\rangle$
$\hat{a}^\dagger$	Raising	$\lvert n\rangle$	H. Oscillator	$\hat{a}^\dagger\lvert n\rangle = \sqrt{n+1}\lvert n+1\rangle$
$\hat{a}$	Lowering	$\lvert n\rangle$	H. Oscillator	$\hat{a}\lvert n\rangle = \sqrt{n}\lvert n-1\rangle$
	State	$\lvert\alpha\rangle$	Glauber	$\hat{a}\lvert\alpha\rangle = \alpha\lvert\alpha\rangle$
Angular Momentum ($\hat{\mathbf{J}}$, $\hat{\mathbf{L}}$, $\hat{\mathbf{S}}$ with $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$)				
$\hat{\mathbf{J}}^2$	Magnitude Squared	$\lvert j, m_j\rangle$	Ang. Momentum	$\hat{\mathbf{J}}^2\lvert j, m_j\rangle = \hbar^2 j(j+1)\lvert j, m_j\rangle$
$\hat{J}_z$	z-Component	$\lvert j, m_j\rangle$	Ang. Momentum	$\hat{J}_z\lvert j, m_j\rangle = \hbar m_j\lvert j, m_j\rangle$
$\hat{J}_+$	Raising	$\lvert j, m_j\rangle$	Ang. Momentum	$\hat{J}_+\lvert j, m_j\rangle = \hbar\sqrt{j(j+1) - m(m+1)}\lvert j, m_j+1\rangle$
$\hat{J}_-$	Lowering	$\lvert j, m_j\rangle$	Ang. Momentum	$\hat{J}_-\lvert j, m_j\rangle = \hbar\sqrt{j(j+1) - m(m-1)}\lvert j, m_j-1\rangle$
Spin-$\frac{1}{2}$ ($\hat{\boldsymbol{\sigma}} = \frac{2}{\hbar}\hat{\mathbf{S}}$)				
$\hat{\boldsymbol{\sigma}}^2$	Magnitude Squared	$\lvert s, m_s\rangle$	Spin-$\frac{1}{2}$	$\hat{\boldsymbol{\sigma}}^2\lvert s, m_s\rangle = 3\lvert s, m_s\rangle$
$\hat{\sigma}_z$	z-Component	$\lvert s, m_s\rangle$	Spin-$\frac{1}{2}$	$\hat{\sigma}_z\lvert s, m_s\rangle = \pm\lvert s, m_s\rangle$
$\hat{\sigma}_+$	Raising	$\lvert s, m_s\rangle$	Spin-$\frac{1}{2}$	$\hat{\sigma}_+\lvert s, -\rangle = \lvert s, +\rangle$
$\hat{\sigma}_-$	Lowering	$\lvert s, m_s\rangle$	Spin-$\frac{1}{2}$	$\hat{\sigma}_-\lvert s, +\rangle = \lvert s, -\rangle$

The following table provides useful commutator relations: It should be noted that the commutator must *always* act on an underlying wave function.

$\hat{A}$	Operator	$\hat{B}$	Operator	Commutator
Common Operators				
$\hat{x}_i$	Position	$\hat{p}_j$	Momentum	$[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}$
Harmonic Oscillator				
$\hat{a}$	Lowering	$\hat{a}^\dagger$	Raising	$[\hat{a}, \hat{a}^\dagger] = 1$
$\hat{N}$	Number State	$\hat{a}$	Lowering	$[\hat{N}, \hat{a}] = -\hat{a}$
$\hat{N}$	Number State	$\hat{a}^\dagger$	Raising	$[\hat{N}, \hat{a}^\dagger] = \hat{a}^\dagger$
Angular Momentum				
$\hat{L}_i$	Angular Momentum	$\hat{V}_j$	Vector	$[\hat{L}_i, \hat{V}_j] = i\hbar\varepsilon_{ijk}V_k$
$\hat{L}_i$	Angular Momentum	$\hat{S}$	Scalar	$[\hat{L}_i, \hat{S}] = 0$
$\hat{L}_\pm$	Raising/Lowering	$\hat{L}_\mp$	Lowering/Raising	$[\hat{L}_\pm, \hat{L}_\mp] = \pm 2\hbar\hat{L}_z$
$\hat{L}_z$	z-Angular Momentum	$\hat{L}_\pm$	Raising/Lowering	$[\hat{L}_z, \hat{L}_\pm] = \pm\hbar\hat{L}_\pm$
$\hat{\mathbf{L}}^2$	Angular Momentum Magnitude	$\hat{L}_\pm$	Raising/Lowering	$[\hat{\mathbf{L}}^2, \hat{L}_\pm] = 0$