Theoretical Physics T2
Quantum Mechanics

Course of Lectures by
Reinhold A. Bertlmann

Script written by
Reinhold A. Bertlmann and Nicolai Friis

T2–Script of Sommersemester 2008

E-mail: Reinhold.Bertlmann@univie.ac.at
Homepage: http://homepage.univie.ac.at/Reinhold.Bertlmann/

Contents

Chapter 1

Wave–Particle Duality

1.1 Planck's Law of Black Body Radiation

1.1.1 Quantization of Energy

The foundation of quantum mechanics was laid in 1900 with Max Planck's discovery of the quantized nature of energy. When Planck developed his formula for black body radiation he was forced to assume that the energy exchanged between a black body and its thermal (electromagnetic) radiation is not a continuous quantity but needs to be restricted to discrete values depending on the (angular-) frequency of the radiation. Planck's formula can explain – as we shall see – all features of the black body radiation and his finding is phrased in the following way:

Proposition 1.1.1 *Energy is quantized and given in units of $\hbar\omega$.* $E = \hbar\omega$

Here ω denotes the angular frequency $\omega = 2\pi\nu$. We will drop the prefix "angular" in the following and only refer to it as the frequency. We will also bear in mind the connection to the wavelength λ given by $c = \lambda\nu$, where c is the speed of light, and to the period T given by $\nu = \frac{1}{T}$.

The fact that the energy is proportional to the frequency, rather than to the intensity of the wave – what we would expect from classical physics – is quite counterintuitive. The proportionality constant $\hbar$ is called *Planck's constant*:

$$\hbar \;=\; \frac{h}{2\pi} \;=\; 1,054 \times 10^{-34}\,\mathrm{Js} \;=\; 6,582 \times 10^{-16}\,\mathrm{eVs} \tag{1.1}$$

$$h \;=\; 6,626 \times 10^{-34}\,\mathrm{Js} \;=\; 4,136 \times 10^{-15}\,\mathrm{eVs}\,. \tag{1.2}$$

1.1.2 Black Body Radiation

A black body is by definition an object that completely absorbs all light (radiation) that falls on it. This property makes a black body a perfect source of thermal radiation. A very good realization of a black body is an oven with a small hole, see Fig. 1.1. All radiation that enters through the opening has a very small probability of leaving through it again.

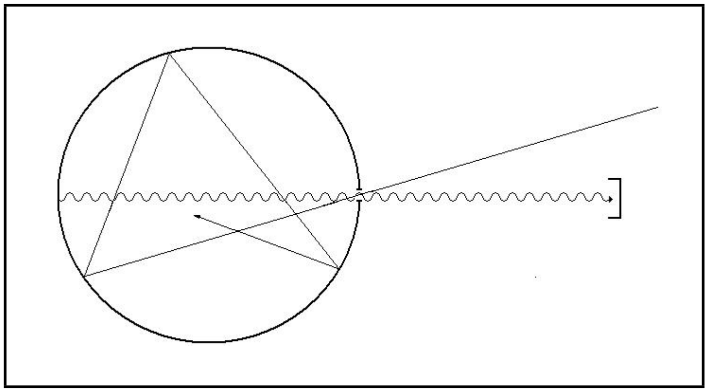

Figure 1.1: Scheme for realization of a black body as a cavity

Thus the radiation coming from the opening is just the thermal radiation, which will be measured in dependence of its frequency and the oven temperature as shown in Fig. 1.2. Such radiation sources are also referred to as (thermal) cavities.

The classical law that describes such thermal radiation is the Rayleigh-Jeans law which expresses the *energy density* $u(\omega)$ in terms of frequency ω and temperature T.

Theorem 1.1 (Rayleigh-Jeans law) $u(\omega) \; = \; \frac{kT}{\pi^2 c^3}\, \omega^2$

where k denotes *Boltzmann's constant*, $k = 1,38 \times 10^{-23}\,\mathrm{J\,K^{-1}}$.

Boltzmann's constant plays a role in classical thermo-statistics where (ideal) gases are analyzed whereas here we describe radiation. The quantity kT has the dimension of energy, e.g., in a classical system in thermal equilibrium, each degree of freedom (of motion) has a mean energy of $E = \frac{1}{2}kT$ – *Equipartition Theorem.*

From the expression of Theorem 1.1 we immediately see that the integral of the energy density over all possible frequencies is divergent,

$$\int_0^\infty d\omega \; u(\omega) \to \infty \,, \tag{1.3}$$

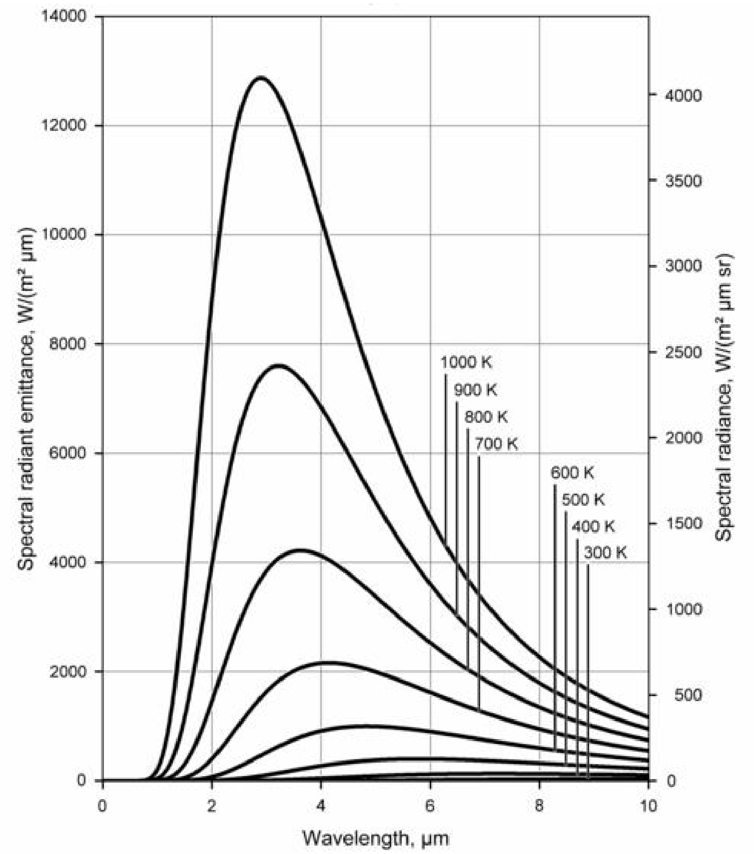

Figure 1.2: Spectrum of a black body for different temperatures, picture from: http://commons.wikimedia .org/wiki/Image:BlackbodySpectrum_lin_150dpi_en.png

which would imply an infinite amount of energy in the black body radiation. This is known as the *ultraviolet catastrophe* and the law is therefore only valid for small frequencies.

For high frequencies a law was found empirically by Wilhelm Wien in 1896.

Theorem 1.2 (Wien's law) $u(\omega) \to A\,\omega^3 e^{-B\frac{\omega}{T}}$ *for* $\omega \to \infty$

where A and B are constants which we will specify later on.

As already mentioned Max Planck derived an impressive formula which interpolates[1] between the Rayleigh-Jeans law and Wien's law. For the derivation the correctness of Proposition 1.1.1 was necessary, i.e., the energy can only occur in quanta of $\hbar\omega$. For this achievement he was awarded the 1919 nobel prize in physics.

---

[1] See Fig. 1.3

Theorem 1.3 (Planck's law) $$u(\omega) \;=\; \frac{\hbar}{\pi^2 c^3}\,\frac{\omega^3}{\exp\left(\frac{\hbar\omega}{kT}\right) - 1}$$

From Planck's law we arrive at the already well-known laws of Rayleigh-Jeans and Wien by taking the limits for $\omega \to 0$ or $\omega \to \infty$ respectively:

$$\begin{array}{lcl} \text{for } \;\omega \to 0 & & \to \quad \text{Rayleigh-Jeans} \\ & u(\omega) = \frac{\hbar}{\pi^2 c^3}\frac{\omega^3}{\exp\left(\frac{\hbar\omega}{kT}\right) - 1} & \\ \text{for } \;\omega \to \infty & & \to \quad \text{Wien} \end{array}$$

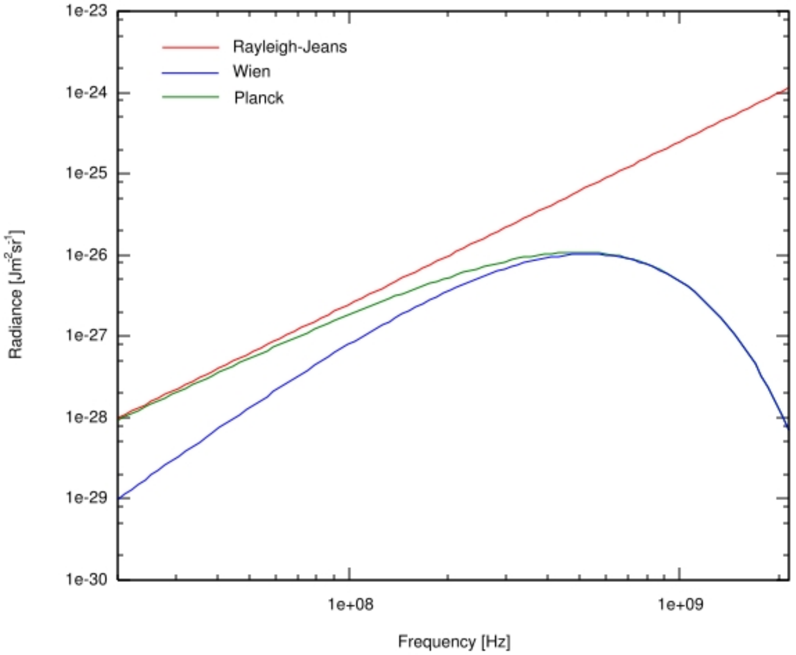

Figure 1.3: Comparison of the radiation laws of Planck, Wien and Rayleigh-Jeans, picture from: http://en.wikipedia.org/wiki/Image:RWP-comparison.svg

We also want to point out some of the consequences of Theorem 1.3.

- **Wien's displacement law**: $\boxed{\lambda_{\max} T \;=\; const. \;=\; 0,29\,\mathrm{cm\,K}}$
- **Stefan-Boltzmann law**: for the radiative power we have

$$\int_0^\infty d\omega\, u(\omega) \;\propto\; T^4 \int_0^\infty d\left(\frac{\hbar\omega}{kT}\right) \frac{\left(\frac{\hbar\omega}{kT}\right)^3}{\exp\left(\frac{\hbar\omega}{kT}\right) \,-\, 1} \tag{1.4}$$

substituting $\frac{\hbar\omega}{kT} = x$ and using formula

$$\int_0^\infty dx \frac{x^3}{e^x - 1} = \frac{\pi^4}{15} \tag{1.4}$$

we find the proportionality $\boxed{\propto T^4}$

1.1.3 Derivation of Planck's Law

Now we want to derive Planck's law by considering a black body realized by a hollow metal ball. We assume the metal to be composed of two energy-level atoms such that they can emit or absorb photons with energy $E = \hbar\omega$ as sketched in Fig. 1.4.

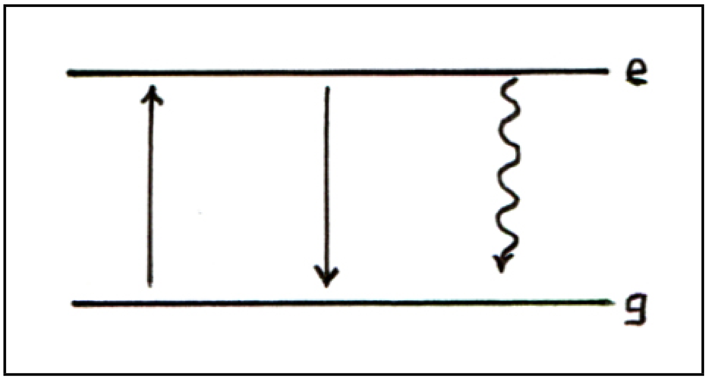

Figure 1.4: Energy states in a two level atom, e... excited state, g... ground state. There are three processes: absorption, stimulated emission and spontaneous emission of photons.

Further assuming thermal equilibrium, the number of atoms in ground and excited states is determined by the (classical) *Boltzmann distribution* of statistical mechanics

$$\frac{N_e}{N_g} = \exp\left(-\frac{E}{kT}\right), \tag{1.5}$$

where $N_{e/g}$ is the number of excited/non-excited atoms and T is the thermodynamical temperature of the system.

As Einstein first pointed out there are three processes: the absorption of photons, the stimulated emission and the spontaneous emission of photons from the excited state of the atom, see Fig. 1.4. The stimulated processes are proportional to the number of photons whereas the spontaneous process is not, it is just proportional to the transition rate. Furthermore, the coefficients of absorption and stimulated emission are assumed to equal and proportional to the probability of spontaneous emission.

Now, in the thermal equilibrium the rates for emission and absorption of a photon must be equal and thus we can conclude that:

Absorption rate:	$N_g\,\overline{n}\,P$	P ... probability for transition
Emission rate:	$N_e\,(\overline{n}+1)\,P$	$\overline{n}$ ... average number of photons

$$\Rightarrow$$

$$N_g\,\overline{n}\,P \;=\; N_e\,(\overline{n}+1)\,P \quad\Rightarrow\quad \frac{\overline{n}}{\overline{n}+1} \;=\; \frac{N_e}{N_g} \;=\; \exp\left(-\frac{E}{kT}\right) \;=\; \exp\left(-\frac{\hbar\omega}{kT}\right)$$

providing the average number of photons in the cavern[2]

$$\overline{n} \;=\; \frac{1}{e^{\frac{\hbar\omega}{kT}}-1}\,, \tag{1.6}$$

and the average photon energy

$$\overline{n}\,\hbar\omega \;=\; \frac{\hbar\omega}{e^{\frac{\hbar\omega}{kT}}-1}\,. \tag{1.7}$$

Turning next to the energy density we need to consider the photons as standing waves inside the hollow ball due to the periodic boundary conditions imposed on the system by the cavern walls. We are interested in the number of possible modes of the electromagnetic field inside an infinitesimal element dk, where $k = \frac{2\pi}{\lambda} = \frac{\omega}{c}$ is the wave number.

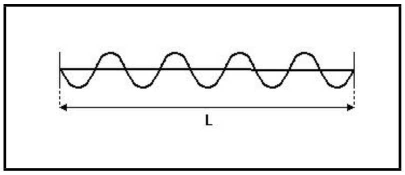

Figure 1.5: Photon described as standing wave in a cavern of length L.

The number of knots N (wavelengths) of this standing wave is then given by

$$N \;=\; \frac{L}{\lambda} \;=\; \frac{L}{2\pi}\frac{2\pi}{\lambda} \;=\; \frac{L}{2\pi}k\,, \tag{1.8}$$

which gives within an infinitesimal element

$$\text{1-dim:}\quad dN \;=\; \frac{L}{2\pi}dk \qquad \text{3-dim:}\quad dN \;=\; 2\,\frac{V}{(2\pi)^3}d^3k\,, \tag{1.9}$$

where we inserted in 3 dimensions a factor 2 for the two (polarization) degrees of freedom of the photon and wrote L^3 as V, the volume of the cell. Furthermore, inserting $d^3k = 4\pi k^2 dk$ and using the relation $k = \frac{\omega}{c}$ for the wave number we get

$$dN \;=\; 2\,\frac{V}{(2\pi)^3}\frac{4\pi\omega^2 d\omega}{c^3}\,. \tag{1.10}$$

We will now calculate the energy density of the photons.

---

[2] Note that in our derivation P corresponds to the Einstein coefficient A and $\overline{n}\,P$ to the Einstein coefficient B in Quantum Optics.

1. **Classically:** In the classical case we follow the *equipartition theorem*, telling us that in thermal equilibrium each degree of freedom contributes $E = \frac{1}{2}kT$ to the total energy of the system, though, as we shall encounter, it does not hold true for quantum mechanical systems, especially for low temperatures.

   Considering the standing wave as harmonic oscillator has the following mean energy:

   $$E \;=\; \langle E_{\text{kin}} \rangle + \langle V \rangle \;=\; \left\langle \frac{m}{2} v^2 \right\rangle + \left\langle \frac{m\omega}{2} x^2 \right\rangle \;=\; \frac{1}{2}kT + \frac{1}{2}kT \;=\; kT\,. \tag{1.11}$$

   For the oscillator the mean kinetic and potential energy are equal[3], $\langle E_{\text{kin}} \rangle = \langle V \rangle$, and both are proportional to quadratic variables which is the equipartition theorem's criterion for a degree of freedom to contribute $\frac{1}{2}kT$. We thus can write $dE = kT\,dN$ and by taking equation (1.10) we calculate the (spectral) energy density

   $$u(\omega) \;=\; \frac{1}{V}\frac{dE}{d\omega} \;=\; \frac{kT}{\pi^2 c^3}\,\omega^2\,, \tag{1.12}$$

   which we recognize as Theorem 1.1, the Rayleigh-Jeans Law.

2. **Quantum-mechanically:** In the quantum case we use the average photon energy, equation (1.7), we calculated above by using the quantization hypothesis, Proposition 1.1.1, to write $dE = \overline{n}\,\hbar\omega\,dN$ and again inserting equation (1.10) we calculate

   $$u(\omega) \;=\; \frac{1}{V}\frac{dE}{d\omega} \;=\; \frac{\hbar}{\pi^2 c^3}\,\frac{\omega^3}{e^{\frac{\hbar\omega}{kT}} \,-\, 1}\,, \tag{1.13}$$

   and we recover Theorem 1.3, Planck's Law for black body radiation.

1.2 The Photoelectric Effect

1.2.1 Facts about the Photoelectric Effect

In 1887 Heinrich Hertz discovered a phenomenon, the photoelectric effect, that touches the foundation of quantum mechanics. A metal surface emits electrons when illuminated by ultraviolet light. The importance of this discovery lies within the inability of classical physics to describe the effect in its full extent based on three observations.

1.) The kinetic energy of the emitted electrons is independent of the intensity of the illuminating source of light and

2.) increases with increasing frequency of the light.

3.) There exists a threshold frequency below which no electrons are emitted.

---

[3] This statement results from the so-called Virial theorem, stating that $\langle 2E_{\text{kin}} \rangle = \left\langle x \frac{dV}{dx} \right\rangle$, which we will encounter later on in Section 2.7.

More accurate measurements were made by Philipp Lenard between 1900 and 1902 for which he received the nobel prize in 1905.

In terms of classical physics this effect was not understood as from classical electrodynamics was known that the

$$\begin{aligned} &\text{energy density:} \quad & u &= \frac{1}{8\pi}\left(\vec{E}^2 + \vec{B}^2\right) & (1.14) \\ \text{and the} \quad &\text{energy flux:} \quad & S &= \frac{c}{8\pi}\,\vec{E} \times \vec{B} & (1.15) \end{aligned}$$

are both proportional to the intensity. Thus “knocking” electrons out of the metal is possible, but there is no threshold that limits this process to certain frequencies. As long as the surface is exposed to the radiation the electrons absorb energy until they get detached from the metal.

1.2.2 Einstein’s Explanation for the Photoelectric Effect

The phenomenon of the photoelectric effect could then be explained by Albert Einstein in 1905 with his photon hypothesis which he was also awarded the nobel prize for in 1921.

Proposition 1.2.1 (Photon hypothesis)
Light consists of quanta (photons) with respective energy of $E = \hbar\omega$ *each.*

Einstein explained the effect in the following way. The incident photons transfer their energy to the electrons of the metal. Since the electrons are bound to the metal the energy need to be sufficient to overcome the electrostatic barrier of the metal. The respective energy, which is dependent on the material used is termed the work function of the material.

Proposition 1.2.2 (Photoelectric formula) $E_{\text{kin}} = \frac{mv^2}{2} = \hbar\omega - W$

where W is the *work function* of the metal. From Proposition 1.2.2 we understand that the kinetic energy of the emitted electrons is bounded by the frequency of the incident light such that $E_{\text{kin,e}^-} < \hbar\omega_{\text{photon}}$ and we conclude that a threshold frequency ω_0 must exist where the emitted electrons are at rest $E_{\text{kin,e}^-} = 0$ after escaping the potential of the metal.

$$\text{Threshold frequency} \quad \boxed{\omega_0 = \frac{W}{\hbar}}$$

For frequencies $\omega < \omega_0$ no electrons are emitted and we have reflection or scattering of the incident light together with electronic and/or thermal excitation instead. Einstein's explanation caused serious debates amongst the community and was not generally accepted due to conceptual difficulties to unify the quantized (particle) nature of light with Maxwell's equations, which stood on solid ground experimentally, suggesting a wave-like behaviour of light. The experiment to change that point of view was performed by Robert Andrews Millikan.

1.2.3 The Millikan Experiment

In 1916 R.A.Millikan developed an experimental setup (sketched in Fig. 1.6) which allowed to check the accuracy of Einstein's formula (Proposition 1.2.2). He was awarded the Nobel prize for his work on the electron charge and the photoelectric effect in 1923.

A photocathode is irradiated with ultraviolet light which causes the emission of electrons thus generating an electric current, between the cathode and the electron collector, that is measured by a galvanometer. Additionally a voltage is installed such that the electrons are hindered in advancing the collector. The electrons now need to raise energy, respective to the voltage U, to get to the collector and generate the current. So if $eU > E_{\text{kin}}$ no electrons arrive at the collector. If however the voltage is regulated to a value U_0 such that the current (as measured by the galvanometer) tends to zero, then U_0 is a measure for the kinetic energy. This allows us to measure the work function of the metal at hand[4].

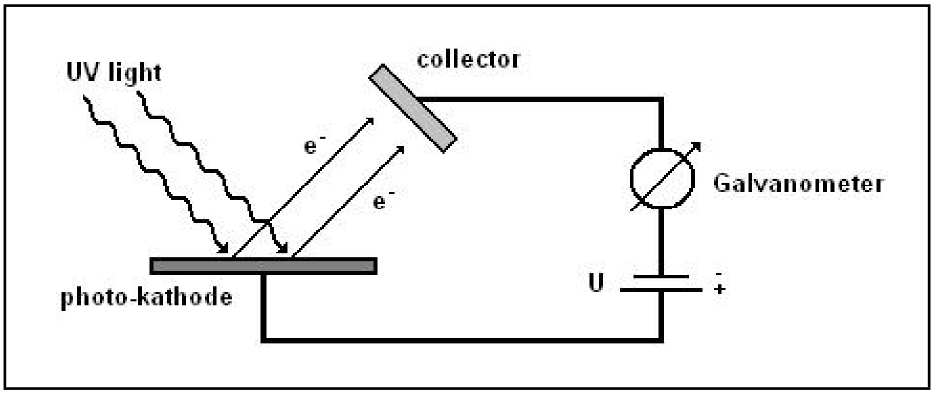

Figure 1.6: Schematic overview of the Millikan experiment

$$\boxed{eU_0 = E_{\text{kin}} = \hbar\omega - W} \quad \rightarrow \text{ work function } W \text{ } (\approx 2\text{–}5 \text{ eV for typical metals})$$

Furthermore, the set up allows for a precise measurement of Planck's constant as can be seen in Fig. 1.7.

---

[4] assuming the collector and emitter are of the same material, otherwise an additional contact voltage needs to be considered

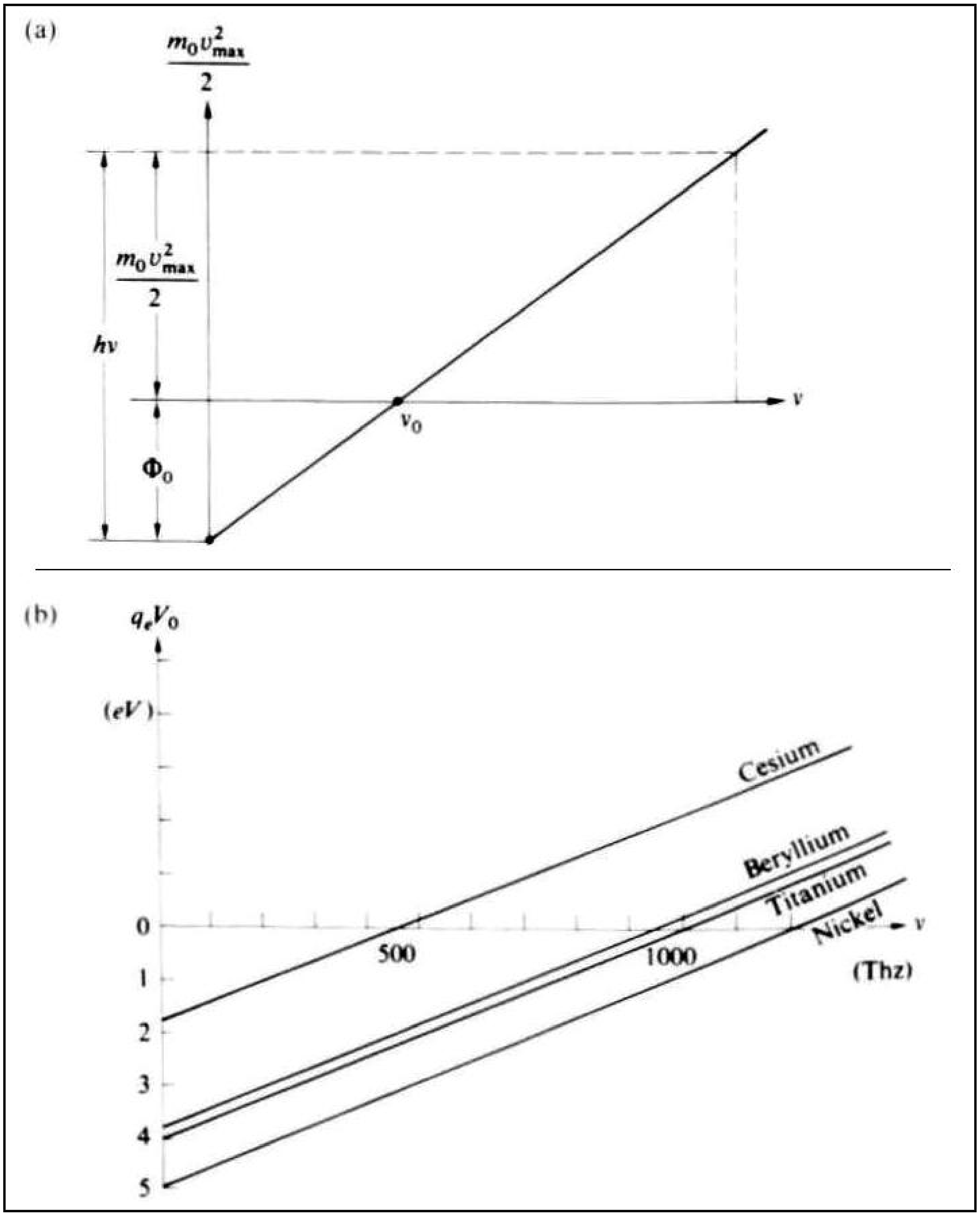

Figure 1.7: Determination of Planck's constant: a) Measuring the kinetic energy of the electrons as a function of the incident radiation frequency, gives Planck's constant as the slope of the resulting linear function. b) Results are shown for different materials displaying the workfunctions as the respective vertical intercepts; pictures from [2]

With the confirmation of Einstein's formula we now want to take a look at the properties of the thus introduced photons. From special relativity we know that energy and velocity are related by

$$1)\quad E=\sqrt{p^2c^2+m^2c^4}\quad \text{and}\quad 2)\quad \vec{v}=\frac{\partial E}{\partial \vec{p}}=\frac{\vec{p}c^2}{\sqrt{p^2c^2+m^2c^4}}\,, \tag{1.16}$$

since for the photon the velocity is $|\vec{v}| = c \cong 2,99.10^{10}\mathrm{cms}^{-1}$ we conclude $\Rightarrow$ $\boxed{m_{\text{photon}} = 0}$
The photon is massless.

If we then compare the energy of the photon as given by special relativity $E = pc$ and quantum mechanics $E = \hbar\omega$ we get the momentum of the photon:

$$k=\frac{\omega}{c}\quad\Rightarrow\quad E=\hbar kc\quad\Rightarrow\quad p=\hbar k\,.$$

$$\text{Momentum of the photon}\qquad \boxed{\vec{p}_{\text{photon}}=\hbar\vec{k}}$$

1.3 The Compton Effect

Another effect that revealed the quantized nature of radiation was the (elastic) scattering of light on particles, called Compton effect or Compton scattering, see Fig. 1.8. For a description we study energy and momentum of the photon, both particle properties. The effect works in total analogy to a scattering process with particles, i.e. energy and momentum are conserved.

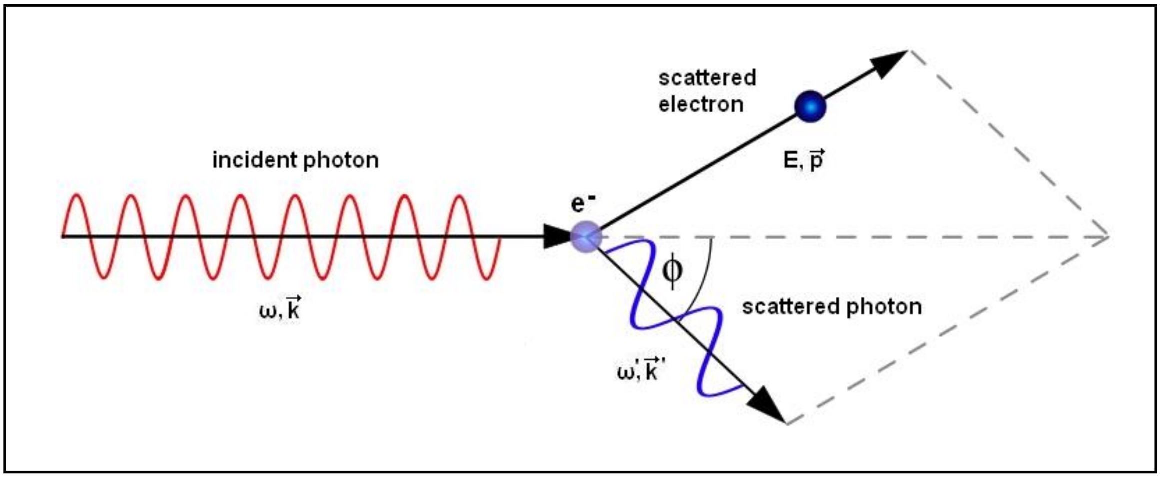

Figure 1.8: Compton effect: scattering of a photon on a resting electron, picture from: http://de.wikipedia.org/wiki/Bild:Compton_scattering-de.svg

1.3.1 The Compton Shift Formula

We assume:

conservation of energy:

$$\hbar\omega + mc^2 = \hbar\omega' + E\,, \tag{1.17}$$

conservation of momentum:

$$\hbar\vec{k} = \hbar\vec{k}\,' + \vec{p}\,. \tag{1.18}$$

From Eqs. (1.17) and (1.18) we find the following relation when using the relativistic energy $E^2 = p^2c^2 + m^2c^4$, Eq. (1.16), for the electron

$$\frac{1}{c^2}\left(\hbar\omega - \hbar\omega' + mc^2\right)^2 \;-\; \hbar^2\left(\vec{k} - \vec{k}\,'\right)^2 \;=\; \frac{1}{c^2}E^2 - \vec{p}^{\,2} \;=\; m^2c^2\,. \tag{1.19}$$

Recalling that $\omega = kc$, $\omega' = k'c$, $k = |\vec{k}|$ and $\vec{k}\vec{k}\,' = kk'\cos\phi$ we calculate

$$\begin{aligned} \frac{1}{c^2}\hbar^2\left(\omega - \omega'\right)^2 \quad &= \hbar^2\left(k - k'\right)^2 \quad = \quad \hbar^2\left(k^2 + k'^{\,2} - 2kk'\right) \\ \text{and} \qquad -\,\hbar^2\left(\vec{k} - \vec{k}\,'\right)^2 \quad &= -\,\hbar^2\left(k^2 + k'^{\,2} - 2\vec{k}\vec{k}\,'\right) = -\,\hbar^2\left(k^2 + k'^{\,2} - 2kk'\cos\phi\right) \end{aligned}$$

which we insert into Eq. (1.19) and obtain

$$\begin{aligned} m^2c^2 + 2mc\hbar\left(k - k'\right) \;&-\; 2\hbar^2kk'\left(1 - \cos\phi\right) \;=\; m^2c^2 \\ k - k' \;&=\; \frac{\hbar}{mc}kk'\left(1 - \cos\phi\right) \\ 2\pi\left(\frac{1}{\lambda} - \frac{1}{\lambda'}\right) \;&=\; \frac{\hbar}{mc}\frac{(2\pi)^2}{\lambda\lambda'}2\sin^2\frac{\phi}{2}\;. \end{aligned}$$

Multiplying finally both sides by $\frac{\lambda\lambda'}{2\pi}$ and denoting $\lambda' - \lambda = \Delta\lambda$, we arrive at the formula:

Lemma 1.1 (Compton shift formula) $\Delta\lambda \;=\; \lambda' - \lambda \;=\; 4\pi\frac{\hbar}{mc}\,\sin^2\frac{\phi}{2}$

Result: We discover that the wavelength of the scattered photon has increased (or the frequency has decreased) relatively to the incoming photon due to the energy transfer to the electron. The difference $\Delta\lambda$ is directly related to the *scattering angle* ϕ.

It is customary to define the *Compton wavelength* $\lambda_{\mathrm{c}} = \frac{h}{mc}$ ($= 2,43 \times 10^{-10}$ cm for electrons).

Since λ_c is very small, high energy radiation (X-rays) is needed to observe the effect. If we choose a wavelength of 7×10^{-9} cm for the X-rays we estimate for a maximal scattering angle an effect of

$$\frac{\Delta\lambda}{\lambda} \;=\; \frac{2\lambda_c}{\lambda_{\text{Xray}}} \;=\; \frac{2 \times 2,43 \times 10^{-10}\text{cm}}{7 \times 10^{-9}\text{cm}} \;\approx\; 0,07 \;=\; 7\,\%\,. \tag{1.20}$$

1.3.2 The Experiment of Compton

In the experiment by A.H.Compton, which he received the Nobel prize for in 1927, X-rays are scattered by nearly free electrons in carbon (graphite) as seen in Fig. 1.9. The intensity of the outgoing radiation is then measured as a function of its wavelength as can be seen in Fig. 1.10.

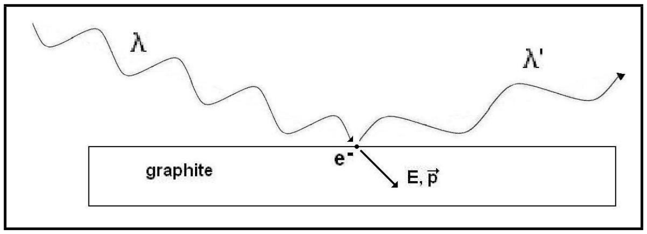

Figure 1.9: Experiment of Compton: scattering of X-rays by electrons in the surface of a carbon block

The intensity of the scattered peak (due to the quasi-free electrons in the material) increases at the expense of the unscattered peak which remains (due to electrons bound to the atom that do not allow a shift in the energy or wavelength). In fact, we have $I^{\max}_{scattered} > I^{\max}_{unscattered}$.

Résumé: The particle character of light is confirmed in Compton's experiment and we assign the energy of $E = \hbar\omega$ and the momentum $\vec{p} = \hbar\vec{k}$ to the (undivisible) photon. The Compton formula (Lemma 1.1) reveals a proportionality to $\hbar$, a quantum mechanical property, which is confirmed in the experiment. Classically no change of the wavelength is to be expected.

1.4 Bohr's Theses

In the early nineteenhundreds many elementary questions about our understanding of matter, especially atoms, and radiation were unanswered and created serious problems. E.g.: Why don't electrons tumble into the atomic nucleus? If the electrons rotate around the nucleus, they would be in accelerated motion thus radiating, i.e., losing their energy which keeps them on their paths. Another important question was why are sharp spectral lines observed. In an attempt to overcome those issues Niels Bohr postulated new rules incompatible with classical notions, which were quite revolutionary at that time.

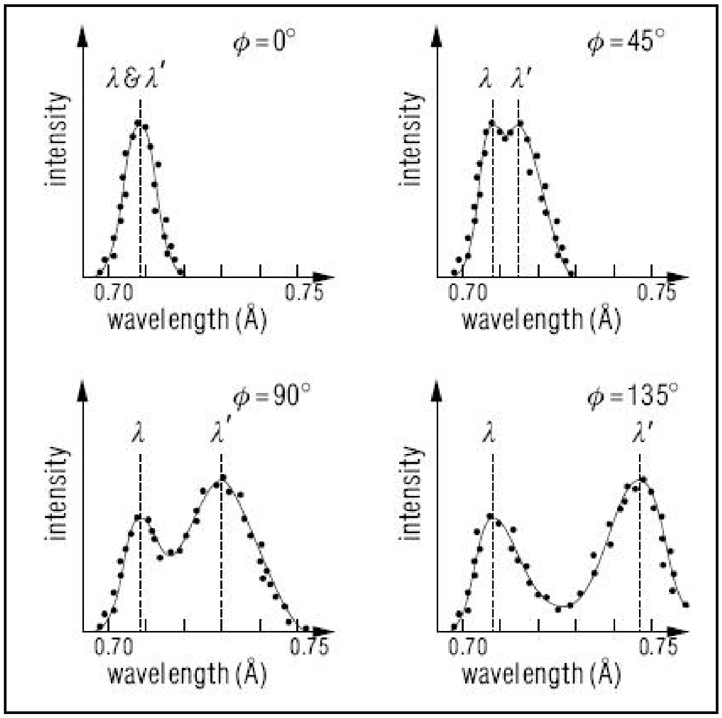

Figure 1.10: Compton shift: intensity of radiation as detected after the scattering process, the scattered line is shifted against the remaining unscatterd line by $\Delta\lambda$ depending on the scattering angle ϕ, picture from Ref. [1]

Proposition 1.4.1 (Bohr's quantum postulate)
Electrons can only take certain discrete energy values in the atom. During the transition from a state with energy E_n *to another state with energy* E_m *a photon with frequency* $\omega = \frac{1}{\hbar}|\,(E_\mathrm{n} - E_\mathrm{m})\,|$ *is emitted* $(E_\mathrm{n} > E_\mathrm{m})$ *or absorbed* $(E_\mathrm{n} < E_\mathrm{m})$.

Bohr's atom model: Having postulated the correspondence between energy levels and the emitted/absorbed photons the next step was to postulate how these energy levels came to be.

Proposition 1.4.2 (Bohr's quantization rule)
States are stationary if their respective classical paths are quantized such that their action functional obeys

$$\oint p\,dq = nh$$

where n *is an integer number* $(n = 1, 2, ...)$.

The energy spectrum for the hydrogen atom, given by the *Rydberg-formula* and empirically already known before Bohr's proposition, was then explained in Bohr's theory as the energy of an electron transition between two orbitals n and m. Furthermore Bohr's model could theoretically predict the value of the Rydberg constant R

$$\hbar\omega \;=\; E_{\mathrm{n}} - E_{\mathrm{m}} \;=\; R\left(\frac{1}{m^2} - \frac{1}{n^2}\right) \quad \text{or} \quad \frac{1}{\lambda} \;=\; \frac{R}{hc}\left(\frac{1}{m^2} - \frac{1}{n^2}\right), \tag{1.21}$$

where $\frac{R}{hc} \;=\; \frac{2\pi^2 m\, e^4}{h^3 c} \;=\; 1,097 \times 10^5\,\mathrm{cm}^{-1}$ is the *Rydberg constant.*

Although the model of Bohr allows a very simple and intuitive visualization, electrons orbiting on fixed paths around the nucleus, creating an analogue to planetary movement[5], it turns out the model is not only too simple but just wrong. Nevertheless Bohr's model represented a great step forward but it was clear that this *ad hoc* prescriptions couldn't be considered as a definitive theory. In fact, as it turned out, the perception of particles following distinct paths altogether has no meaning for quantum mechanical calculations and only certain interpretations of quantum mechanics allow for their existence at all. But generally, as we shall see in Sect. 1.7, we need to cast aside the concept of a trajectory in quantum mechanics.

Another issue that arose was the debate whether light was a wave or composed of particles, concepts which seemed to contradict each other but were both observed under certain conditions. Bohr tried to formulate this problem in the following way:

Proposition 1.4.3 (Bohr's complementarity principle)
Wave and particle are two aspects of describing physical phenomena, which are complementary to each other.

Depending on the measuring instrument used, either waves or particles are observed, but never both at the same time, i.e. wave- and particle-nature are not simultaneously observable.

Similar statements can be made about other *complementary quantities* like position and momentum, energy and time, spin in orthogonal directions (e.g.: σ_{x} and σ_{y}).

Some of these questions are still subject of interest and play important roles in experiments regarding "which way detectors", "quantum erasers" and "continuous complementarity" to name but a few.

---

[5] Arnold Sommerfeld, 1868 – 1951, a prominent physicists at that time and teacher of many Nobel prize winners like Heisenberg, Pauli and Bethe, has extended Bohr's atom model to explain the fine structure of hydrogen.

A question remaining, which still causes discussion, is the connection of quantum mechanics to classical physics which Bohr phrased as such:

Proposition 1.4.4 (Bohr's correspondence principle)
In the limit of large quantum numbers classical physics is reproduced.

The propositions of this section form the basis for what is usually referred to as the *Copenhagen interpretation of quantum mechanics*, which was developed mainly by Niels Bohr and Werner Heisenberg.

Bohr was *the* leading figure of quantum physics at that time and his institute in Copenhagen became the center of the avantgarde of physics where physicists all over the world met, just to mention some names: Heisenberg, Pauli, Gamow, Landau, Schrödinger, Kramers, ... Bohr's dispute with Einstein about the foundations of quantum mechanics became legendary. Bohr was awarded the Nobel prize in 1922.

1.5 Wave Properties of Matter

As we will see in this section, not only radiation, but also massive particles are in need of a more sophisticated description than given by classical mechanics. To associate microscopical (quantum) objects, as for example electrons, with idealized (especially localized) point-particles, carrying sharp momenta, is not only misleading, but simply wrong (see Sect. 1.6 and Sect. 2.6) and can not account for all observed phenomena. A very important step towards a more complete description was Louis de Broglie's proposal of wavelike behaviour of matter in 1923, which he received the Nobel prize for in 1929.

1.5.1 Louis de Broglie's Hypothesis

In view of particle properties for light waves – photons – Louis de Broglie ventured to consider the reverse phenomenon, he proposed to assign wave properties to matter, which we will formulate here in the following way:

Proposition 1.5.1 (Louis de Broglie's hypothesis)
To every particle with mass m, momentum $\vec{p}$ and energy E a wavelength of $\lambda_{\text{deBroglie}} = \frac{h}{|\vec{p}|} = \frac{h}{\sqrt{2mE}}$ is associated, where $E = E_{\text{kin}} = \frac{\vec{p}^{\,2}}{2m}$.

The above statement can be easily understood when assigning energy and momentum

$$E = \hbar\omega \quad \text{and} \quad p = \hbar k = \frac{h}{\lambda} \tag{1.22}$$

to matter in (reversed) analogy to photons. If we then express the wavelength λ through the momentum p and use the form of the kinetic energy $E = p^2/2m$ to write $p = \sqrt{2mE}$ we directly get the *de Broglie wavelength* $\lambda_{\mathrm{deBroglie}}$ of massive particles.

In this connection the notion of *matter waves* was introduced. De Broglie's view was that there exists a *pilot wave* which leads the particle on definite trajectories. This point of view – wave *and* particle – being in contrast to Bohr's view leads, however, into serious difficulties as we shall see.

Note that above wave assignment was made for free particles, i.e. particles that are not subjected to any outer potential. The question whether the potential energy would influence his hypothesis was also raised by de Broglie and will be tangible when we consider Schrödinger's theory (see Chapt. 2) where also the nature of the waves becomes more evident in terms of Max Born's probability interpretation (see Sect. 1.7).

1.5.2 Electron Diffraction from a Crystal

To test his hypothesis de Broglie proposed an experiment with electrons. He observed that, following Proposition 1.5.1, electrons with a kinetic energy of several eV and mass $m_{\mathrm{e}} = 0,5\,\mathrm{MeV}$ would have a de Broglie wavelength of a few $\mathring{A}$. For example, for an energy of $10\,\mathrm{eV}$ we obtain $\lambda_{\mathrm{deBroglie}} = 3,9\ \mathring{A}$, which is the same order of magnitude as the lattice spacing of atoms in crystals, thus making it possible to diffract electrons from the lattice analogously to the diffraction of light from a grating.

The corresponding experiment has been performed by C.Davisson and L.Germer in 1927 and independently by G.P. Thomson. It involved electrons which were sent with appropriate velocity onto a nickel crystal with lattice spacing $d \cong 0,92\ \mathring{A}$, see Fig. 1.11.

The intensity of the outgoing electron beam was then measured for different angles, reproducing the diffraction pattern postulated by W.H.Bragg and (his son) W.L.Bragg for X-rays. The similarity of X-ray- and electron-diffraction can be seen in Fig. 1.12.

The Bragg condition for constructive interference is $\boxed{\sin\theta = \frac{n\lambda}{2d}}\,,\quad n \in N\,.$

The observation of an intensity maximum of the diffraction (Bragg peak) for a scattering angle $\varphi = 50°$, which translates to the angle in the Bragg condition of $\Theta = 65°$, gives us

$$\Rightarrow \quad \lambda \;=\; 2 \times 0,92\ \mathring{A} \times \sin 65° \;=\; 1,67\ \mathring{A}\,,$$

which is in perfect accordance with the de Broglie wavelength for an acceleration voltage of $U = 54\,\mathrm{V}$ used in this experiment.

The Davisson-Germer experiment thus helped to confirm the wavelike nature of matter

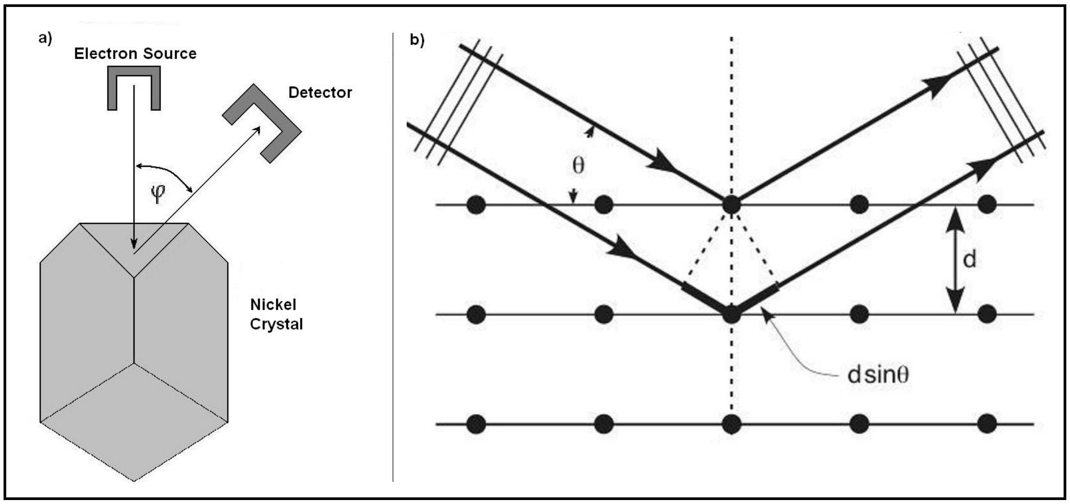

Figure 1.11: Davisson-Germer Experiment: a) An electron beam is diffracted from a nickel crystal and the intensity of the outgoing beam is measured. b) Scheme of the Bragg diffraction from a crystal with lattice spacing d, picture (b) from: http://en.wikipedia.org/wiki/Image:Bragg_diffraction.png

and to validate the claims of early quantum mechanics. Davisson and Thomson[6] were awarded the Nobel prize in 1937.

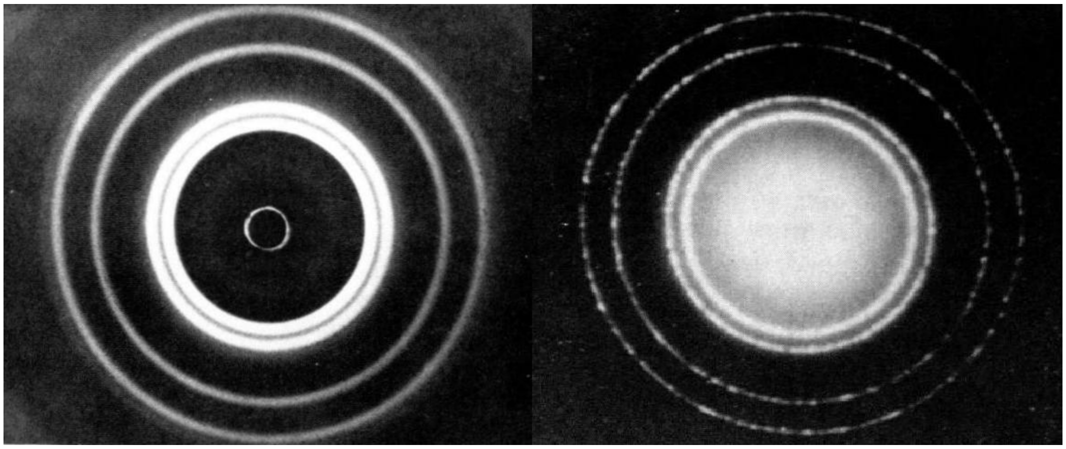

Figure 1.12: Comparison of X-ray- (left) and electron- (right) diffraction patterns caused by the same aperture, namely a small hole in an aluminium foil; pictures from Ref. [2]

---

[6] It's quite amusing that his father J.J. Thomson received the Nobel prize in 1906 for showing the "opposite", namely that the electron is a particle.

1.6 Heisenberg's Uncertainty Principle

We now want to introduce a quantum mechanical principle, *Heisenberg's uncertainty principle* that is somehow difficult to grasp conceptually even though the mathematics behind is straightforward. Before we will derive it formally, which we will do in Sect. 2.6, we try to make it plausible by more heuristic means, namely by the *Heisenberg microscope.* The Gedankenexperiment of the Heisenberg microscope needs to be seen more as an example for the application of the uncertainty principle, than a justification of the principle itself.

1.6.1 Heisenberg's Microscope

Let's start by detecting the position of an electron by scattering of light and picturing it on a screen. The electron will then appear as the central dot (intensity maximum) on the screen, surrounded by bright and dark concentric rings (higher order intensity maxima/minima). Since the electron acts as a light source we have to consider it as an aperture with width d where we know that the condition for destructive interference is

$$\sin\phi \,=\, \frac{n\lambda}{d} \;\;,\; n \in \mathbb{N}\,. \tag{1.23}$$

So following Eq. (1.23) the smallest length resolvable by a microscope is given by $d = \lambda / \sin\phi$ and thus the *uncertainty of localization* of an electron can be written as

$$\Delta x \,=\, d \,=\, \frac{\lambda}{\sin\phi}\,. \tag{1.24}$$

It seems as if we chose the wavelength λ to be small enough and $\sin\phi$ to be big, then Δx could become arbitrarily small. But, as we shall see, the accuracy increases at the expense of the momentum accuracy. Why is that? The photons are detected on the screen but their direction is unknown within an angle ϕ resulting in an uncertainty of the electron's recoil within an interval Δ. So we can identify the momentum uncertainty (in the direction of the screen) of the photon with that of the electron

$$\Delta p_{\mathrm{x}}^{\mathrm{e}} \,=\, \Delta p_{\mathrm{x}}^{\mathrm{Photon}} \,=\, p^{\mathrm{Photon}} \sin\phi \,=\, \frac{h}{\lambda}\sin\phi\,, \tag{1.25}$$

where we inserted the momentum of the photon $p^{\mathrm{Photon}} \,=\, \hbar k \,=\, h/\lambda$ in the last step. Inserting the position uncertainty of the electron from Eq. (1.24) into Eq. (1.25) we get

$$\boxed{\text{Heisenberg's Uncertainty relation:} \qquad \Delta\, x\, \Delta\, p_{\mathrm{x}} \,=\, h}$$

which he received the Nobel prize for in 1932. We will further see in Sect. 2.6 that the accuracy can be increased by a factor 4π and that the above relation can be generalized to the statement

$$\boxed{\Delta x \, \Delta p_{\mathrm{x}} \;\geq\; \tfrac{\hbar}{2}}$$

This is a fundamental principle that has nothing to do with technical imperfections of the measurement apparatus. We can phrase the uncertainty principle in the following way:

Proposition 1.6.1
Whenever a position measurement is accurate (i.e. precise information about the current position of a particle), the information about the momentum is inaccurate – uncertain – and vice versa.

1.6.2 Energy–Time Uncertainty Principle

We now want to construct another uncertainty relation, the energy–time uncertainty, which describes the relation between the uncertainties Δt for the time a physical process lasts and ΔE for the respective energy. Consider for example a wave packet traveling along the x-axis with velocity v. It takes the time Δt to cover the distance Δx. We can thus write

$$\text{energy:} \quad E \;=\; \frac{p^2}{2m}\,, \qquad \text{velocity:} \quad v \;=\; \frac{\Delta x}{\Delta t}\,. \tag{1.26}$$

Calculating the variation ΔE of the energy E and expressing Δt from the right hand side of Eq. (1.26) by the velocity v and substituting $v \;=\; p/m$ we get

$$\text{variation:} \quad \Delta E \;=\; \frac{p}{m}\Delta p, \qquad \text{time period:} \quad \Delta t \;=\; \frac{\Delta x}{v} \;=\; \frac{m}{p}\,\Delta x\,. \tag{1.27}$$

The right hand side represents the period of time where the wave is localizable. When we now multiply Δt with ΔE we arrive at:

$$\Delta t\,\Delta E \;=\; \frac{m}{p}\,\Delta x\,\frac{p}{m}\Delta p \;=\; \Delta x\,\Delta p \;\geq\; \frac{\hbar}{2}\,, \tag{1.28}$$

$$\boxed{\Delta E \, \Delta t \;\geq\; \tfrac{\hbar}{2}}$$

We can conclude that there is a fundamental complementarity between energy and time. An important consequence of the energy–time uncertainty is the finite “natural” width of the spectral lines.

1.7 Double–Slit Experiment

1.7.1 Comparison of Classical and Quantum Mechanical Results

We will now take a look at the double–slit experiment, which is well-known from classical optics and whose interference pattern is completely understood when considering light as electromagnetic waves. The experiment can be performed with massive particles as well but then a rather strange phenomenon occurs. It turns out that it is impossible, *absolutely* impossible, to explain it in a classical way or as Richard Feynman [3] phrased it emphasizing the great significance of the double–slit as *the* fundamental phenomenon of quantum mechanics: " ... the phenomenon [of the double–slit experiment] has in it the heart of quantum mechanics. In reality, it contains the *only* mystery."

The associated experiments have been performed with electrons by Möllenstedt and Sönsson in 1959 [4], with neutrons by Rauch [5] and Shull [6] in 1969 and by Zeilinger et al. in the eighties [7]. More recently, in 1999 and subsequent years, Arndt et al. [8] performed a series of experiments using large molecules, so-called fullerenes.

Classical considerations: Let's consider classical particles that pass through an array of two identical slits and are detected on a screen behind the slits. When either the first slit, the second or both slits are open then we expect a probability distribution to find the particle on the screen for the third case (both open) to be just the sum of the distributions for the first and second case (see left side of Fig. 1.13). We can formally write

$$P_{1,2}^{\text{class}} = P_1 + P_2 \,, \tag{1.29}$$

where P_{k} $(k = 1, 2)$ is the distribution when the k-th slit is open. For classical objects, like e.g. marbles, Eq. (1.29) is valid (for accordingly high numbers of objects).

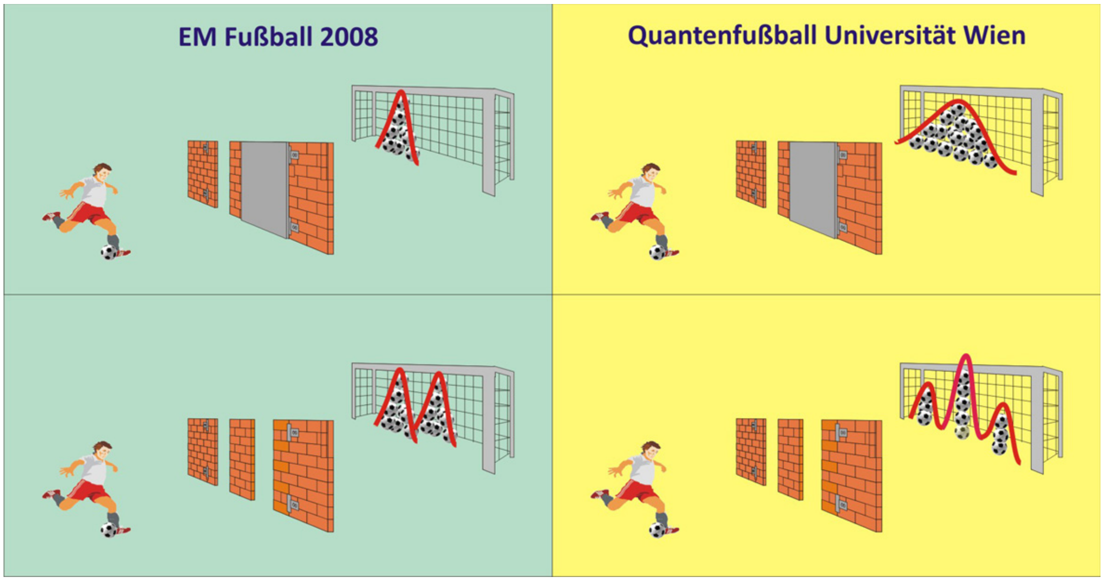

Figure 1.13: Illustration of the double–slit experiment with classical particles (left) and quantum mechanical particles (right); picture used with courtesy of M. Arndt

Quantum version: We now consider the same setup as before, but instead of classical particles we use small objects like electrons or neutrons together with the corresponding detection device and we count the number of clicks from the detector. When either one or the other slit is open, we obtain from the clicking rate some kind of probability distribution similar to the classical case (of course, the intensity need not be a single peak, see for example Fig. 1.12, but the explicit form does not matter here). More important, however, is the case when both slits are open. Then the probability distribution produced is not equal to the sum of the single distributions of either slit as in the classical case, which is illustrated on the right side of Fig. 1.13

$$P^{\mathrm{QM}}_{1,2} \neq P_1 + P_2 \,. \tag{1.30}$$

It turns out that the intensity, the probability distribution, resembles an interference pattern of waves. The intensity, the clicking rate, is high where constructive interference of the waves is observed and low where destructive interference occurs.

1.7.2 Interpretation of Quantum Mechanical Results

We now want to formulate the ideas of the previous section in a more rigorous way, concentrating on electrons as representatives of quantum mechanical objects.

We describe the behaviour of the electron by a wave $\psi(x,t)$ which is a complex-valued function.

Proposition 1.7.1 (Born's probability interpretation)
The probability P of finding the electron at a certain position x at the time t, is given by the squared amplitude $|\psi(x,t)|^2 = \psi^*(x,t)\,\psi(x,t)$ – the intensity.

For this interpretation of the wave function Max Born was awarded the 1954 Nobel prize.

From Prop. 1.7.1 we then have the following probability distributions:

$$\begin{aligned}
&\text{slit 1 open} & P_1 &= |\psi_1|^2 \\
&\text{slit 2 open} & P_2 &= |\psi_2|^2 \\
&\text{slit 1 and 2 open} & P &= |\psi|^2 = |\psi_1 + \psi_2|^2 = |\psi_1|^2 + |\psi_2|^2 + \underbrace{2\,\mathrm{Re}\,(\psi_1^*\psi_2)}_{\text{interference term}}
\end{aligned} \tag{1.31}$$

We see that for both slits open the resulting total wave function is a superposition of the individual wave functions $\psi = \psi_1 + \psi_2$. Thus the total probability corresponding to the intensity is not just the sum of the individual probabilities. From the interference term in Eq. (1.31) – which may have any sign – we conclude that if the term vanishes we won't observe any interference and if it is positive or negative we can expect constructive or destructive interference.

Résumé: An electron as a quantum system passes through the double–slit, then hits a screen or a detector. It produces a very localized lump or causes a click in the detector thus occurring as a definite particle. But the probability for the detection is given by the intensity $|\psi|^2$ of the wave ψ. It is in this sense that an electron behaves like a particle or like a wave.

Remark I: The probability distribution – the interference pattern – is not created, as one could deduce, by the simultaneously incoming electrons but does arise through the interference of the single electron wave function and thus does not change when the electrons pass through the double–slit one by one. The spot of a single electron on the screen occurs totally at random and for few electrons we cannot observe any structure, only when we have gathered plenty of them, say thousands, we can view an interference pattern, which can be nicely seen in Fig. 1.14.

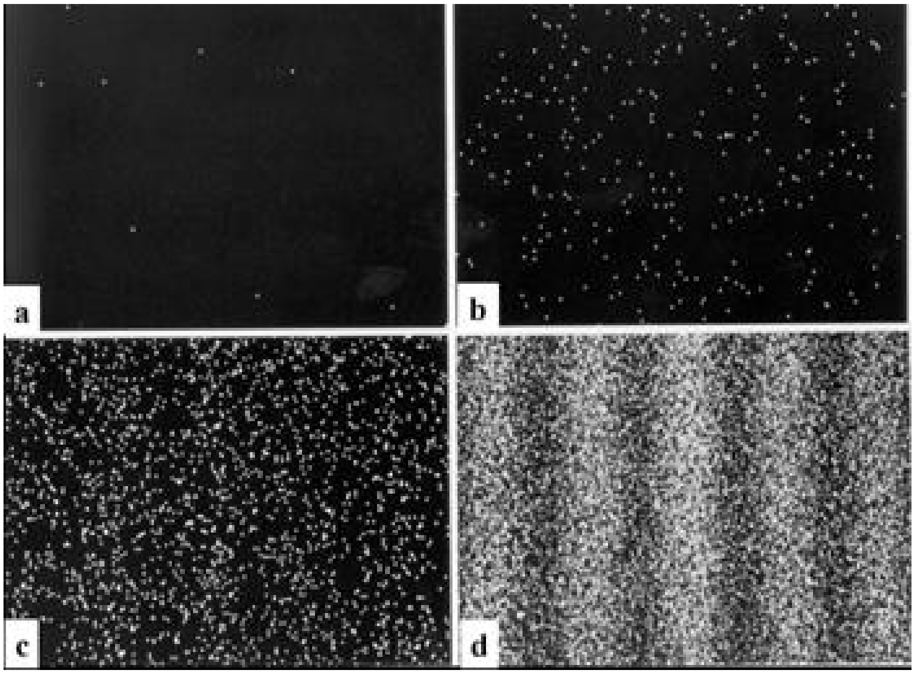

Figure 1.14: Double–slit experiment with single electrons by Tonomura: The interference pattern slowly builds up as more and more electrons are sent through the double slit one by one; electron numbers are a) 8, b) 270, c) 2000, d) 60000; picture from http://de.wikipedia.org/wiki/Bild:Doubleslitexperiment_results_Tanamura_1.gif

Remark II: Path measurement If we wish to gain which-way information, i.e. determine whether the electron passes slit 1 or 2, e.g. by placing a light–source behind the double–slit to illuminate the passing electrons, the interference pattern disappears and we end up with a classical distribution.

Proposition 1.7.2 (Duality)
Gaining path information destroys the wave like behaviour.

It is of crucial importance to recognize that the electron does not split. Whenever an electron is detected at any position behind the double–slit it is always the whole electron and not part of it. When both slits are open, we do not speak about the electron as following a distinct path since we have no such information[7].

1.7.3 Interferometry with C_{60}−Molecules

Small particles like electrons and neutrons are definitely quantum systems and produce interference patterns, i.e. show a wave-like behaviour. We know from Louis de Broglie's hypothesis that every particle with momentum p can be assigned a wavelength. So it's quite natural to ask how big or how heavy can a particle be in order to keep its interference ability, what is the boundary, does there exist an upper bound at all ?

This question has been addressed by the experimental group Arndt, Zeilinger et al. [8] in a series of experiments with fullerenes. Fullerenes are C_{60}–molecules with a high spherical symmetry, resembling a football with a diameter of approximately $D \approx 1\,\text{nm}$ (given by the outer electron shell) and a mass of 720 amu, see Fig. 1.15.

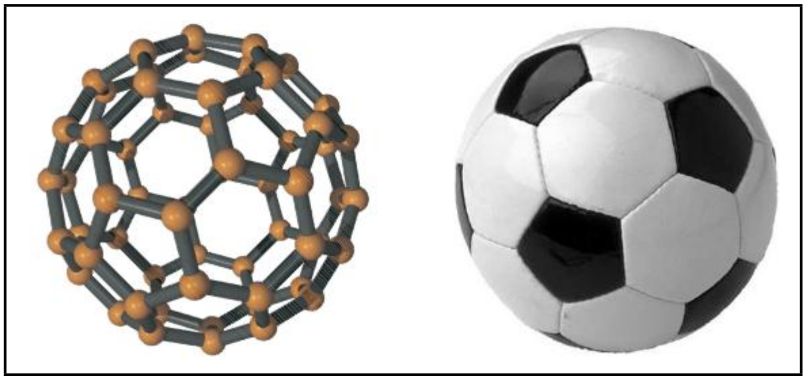

Figure 1.15: Illustration of the structure of fullerenes: 60 carbon atoms are arranged in pentagonal and hexagonal structures forming a sphere analogously to a football; picture used with courtesy of M. Arndt

In the experiment fullerenes are heated to about 900 °K in an oven from where they are emitted in a thermal velocity distribution. With an appropriate mechanism, e.g. a set of

---

[7] Certain interpretations of quantum mechanics might allow one to assign a definite path to an electron, though one would also need to introduce additional parameters which cannot be sufficiently measured and thus do not improve the predictive power of the theory. For references to the Bohmian interpretation see for example [9] and [10] and for the many worlds interpretation [11]

constantly rotating slotted disks, a narrow range of velocities is selected from the thermal distribution. After collimation the fullerenes pass a SiN lattice with gaps of $a = 50\,\mathrm{nm}$ and a grating period of $d = 100\,\mathrm{nm}$ and are finally identified with help of an ionizing laser detector, see Fig. 1.16.

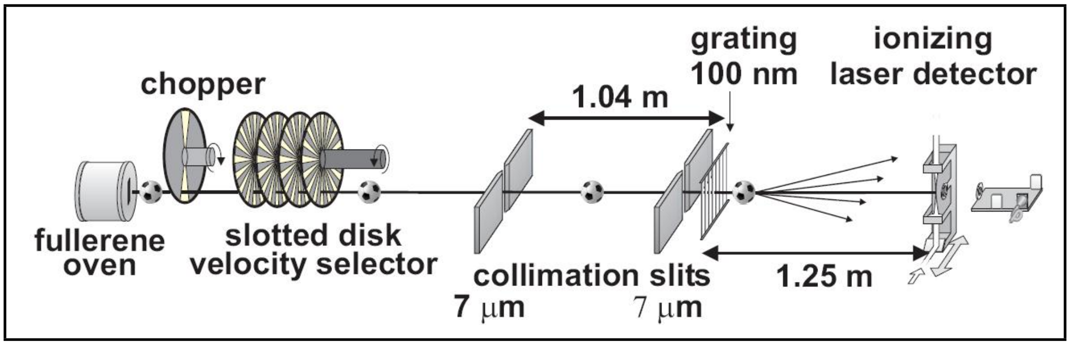

Figure 1.16: Experimental setup for fullerene diffraction; picture by courtesy of M. Arndt

For a velocity $v_{\max} = 220\,\mathrm{m/s}$ the de Broglie wavelength is

$$\lambda_{\mathrm{deBroglie}} = \frac{h}{mv} = 2,5\,\mathrm{pm} \approx \frac{1}{400}\,\mathrm{D}\,, \tag{1.32}$$

thus about 400 times smaller than the size of the particle.

As the angle between the central and the first order intensity maximum is very small (thus $\sin\Theta \approx \Theta$)

$$\Theta = \frac{\lambda_{\mathrm{deBroglie}}}{d} = 25\,\mu\mathrm{rad} = 5" \tag{1.33}$$

a good collimation is needed to increase the spatial (transverse) coherence. The whole experiment also has to be performed in a high vacuum chamber, pictured in Fig. 1.17, to prevent decoherence of the fullerene waves by collision with residual gas molecules.

The fullerenes are ionized by a laser beam in order to simplify the detection process, where an ion counter scans the target area, see Fig. 1.16.

We can also conclude that the fullerenes do not interfere with each other due to their high temperature, i.e. they are in high modes of vibration and rotation, thus occurring as classically distinguishable objects. Even so to prevent collisions the intensity of the beam is kept very low to ensure a mean distance of the single fullerenes of some mm (which is about 1000 times further than the intermolecular potentials reach). As a side effect of the high temperature the fullerenes emit photons[8] (black body radiation) but, fortunately, these photons don't influence the interference pattern because their wavelengths $\lambda \approx 5 - 20\,\mu\mathrm{m}$ are much bigger than the grating period (the double–slit). Thus we don't get any path information and the interference phenomenon remains.

---

[8] The probality for photon emission is rather small for the 900 K mentioned, for temperatures higher than 1500 K the probability of emitting (visible) photons is considerably higher but their effects can be neglected for the setup we have described here.

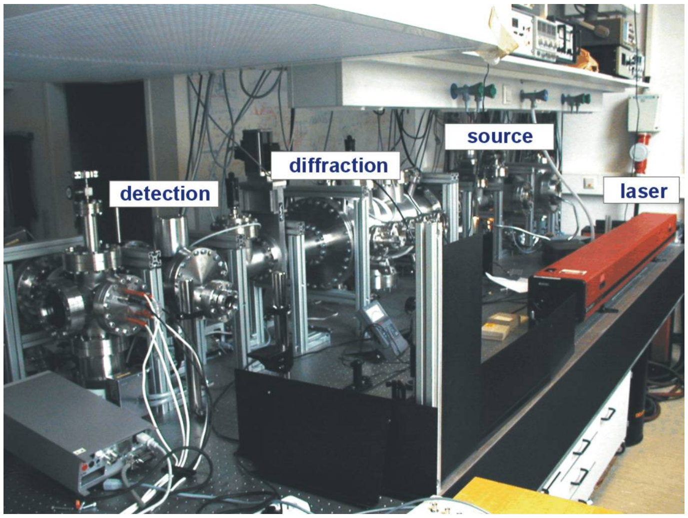

Figure 1.17: Photography of the experimental setup; picture by courtesy of M. Arndt

Looking finally at the result of the fullerene interferometry in Fig. 1.18 we see that the detection counts do very accurately fit the diffraction pattern[9] of the grating, thus demonstrating nicely that quite massive particles behave as true quantum systems.[10]

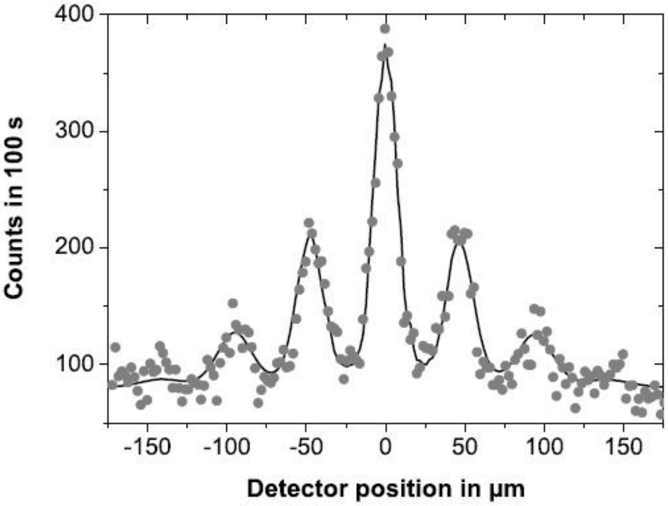

Figure 1.18: Results of the fullerene experiments, see [12]; picture by courtesy of M. Arndt

---

[9] The inquisitive reader might have noticed that for a grating with a 100 nm period and 50 nm wide slits, the second order diffraction maximum shouldn't exist at all, since the first order minimum of the single slit should cancel the second order maximum of the grating. The reason for its existence lies in the effective reduction of the slit width by Van-der-Waals interaction.

[10] It's quite amusing to notice, while the mass of a fullerene, also called bucky ball, does not agree with the requirements for footballs the symmetry and shape actually does, and furthermore the ratio between the diameter of a buckyball (1 nm) and the width of the diffraction grating slits (50 nm) compares quite well with the ratio between the diameter of a football (22 cm) and the width of the goal (732 cm).

Chapter 2

Time–Dependent Schrödinger Equation

2.1 Wave Function and Time–Dependent Schrödinger Equation

In Chapt.1 we discussed how to understand the wave–particle duality. According to Planck and Einstein the energy and frequency of light are related by $E = \hbar\omega$. De Broglie extended this dualism to massive particles by relating in addition the momentum to the wave vector $p = \hbar k$.

It was Erwin Schrödinger who reconsidered de Broglie's matter waves and discovered in 1926 a wave equation, the equation of motion corresponding to the wave nature of particles, which fits the physical observations. This differential equation is a fundamental equation and cannot be derived from first principles but we can make its form plausible.

Let us consider plane waves or rather wave packets which are of the form

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

$$\psi\,(t,x) \;=\; \int\limits_{-\infty}^{\infty} \frac{dk}{\sqrt{2\pi}}\,\tilde{\psi}\,(k)\,e^{i(kx-\omega t)} \quad . \tag{2.2}$$

We differentiate these waves with respect to t and x and recall the relations of wave and particle properties, Eq. (1.22)

$$i\hbar\,\frac{\partial}{\partial t}\,\psi\,(t,x) \;=\; \underbrace{\hbar\omega}_{E}\,\psi = E\,\psi \quad\to\quad \int \frac{dk}{\sqrt{2\pi}}\,\underbrace{\hbar\omega}_{E}\,\tilde{\psi}\,(k)\,e^{i(kx-\omega t)} \tag{2.3}$$

$$-i\hbar\,\nabla\,\psi\,(t,x) \;=\; \underbrace{\hbar k}_{p}\,\psi = p\,\psi \quad\to\quad \int \frac{dk}{\sqrt{2\pi}}\,\underbrace{\hbar k}_{p}\,\tilde{\psi}\,(k)\,e^{i(kx-\omega t)} \tag{2.4}$$

$$-\hbar^2\,\Delta\,\psi\,(t,x) \;=\; \underbrace{(\hbar k)^2}_{p^{\,2}}\,\psi = p^2\,\psi \quad\to\quad \int \frac{dk}{\sqrt{2\pi}}\,\underbrace{(\hbar k)^2}_{p^{\,2}}\,\tilde{\psi}\,(k)\,e^{i(kx-\omega t)} \quad . \tag{2.5}$$

The nonrelativistic energy–momentum relation for massive particles, where we assume for simplicity that the potential $V = V(x)$ is independent of time

$$E = \frac{p^2}{2m} + V(x)\,, \tag{2.6}$$

then provides a differential equation for ψ which Schrödinger assumed to hold quite generally for massive particles in an external potential $V(x)$[1].

Proposition 2.1.1 (Time-dependent Schrödinger equation)

$$i\hbar\,\tfrac{\partial}{\partial t}\,\psi\,(t,x) \;=\; \Big(-\tfrac{\hbar^2}{2m}\,\Delta \,+\, V(x)\Big)\,\psi\,(t,x) \;=\; H\,\psi\,(t,x)$$

The operator $H = -\frac{\hbar^2}{2m}\,\Delta + V(x)$ is called the *Hamiltonian* of the system, $\hbar$ is Planck's constant and m is the mass of the particle.

The solution $\psi(t,x)$ of the Schrödinger equation is called the *wave function*. It contains all the information about a physical system. The physical interpretation of the wave function is due to Max Born (see Prop. 1.7.1) and can be phrased in the following way: The probability for finding the particle in an interval $[x, x+dx]$ is given by $|\psi(t,x)|^2 dx$, which we have illustrated in Fig. 2.1.

Remark I: In deducing the Schrödinger equation within plane waves and wave packets we see that we can assign operators to the physical quantities energy and momentum

$$i\hbar\,\frac{\partial}{\partial t}\,\psi \;=\; E\,\psi \qquad \text{and} \qquad -\,i\hbar\,\nabla\,\psi \;=\; p\,\psi\,. \tag{2.7}$$

This *quantum mechanical correspondence* between physical quantities and operators

$$E \;\;\rightarrow\;\; i\hbar\,\frac{\partial}{\partial t} \qquad \text{and} \qquad p \;\;\rightarrow\;\; -i\hbar\,\nabla \tag{2.8}$$

is quite generally valid (when applied to any wave function). To classical relations correspond quantum mechanical ones. Thus we can quickly find the Schrödinger equation by starting with the classical Hamilton function $H(x,p) = \frac{p^2}{2m} + V = E$. We substitute Eq. (2.8), get exactly the Hamilton–operator H which applied to a wave function represents the Schrödinger equation.

---

[1] Historically, Schrödinger who was informed by Einstein about de Broglie's thesis gave a talk at the ETH Zürich in 1925. There Debye encouraged him to look for a wave equation and Schrödinger indeed found quickly one, nowadays known as Klein-Gordon equation, by using the relativistic energy–momentum relation, which however did not reproduce the spectrum of the H-atom correctly. During his Christmas vacations in a Pension near Arosa Schrödinger tried again by working now with nonrelativistic relations and discovered the successful "Schrödinger equation" which describes the spectrum correctly.

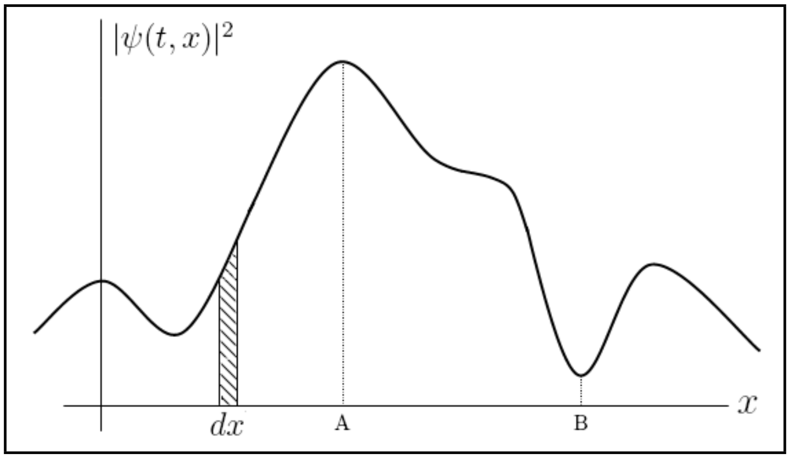

Figure 2.1: Probability interpretation of the wave function: The probability of finding the particle in the interval dx is given by the striped area. The probability to find the particle at location A is highest, whereas it is lowest for location B.

Remark II: The statistical interpretation of the wave function leads to a principal uncertainty to localize a particle. We cannot predict the definite measurement outcome for a specific particle, where it is localized at a certain time, thus we cannot assign a path to the particle.

Since we want to associate the wave function with a single particle, it is reasonable to require the probability to find the electron anywhere in space to be exactly equal to one. Since this may not be the case for all ψ a priori, we introduce a *normalization* condition:

$$\int\limits_{-\infty}^{\infty} dx \, |\psi(t,x)|^2 = 1 \,. \tag{2.9}$$

The normalization thus imposes a condition on the wave function, namely on its asymptotic behaviour, i.e. for $|x| \to \infty$:

$$\text{1-dim.} \quad |\psi(t,x)| \,\leq\, \frac{const.}{|x|^{\frac{1}{2}\,+\,\epsilon}} \qquad \text{3-dim.} \quad |\psi(t,\vec{x})| \,\leq\, \frac{const.}{|\vec{x}|^{1\,+\,\epsilon}} \;. \tag{2.10}$$

In other words, the wave function must be a square integrable function

$$\psi(t,x) \in L_2 \quad \text{space of all square integrable functions}\,. \tag{2.11}$$

We conclude that formal solutions of the Schrödinger equation, that are not normalizable, have no physical meaning in this context.

Example: Plane waves $e^{i(kx-\omega t)}$ are solutions of the Schrödinger equation, which can only be normalized in a box. For a box of length L we write

$$\begin{aligned} \psi(t,x) &= \frac{1}{\sqrt{L}}\, e^{i(kx-\omega t)} && \text{for } 0 \leq x \leq L \\ \psi(t,x) &= 0 && \text{outside of box} \end{aligned} \tag{2.12}$$

$$\Rightarrow \quad \int_{-\infty}^{\infty} dx\, \psi^*(t,x)\psi(t,x) \;=\; \frac{1}{L}\int_0^L dx\, \underbrace{e^{-i(kx-\omega t)}\, e^{i(kx-\omega t)}}_{1} = \frac{L}{L} = 1\,. \tag{2.13}$$

So we conclude that plane waves can only represent particles in a box, quite generally we need wave packets for their description.

Résumé:
The Schrödinger equation is a partial differential equation with the following properties:

1. 1^{st} **order in time** t: the wave function is determined by initial conditions, which is desirable from a physical point of view,
2. **linear in** ψ: $\Rightarrow$ superposition principle: linear combinations of solutions are again solutions, i.e. if ψ_1, ψ_2 are solutions $\Rightarrow \psi = c_1\psi_1 + c_2\psi_2$ with $c_1, c_2 \in \mathbb{C}$ is a solution,
3. **homogen**: the normalization holds for all times t.

2.2 Continuity Equation

With the probability density $|\psi(t,\vec{x})|^2$ we can associate a current density $\vec{j}(t,\vec{x})$ analogously to the charge density in electrodynamics. These densities will satisfy a continuity equation.

Let us start from the time dependent Schrödinger equation in 3 dimensions

$$i\hbar \frac{\partial}{\partial t}\, \psi(t,\vec{x}) \;=\; H\, \psi(t,\vec{x}) \tag{2.14}$$

and form its complex conjugate, where we use that the Hamiltonian H is a hermitian operator, which, in the present context[2], means that $H = H^*$, to gain

$$-i\hbar \frac{\partial}{\partial t}\, \psi^*(t,\vec{x}) \;=\; H\, \psi^*(t,\vec{x})\,. \tag{2.15}$$

---

[2] A mathematically more precise formulation of this criterion will be given later in Sec. 2.3

Definition 2.1 *The* ***probability density*** ρ *is given by the modulus squared of the wave function* ψ

$$\rho(t,\vec{x}) \;:=\; |\psi(t,\vec{x})|^2 \;=\; \psi^*\psi(t,\vec{x}) \;\; .$$

Performing the derivative of the probability density ρ with respect to time, we get

$$\frac{\partial}{\partial t}\,\rho(t,\vec{x}) \;=\; \frac{\partial}{\partial t}\,\psi^*\psi(t,\vec{x}) \;=\; \dot{\psi}^*\psi \,+\, \psi^*\dot{\psi}\,, \tag{2.16}$$

where we have used the notation $\frac{\partial}{\partial t}\psi \;=\; \dot{\psi}$. Inserting the Schrödinger equation (2.14) and its complex conjugate (2.15) into Eq. 2.16 we find

$$\begin{aligned} \dot{\rho}(t,\vec{x}) \;&=\; -\frac{1}{i\hbar}\,[\,(H\,\psi^*)\,\psi \,-\, \psi^* H\psi\,] \;= \\ &=\; \frac{\hbar}{2mi}\,[\,(\Delta\psi^*)\,\psi \,-\, \psi^*\Delta\psi\,] \,-\, \frac{1}{i\hbar}[\,V\psi^*\psi \,-\, \underbrace{\psi^* V\psi}_{V\psi^*\psi}\,] \;= \\ &=\; \frac{\hbar}{2mi}\,\vec{\nabla}\,[\,(\vec{\nabla}\psi^*)\psi \,-\, \psi^*\vec{\nabla}\psi\,]\,. \end{aligned} \tag{2.17}$$

Definition 2.2 *We define the* ***probability current*** $\vec{j}$ *as*

$$\vec{j}(t,\vec{x}) \;:=\; \tfrac{\hbar}{2mi}\,[\,\psi^*\vec{\nabla}\psi \,-\, \psi\vec{\nabla}\psi^*\,] \;\; .$$

With this definition we can write down a *continuity equation* analogously to the continuity equation in electrodynamics:

Theorem 2.1 (Continuity equation) $\frac{\partial}{\partial t}\,\rho(t,\vec{x}) \,+\, \vec{\nabla}\,\vec{j}(t,\vec{x}) \;=\; 0$

Theorem 2.1 implies the *conservation of probability* for all times

$$\int_{-\infty}^{\infty} d^3x\,|\psi(t,\vec{x})|^2 \;=\; 1 \qquad \forall\, t \geq 0\,, \tag{2.18}$$

analogously to the charge conservation. Thus the continuity equation, Theorem 2.1, means that a change of the probability in a volume V is balanced by a probability flux leaving V. So the probability and current densities behave like actual densities.

Proof: To prove the conservation of probability (2.18) we assume that ψ is normalized to 1 at $t = 0$ and that the probability density is only nonzero in a finite region $V \in \mathbb{R}^3$. Performing the derivative of ρ with respect to time and using the continuity equation, Theorem 2.1, we get

$$\frac{\partial}{\partial t} \int\limits_V d^3x\, \rho(t, \vec{x}) \;=\; - \int\limits_V d^3x\, \vec{\nabla}\, \vec{j}(t, \vec{x}) \;\overset{\text{Gauß}}{=}\; \int\limits_{\partial V} d\vec{f}\, \vec{j}(t, \vec{x})\,. \tag{2.19}$$

We have used the Theorem of Gauß to convert the integral over a 3–dimensional space V into an integral over its (2–dimensional) boundary ∂V which, for convenience, we can assume to be a sphere S_2 with radius R. By requiring the wave function to be square integrable it has to decrease stronger than $\frac{1}{R}$ for $R \to \infty$. Therefore the probability current has to fall off like $\frac{1}{R^3}$ (since the nabla operator is proportional[3] to $1/R$: $\vec{\nabla} \propto \frac{1}{R}$).

If we consider the last integral of Eq. (2.19) in the limit $R \to \infty$, using spherical coordinates, we find it to be proportional to

$$\propto \int dR\, R^2\, \frac{1}{R^3} \;\to\; 0 \qquad \text{for} \quad R \to \infty\,. \tag{2.20}$$

So we can conclude

$$\Rightarrow \int\limits_{-\infty}^{\infty} d^3x\, \rho(t, \vec{x}) \;=\; \text{const.} \;=\; 1 \qquad \text{q.e.d.} \tag{2.21}$$

Remark: The probability current $\vec{j}$ can be expressed by the *momentum operator* $\vec{P} = -i\hbar\vec{\nabla}$ by noting that

$$\begin{aligned} \vec{j}(t, \vec{x}) \;&=\; \frac{\hbar}{2mi}\, [\, \psi^* \vec{\nabla} \psi \;-\; \psi \vec{\nabla} \psi^* \,] \;= \\ &=\; \frac{1}{2m} [\, \psi^* (-i\hbar\vec{\nabla}) \psi \;+\; (\psi^* (-i\hbar\vec{\nabla}) \psi)^* \,] \;= \\ &=\; \frac{1}{m} \mathrm{Re}(\psi^* (-i\hbar\vec{\nabla}) \psi) \;=\; \frac{1}{m} \mathrm{Re}(\psi^* \vec{P} \psi)\,. \end{aligned} \tag{2.22}$$

As an example we calculate the probability current of a plane wave

$$\vec{j}(t, \vec{x}) \;=\; \frac{1}{m} \mathrm{Re}\big(e^{-i(\vec{k}\vec{x} - \omega t)} (-i\hbar\vec{\nabla}) e^{i(\vec{k}\vec{x} - \omega t)}\big) \;=\; \frac{1}{m} \mathrm{Re}\big(e^{-i(\vec{k}\vec{x} - \omega t)}\, \underbrace{\hbar\vec{k}}_{\vec{p}}\, e^{i(\vec{k}\vec{x} - \omega t)}\big) \;=\; \frac{\vec{p}}{m} \;=\; \vec{v} \tag{2.23}$$

and see that probability "flows" along the direction of the particle's velocity.

---

[3] Of course since this is an operator, this has to be understood as acting on a wave function.

2.3 Observables

In quantum mechanics there are two important concepts, which we use to describe physical objects:

1. **State of a system,** which is described by the wave function $\psi(t, \vec{x})$.

2. **Observables,** which are physical quantities we want to measure, like e.g. position, momentum, energy. These observables are represented by hermitian operators acting on the wave function.

The rules to combine these concepts are provided by quantum mechanics. We now introduce some mathematical background which we will need to set up these rules.

2.3.1 Operators

Definition 2.3 *A is called a **linear operator**, if for $A\,\psi_1(x) = \phi_1(x)$ and $A\,\psi_2(x) = \phi_2(x)$, where $\psi_1,\ \psi_2,\ \phi_1,\ \phi_2 \in L_2$, follows that*

$$A(c_1\psi_1 + c_2\psi_2) = c_1\phi_1 + c_2\phi_2 \qquad c_1, c_2 \in \mathbb{C}\,.$$

Remark: Linear operators, such as e.g. $\vec{\nabla}$, Δ, $\frac{\partial}{\partial t}$, $V(x)$, obey the law of distributivity (with respect to addition and multiplication in the space of linear operators) in the sense that

$$\begin{aligned} (A + B)\,\psi &= A\psi + B\,\psi \\ A\,B\,\psi &= A\,(B\,\psi) \end{aligned} \tag{2.24}$$

Furthermore the space of linear operators is equipped with an absorbing element (or zero operator) $\hat{0}$ and a neutral element (or unit element, or identity) $\mathbb{1}$, satisfying

$$\mathbb{1}\,\psi = \psi \quad \text{and} \quad \hat{0}\,\psi = 0 \quad \forall\,\psi \tag{2.25}$$

and which commute with every operator

$$\mathbb{1}A = A\,\mathbb{1} \quad \text{and} \quad \hat{0}A = A\,\hat{0} \quad \forall\,A\,. \tag{2.26}$$

The commutativity property, however, does not hold generally for arbitrary operators A and B

$$A\,B \neq B\,A\,. \tag{2.27}$$

2.3.2 Scalar Product

We will now introduce the notion of the *scalar product* used in quantum mechanics. This will be done in the so called "Dirac notation" of "bra's" and "ket's" which will be thoroughly investigated in Chapter 3.

Definition 2.4 *We define the **scalar product** of two wave functions ψ and ϕ by*

$$\langle \psi \,|\phi\rangle \;:=\; \int\limits_{-\infty}^{\infty} d^3x\, \psi^*(\vec{x})\phi(\vec{x}) \;.$$

Remark: The scalar product is formally written as the product[4] of "bra" $\langle\,.\,|$ with "ket" $|\,.\,\rangle$. Other notations of this scalar product include $(\psi\,,\,\phi)$ and $\langle\psi\,,\,\phi\rangle$. If the wave function depends on discrete variables, e.g. spin orientation, instead of continous ones, like e.g. position, the scalar product from Definition 2.4 reduces to

$$\langle \psi \,|\phi\rangle \;=\; \sum_k \psi_k^*\,\phi_k \;. \tag{2.28}$$

Properties of the scalar product:

$$\langle \phi \,|\psi\rangle^* \;=\; \langle \psi \,|\phi\rangle \tag{2.29}$$

$$\langle \phi \,|c_1\psi_1 \,+\, c_2\psi_2\rangle \;=\; c_1\,\langle \phi \,|\psi_1\rangle \,+\, c_2\,\langle \phi \,|\psi_2\rangle \quad \text{linear in "ket"} \tag{2.30}$$

$$\langle c_1\phi_1 \,+\, c_2\phi_2 \,|\psi\rangle \;=\; c_1^*\,\langle \phi_1 \,|\psi\rangle \,+\, c_2^*\,\langle \phi_2 \,|\psi\rangle \quad \text{antilinear in "bra"} \tag{2.31}$$

$$\langle \psi \,|\psi\rangle \;>\; 0 \quad \forall\,\psi \,\neq\, 0, \quad \langle \psi \,|\psi\rangle \;=\; 0 \quad \Leftrightarrow \quad \psi \,\equiv\, 0 \quad \text{positive definite} \tag{2.32}$$

Operators in the scalar product: [5]

Suppose A is an operator satisfying $A\,\psi \,=\, \phi$, where $\psi,\, \phi \,\in\, L_2$, then

$$\langle \psi \,|A\,\phi\rangle \;=\; \int d^3x\, \psi^*(\vec{x})A\,\phi(\vec{x}) \;. \tag{2.33}$$

---

[4] Strictly speaking, this is only a product in the sense that one object acts on the other (from the left side).

[5] We will omit the limits of the integration in the scalar product from now on and implicitly assume them to be ±∞ unless stated otherwise.

Definition 2.5 *$A^\dagger$ is called the* ***adjoint operator*** *to A, if $\forall\ \psi,\ \phi\ \in\ L_2$*

$$\left\langle A^\dagger\psi\,|\phi\right\rangle\;=\;\langle\psi\,|A\,\phi\rangle\;.$$

That means more explicitly

$$\int d^3x\,A^\dagger\,\psi^*(\vec{x})\phi(\vec{x})\;=\;\int d^3x\,\psi^*(\vec{x})A\,\phi(\vec{x})\,. \tag{2.34}$$

In matrix notation we can write the the adjoint[6] matrix as the complex conjugation of the transposed matrix or vice versa, i.e. $A^\dagger\;=\;(A^{\mathrm{T}})^*\;=\;(A^*)^{\mathrm{T}}$.

Here the operator in the scalar product is assumed to act to the right side[7], justifying the notation

$$\langle\psi|\,A\,|\phi\rangle\;=\;\langle\psi\,|A\,\phi\rangle\;. \tag{2.35}$$

Definition 2.6 *An Operator A is called* ***hermitian****, if*

$$A^\dagger\;=\;A$$

and the domains satisfy $D(A^\dagger)\supset D(A)$

If $D(A^\dagger)=D(A)$, then A is called self-adjoint.

Since the difference between hermitian and self-adjoint operators only plays a role for unbounded operators, we will neglect it here and use the terms synonymously.

Examples:

- **Position operator X** $\qquad X\psi(x)\;=\;x\,\psi(x)$

  defined for functions ψ, satisfying:

  $$\langle X\psi\,|X\psi\rangle\;=\;\int dx\,x^2\,|\psi(x)|^2\;<\;\infty\;. \tag{2.36}$$

  The generalization to 3 dimensions[8] is straightforward by labeling the above operator X as X_{i}, the i-th compononent of a vector operator $\vec{X}$. In this case we also require the integral in Eq. 2.36 to be finite when carried out over d^3x.

  The position operator X is a **hermitian** operator.

---

[6] The adjoint operator is also called the hermitian conjugate (mainly by physicists).

[7] This is not always the case but operators acting otherwise are often noted with arrows indicating their direction of action, e.g. $\overleftarrow{D}$, $\overleftrightarrow{D}$

[8] In the following we will mostly ignore the generalizations to 3 dimensions, since these can (mostly) be done by simple replacements, e.g. $x\to\vec{x}$, $dx\to d^3x$, at the most accompanied by constant factors due to Fourier transformations, but won't help us to gain any more insight in the physical or mathematical processes involved.

- **Momentum operator P** $\quad P\psi(x) \,=\, -i\hbar\nabla\,\psi(x)$

defined for functions ψ, satisfying:

$$\langle P\psi\,|P\psi\rangle \;=\; \hbar^2\int dx\,|\nabla\psi(x)|^2 \;<\;\infty\,. \tag{2.37}$$

For example, for a plane wave $\psi(x) = e^{ixp/\hbar}$, we get[9] $\quad P\psi(x) = p\,\psi(x)$.

The momentum operator also is **hermitian**: $P^\dagger = P$.

- **Hamiltonian H** $\quad H\psi(x) \;=\; \Big(-\frac{\hbar^2}{2m}\Delta \,+\, V(x)\Big)\;\psi(x) \;\;=\;\; E\psi(x)$

which is defined for $\psi\,'s$ such that

$$\langle H\psi\,|H\psi\rangle \;=\; E^2\int dx\,|\psi(x)|^2 \;<\;\infty\;. \tag{2.38}$$

2.4 Expectation Values

Having now explained some basic concepts in quantum mechanics, wave functions, scalar products and observables, we now need to discuss the rules that quantum mechanics provides to make physical predictions. Let's first remember the concepts from Section 2.3 and make them a little more precise:

Definition 2.7 *The **state** of a quantum mechanical system is represented by a vector $|\psi\rangle$ (ket) in a complex vectorspace, called Hilbertspace, equipped with a scalar product.*

Definition 2.8 ***Observables** in quantum mechanics are represented by hermitian operators acting on the state vectors in Hilbertspace.*

2.4.1 Expectation Values of Operators

We now want to find a way to relate these theoretical concepts with experimental outcomes, i.e., measurements. Since we already assume quantum mechanics to be of a statistical character (remember Born's probability interpretation, Prop. 1.7.1), let's take a brief look at some concepts of classical statistics.

If we want to calculate the average of a set of values $\{x_i\}$, we do so by calculating the *arithmetic mean*

$$\overline{x} \;=\; \frac{1}{n}\sum_{i=1}^{n} x_i\,. \tag{2.39}$$

---

[9] remember $p/\hbar = k$ (Eq. 1.22)

If the possible values are not equally probable, we use the *weighted arithmetic mean*

$$\overline{x} \;=\; \frac{\sum_{i=1}^{n} w_i \, x_i}{\sum_{i=1}^{n} w_i} \,. \tag{2.40}$$

For a (Riemann-integrable) function $f(x)$ we can generalize the arithmetic mean to

$$\overline{f} \;=\; \frac{1}{b-a} \int\limits_a^b dx \, f(x) \,. \tag{2.41}$$

Analogously to Eq. (2.40) we introduce a weight-function $w(x)$, which we can interpret as probability distribution, to get

$$\overline{f} \;=\; \frac{\int\limits_a^b dx \, w(x) \, f(x)}{\int\limits_a^b dx \, w(x)} \,. \tag{2.42}$$

Returning now to quantum mechanics, we notice the resemblance of Eq. (2.42) to the operator in the scalar product (Eq. 2.33), when interpreting $\psi^* \psi$ as a probability distribution. Thus we can write the weighted mean value of an operator A as

$$\langle A \rangle \;=\; \frac{\langle \psi | \, A \, | \psi \rangle}{\langle \psi \, | \psi \, \rangle} \,. \tag{2.43}$$

Since we require our wave function to be normalized to 1 (Eq. 2.9) we define:

Definition 2.9 *The* ***expectation value*** *of an observable A in the state $|\psi\rangle$ is given by*

$$\langle A \rangle_\psi \;=\; \langle \psi | \, A \, | \psi \rangle \,.$$

The expectation value $\langle A \rangle_\psi$ gives us the outcome of a series of measurements of the observable A of a physical system prepared in the state ψ. It thus has a physical meaning with respect to predicting measurement outcomes for an ensemble of identically prepared objects, even if the individual results are at random.

Properties of the expectation value:

$$\text{I)} \quad \langle \mathbb{1} \rangle \;=\; \langle \psi \,|\psi\rangle \;=\; 1 \tag{2.44}$$

$$\text{II)} \quad \begin{array}{ll} \langle A \rangle \in \mathbb{R} & \text{if } A \text{ is hermitian} \quad A = A^\dagger \\ \langle A \rangle \in i \cdot \mathbb{R} & \text{if } A \text{ is anti-hermitian} \quad A = -A^\dagger \end{array} \tag{2.45}$$

$$\text{III)} \quad \langle A \rangle \geq 0 \qquad \text{if } A \text{ is a positive operator} \quad A \geq 0 \tag{2.46}$$

$$\text{IV)} \quad \langle A \rangle \text{ is linear} \qquad \begin{array}{l} \langle \alpha A + \beta B\rangle \;=\; \alpha \,\langle A \rangle + \beta \,\langle B \rangle \\ \alpha, \beta \,\in\, \mathbb{C}, \quad A, B \;\text{ linear operators} \end{array} \tag{2.47}$$

Examples:

- **Expectation value of a potential $V(x)$**

$$\langle V(x) \rangle \;=\; \langle \psi |\, V(x)\, |\psi \rangle \;=\; \int dx\, V(x)\, |\psi(x)|^2 \tag{2.48}$$

In accordance with our consideration of classical statistics this is exactly the classical expectation for a potential controlled by a distribution $|\psi(x)|^2$.

- **Expectation value of the position operator X**

$$\langle X \rangle \;=\; \langle \psi |\, X\, |\psi \rangle \;=\; \int dx\, x\, |\psi(x)|^2 \tag{2.49}$$

For the postition operator we again get the classical result for the variable x, distributed by the probability density $|\psi(x)|^2$.

- **Expectation value of the momentum operator P**

$$\langle P \rangle \;=\; \langle \psi |\, P\, |\psi \rangle \;=\; -i\hbar \int dx\, \psi(x)^*\, \nabla\, \psi(x) \;=\; \int\limits_{-\infty}^{\infty} dp\, p\, |\tilde{\psi}(p)|^2 \tag{2.50}$$

$|\tilde{\psi}(p)|^2$ takes over the role of the probability density for the momentum-variable.

The function $\tilde{\psi}(p)$ is the Fourier-transform[10] of the wave function $\psi(x)$

$$\tilde{\psi}(p) \;=\; \frac{1}{\sqrt{2\pi\hbar}} \int dx\; \exp\,(-\frac{i}{\hbar} px)\psi(x)\,. \tag{2.51}$$

---

[10] The change of variables, $k = p/\hbar$, also requires the inclusion of $\hbar$ in the factor $1/\sqrt{2\pi}$ that normally occurs in the Fourier transformation.

Proof: To prove the last step of Eq. (2.50) we form the inverse transformation to Eq. (2.51)

$$\psi(x) \;=\; \frac{1}{\sqrt{2\pi\hbar}} \int dp \;\exp\,(\frac{i}{\hbar} p x) \tilde{\psi}(p) \tag{2.52}$$

and insert it into the expectation value of P, Eq. (2.50),

$$\begin{aligned} \langle\, P \,\rangle \;&=\; \frac{1}{2\pi\hbar} \int dx \int dp \int dp' \;\exp\,(-\frac{i}{\hbar} p' x) \tilde{\psi}^*(p') (-i\hbar\nabla) \;\exp\,(\frac{i}{\hbar} p x)\, \tilde{\psi}(p) \;= \\ &=\; \frac{1}{2\pi\hbar} \int dx \int dp \int dp' \;\exp\,(-\frac{i}{\hbar} p' x) \tilde{\psi}^*(p')\, p \;\; \exp\,(\frac{i}{\hbar} p x) \tilde{\psi}(p) \;= \\ &=\; \frac{1}{2\pi\hbar} \int dp \int dp' \;\int dx \; p \;\exp\,(\frac{i}{\hbar} (p - p') x)\, \tilde{\psi}^*(p')\, \tilde{\psi}(p)\,. \end{aligned} \tag{2.53}$$

Now we use the Dirac delta distribution

$$\delta(p - p') \;=\; \frac{1}{2\pi\hbar} \int dx \;\exp\,(\frac{i}{\hbar} (p - p') x)\,, \tag{2.54}$$

which is defined[11] by its action in an integration

$$\int\limits_{-\infty}^{\infty} dx'\, \delta(x - x')\, f(x') \;=\; f(x)\,. \tag{2.55}$$

With help of formulas (2.54) and (2.55) we can rewrite Eq. (2.53) as

$$\text{Eq. (2.53)} \;=\; \int dp \int dp' \;\, p\, \delta(p - p') \;\tilde{\psi}^*(p')\, \tilde{\psi}(p) \;=\; \int dp \; p\, |\tilde{\psi}(p)|^2 \qquad \text{q.e.d.} \tag{2.56}$$

2.4.2 Uncertainty of Observables

As we now know how to compute the quantum mechanical predictions for measurement outcomes, we want to go a step further by calculating how trustworthy our predictions are. That means, how much possible measurement outcomes will vary or deviate, leading to a variance, or a mean-square deviation, or an uncertainty. To do this we again investigate the methods of classical statistics.

Performing a series of measurements with outcomes $\{x_{\mathrm{i}}\}$ we already know from Eq. (2.39) and Eq. (2.40) how to calculate the mean values. If we started from there by calculating the deviations of the individual values from their mean value, then some of the deviations might cancel each other in the sum since these deviations could have either sign. Therefore it's better to compute the arithmetic mean of the squared deviations, the *variance*

$$\sigma^2 \;=\; \frac{1}{n} \sum_{i=1}^{n} (x_{\mathrm{i}} - \overline{x})^2\,. \tag{2.57}$$

---

[11] For a more rigorous definition we must refer the interested reader to a standard lecture, see e.g. Ref. [14], and/or textbook on functional analysis.

We then take its square-root, the *standard deviation*, which gives us a measure for the width of the distribution, i.e., how good the mean value approximates the individual results, see Fig. 2.2,

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{\mathrm{i}}-\overline{x})^2}\,. \tag{2.58}$$

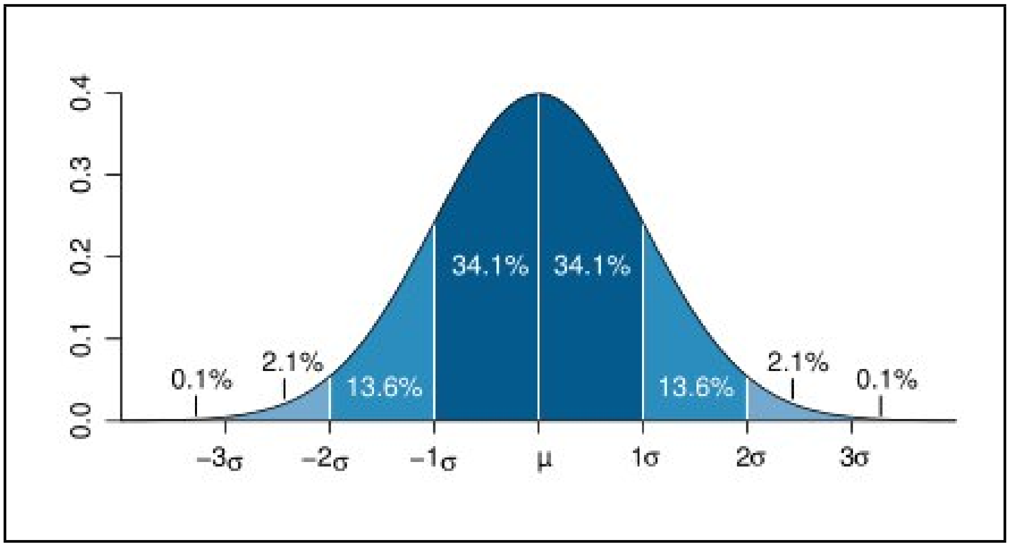

Figure 2.2: Standard deviation: In this normal (or "Gaussian") distribution we can see how the standard deviation σ quantifies the probability of finding the next result in σ-intervals around the mean value; figure from http://en.wikipedia.org/wiki/Image:Standard_deviation_diagram.svg

To convey this concepts to quantum mechanics we start by defining a deviation operator $\overline{A}$ associated with our observable A

$$\overline{A} := A - \langle\, A\,\rangle\ . \tag{2.59}$$

To compute the variance we calculate the expectation value of the squared deviation operator

$$\begin{aligned}\left\langle\, \overline{A}^{\,2} \right\rangle &= \langle\, A^2 - 2A\,\langle\, A\,\rangle + \langle\, A\,\rangle^2\,\rangle = \\ &= \langle\, A^2\,\rangle - 2\,\langle\, A\,\rangle\langle\, A\,\rangle + \langle\, A\,\rangle^2 = \\ &= \langle\, A^2\,\rangle - \langle\, A\,\rangle^2 =: (\Delta A)^2\,. \end{aligned} \tag{2.60}$$

Finally, we take the square root and define the variance or uncertainty.

Definition 2.10 *The* ***uncertainty*** *of an observable A in the state $|\psi\rangle$ is given by*

$$\Delta A \,=\, \sqrt{\langle\, A^2 \,\rangle_\psi \,-\, \langle\, A \,\rangle_\psi^2} \ .$$

The uncertainty of an observable represents a variation or fluctuation of possible measurement outcomes and it depends on the state $|\psi\rangle$ of the system. Although we followed here ideas of classical statistics we must stress that the uncertainty of an operator is not related to any technical imperfections of the measurement process. It is an important, genuine quantum feature !

The uncertainty can even vanish (much unlike technical imperfections) for certain states, which we will formulate by the following theorem:

Theorem 2.2 *The uncertainty of an observable A vanishes if, and only if, the state $|\psi\rangle$ of the system is an eigenvector of the operator A.*

$$\Delta A \,=\, 0 \qquad \Leftrightarrow \qquad A\,|\psi\rangle \,=\, a\,|\psi\rangle$$

The value $a \in \mathbb{R}$ is called the *eigenvalue*[12] of A, corresponding to the eigenvector $|\psi\rangle$.

Proof:

$$\begin{aligned}
\langle\, A \,\rangle_\psi \,&=\, \langle\psi|\, A\,|\psi\rangle \,=\, \langle\psi|\, a\,|\psi\rangle \overset{a\in\mathbb{R}}{=} a\,\langle\psi\,|\psi\rangle \,=\, a \qquad \Rightarrow \ \langle\, A \,\rangle^2 \,=\, a^2 \\
\left\langle\, A^2 \,\right\rangle_\psi \,&=\, \langle\psi|\, A\,A\,|\psi\rangle \,=\, \langle\psi|\, a^2\,|\psi\rangle \overset{a^2\in\mathbb{R}}{=} a^2\,\langle\psi\,|\psi\rangle \,=\, a^2 \qquad \Rightarrow \ \left\langle\, A^2 \,\right\rangle \,=\, a^2 \\
&\Rightarrow \quad \Delta A \,=\, \sqrt{\langle\, A^2 \,\rangle_\psi \,-\, \langle\, A \,\rangle_\psi^2} \,=\, \sqrt{a^2 \,-\, a^2} \,=\, 0 \qquad \text{q.e.d.}
\end{aligned}$$

2.5 Commutator of Operators

Up to now we have only considered individual operators or at the most products of the same operator, which didn't cause a problem. However, if we multiply different operators, we face an interesting consequence. Generally, those products are not commutative, which means that the order of application of the operators does matter, i.e.:

$$A\,B \,\neq\, B\,A\,. \tag{2.61}$$

---

[12] For general operators the eigenvalues would be complex numbers, but for hermitian operators, i.e. observables, the eigenvalues are real.

Take, for example, position and momentum operator X and P

$$X\,\psi(x) \,=\, x\,\psi(x) \qquad P\,\psi(x) \,=\, -i\hbar\nabla\,\psi(x)\,, \tag{2.62}$$

and consider the action of a combination of these operators on a wave function

$$P\,X\,\psi(x) \,=\, -i\hbar\nabla\,(x\,\psi(x)) \,=\, -i\hbar\,\psi(x) - i\hbar\,x\,\nabla\,\psi(x) \,=\, -i\hbar\,\psi(x) + X\,P\,\psi(x)\,, \tag{2.63}$$

then we can write formally

$$X\,P \,=\, P\,X \,+\, i\hbar\,. \tag{2.64}$$

To further formalize relations between operators we introduce the so-called *commutator* of operators.

Definition 2.11 *The* ***commutator*** $[\,.\,,\,.\,]$ *of two operators* A *and* B *is given by*

$$[\,A\,,\,B\,] \,=\, A\,B \,-\, B\,A\;.$$

Thus we rewrite Eq. (2.64) as commutator of X and P

$$[\,X\,,\,P\,] = i\hbar\,, \tag{2.65}$$

which is part of an important theorem.

Theorem 2.3 (Canonical commutation relations)
The components of the 3–dimensional position and momentum operator are related by

$$[\,X_{\mathrm{i}}\,,\,P_{\mathrm{j}}\,] \,=\, i\hbar\,\delta_{\mathrm{ij}}\;, \qquad [\,X_{\mathrm{i}}\,,\,X_{\mathrm{j}}\,] \,=\, [\,P_{\mathrm{i}}\,,\,P_{\mathrm{j}}\,] \,=\, 0$$

Properties of the commutator:

I) $$[A\,,B\,] = -\,[B\,,A\,] \qquad \Rightarrow [A\,,A\,] = 0 \tag{2.66}$$

II) $$[A\,,B\,] \quad \text{is linear in } A \text{ and } B \tag{2.67}$$

III) $$[A\,,B\,]^\dagger = \left[A^\dagger\,,B^\dagger\,\right] \tag{2.68}$$

IV) $$[A\,,BC\,] = B\,[A\,,C\,] + [A\,,B\,]\,C \tag{2.69}$$

V) $$[A\,,[B\,,C\,]] + [B\,,[C\,,A\,]] + [C\,,[A\,,B\,]] = 0 \qquad \text{Jacobi-identity} \tag{2.70}$$

VI) $$e^A\,B\,e^{-A} = B + [A\,,B\,] + \tfrac{1}{2!}\,[A\,,[A\,,B\,]] + \cdots \qquad \text{Baker-Campbell-Hausdorff formula} \tag{2.71}$$
$$\text{where} \quad e^A = \sum \tfrac{1}{n!}A^n$$

VII) $$e^A\,e^B = e^B\,e^A\,e^{[A\,,B\,]} \quad \text{and} \quad e^{A+B} = e^A\,e^B\,e^{-[A\,,B\,]/2} \tag{2.72}$$
$$\text{if} \quad [[A\,,B\,]\,,A\,] = [[A\,,B\,]\,,B\,] = 0$$

2.6 Uncertainty Principle

As we have learned in Section 2.4.2 the uncertainty of an observable is the quantum mechanical analogue of the classical standard deviation in the results of repeated measurements on identically prepared objects.

2.6.1 Uncertainty Relation for Observables

The non-commutativity of two observables has profound consequences, it is deeply related to the uncertainty of these observables. Let's recall the deviation operator (2.59) and the corresponding variance (2.60) then we can formulate the following theorem:

Theorem 2.4 (Uncertainty relation)
Let A and B be two observables, then the following inequality, the ***uncertainty relation****, is valid for all states*

$$\Delta A\ \Delta B \geq \tfrac{1}{2}|\,\langle[A\,,B\,]\rangle\,|$$

Whenever the commutator of two observables is nonvanishing, there is an uncertainty of these observables.

Proof: Let's start by defining the following non-hermitian operator Z

$$Z := \frac{\overline{A}}{\Delta A} + i\,\frac{\overline{B}}{\Delta B}\,, \tag{2.73}$$

where $\overline{A}$ and $\overline{B}$ are defined in Eq. (2.59). We consider the scalar product

$$\langle Z\psi\,|Z\psi\rangle \,=\, \langle\psi|\,Z^{\dagger}Z\,|\psi\rangle \,\geq\, 0\,, \tag{2.74}$$

which is definitely positive, Eq. (2.32), and use the definition of the adjoint operator, Definition 2.5. Inserting operator Z, Eq. (2.73), into the scalar product and noting that $Z^{\dagger} = (Z^{\mathrm{T}})^{*}$ we get

$$\langle\psi|\left(\frac{\overline{A}}{\Delta A} - i\,\frac{\overline{B}}{\Delta B}\right)\left(\frac{\overline{A}}{\Delta A} + i\,\frac{\overline{B}}{\Delta B}\right)|\psi\rangle \,\geq\, 0\,. \tag{2.75}$$

Multiplying the brackets and recalling that operators generally do not commute, gives

$$\langle\psi|\left(\frac{\overline{A}^{\,2}}{(\Delta A)^2} + i\,\frac{\overline{A}\,\overline{B} - \overline{B}\,\overline{A}}{\Delta A\Delta B} + \frac{\overline{B}^{\,2}}{(\Delta B)^2}\right)|\psi\rangle \,\geq\, 0\,. \tag{2.76}$$

Since the expectation values of $\overline{A}^{\,2}$ and $\overline{B}^{\,2}$ are just the squares of the corresponding variances, Eq. (2.60), the first and third term (each separately) give just 1. Using the commutator we may write

$$2 \,\geq\, -i\,\frac{\langle\psi|\left[\,\overline{A}\,,\,\overline{B}\,\right]|\psi\rangle}{\Delta A\,\Delta B} \,=\, -i\,\frac{\langle\psi|\,[\,A - \langle\,A\,\rangle\;,\,B - \langle\,B\,\rangle\,]\,|\psi\rangle}{\Delta A\,\Delta B} \,=\, -i\,\frac{\langle\psi|\,[\,A\,,\,B\,]\,|\psi\rangle}{\Delta A\,\Delta B}\,. \tag{2.77}$$

Rewriting Eq. (2.77) we finally obtain

$$\Delta A\,\Delta B \,\geq\, \frac{1}{2}\,|\,\langle[\,A\,,\,B\,]\rangle\,| \qquad \text{q.e.d.} \tag{2.78}$$

Note: Inequality (2.77) is certainly valid since the factor $(-i)$ is compensated by the pure imaginary scalar product $\langle\psi|\,[\,A\,,\,B\,]\,|\psi\rangle$, Eq. (2.45), due to the anti-hermiticity of the commutator of two hermitian operators

$$[\,A\,,\,B\,]^{\dagger} \,=\, \left[\,B^{\dagger}\,,\,A^{\dagger}\,\right] \,=\, [\,B\,,\,A\,] \,=\, -\,[\,A\,,\,B\,]\;. \tag{2.79}$$

Example: **Position and momentum**

Choosing in Theorem 2.3 as observables position and momentum $A = X$ and $B = P$ and recalling the canonical commutation relation (2.65) we recover the uncertainty relation we proposed in Section 1.6

$$\Delta X\,\Delta P \,\geq\, \frac{\hbar}{2}\,. \tag{2.80}$$

This inequality holds for all states, but it is saturated (it becomes an equality) for Gaussian wave packets (see Section 2.8) and, in particular, for the Gaussian ground state of the harmonic oscillator (see Section 5.1).

2.6.2 Energy–Time Uncertainty

In order to derive the energy–time uncertainty relation, we need to take a closer look on how time is treated within quantum mechanics. We know that in special relativity time and space are treated on equal footing by introducing the notion of so-called 4-vectors, $x^\mu = (\, ct \,, \vec{x}\,)^T$, $p^\mu = (\, \frac{E}{c} \,, \vec{p}\,)^T$, which include both time- and space-coordinates.

In nonrelativistic quantum mechanics, however, time t and space x have a different significance. Whereas x and p are the eigenvalues of the observables X and P respectively – physical quantities, for this reason – the time t is only an independent variable, a parameter. We are measuring space x and momentum p at a certain time t. Thus there does not exist such a thing as a "time operator" [13], whose eigenvalues represent possible time measurement results.

We can, however, analyze the change of a physical system within a short interval of time. Let's denote Δt as the time variation that is necessary for the system to undergo a substantial change.

As a measure for the change of the system we calculate the derivative with respect to time of the expectation value of some observable $A(t, x, p)$

$$\frac{d}{dt} \langle\, A \,\rangle \;=\; \frac{d}{dt} \langle\, \psi |\; A \;| \psi \,\rangle \;=\; \langle\, \dot{\psi} \,|\, A \,|\, \psi \,\rangle \,+\, \langle\, \psi \,|\, \dot{A} \,|\, \psi \,\rangle \,+\, \langle\, \psi \,|\, A \,|\, \dot{\psi} \,\rangle \;, \tag{2.81}$$

where we have used the notation $\dot{\psi} = \frac{\partial}{\partial t}\psi$. Using the Schrödinger equation (Eq. 2.14) and its complex conjugate (Eq. 2.15) we get

$$\frac{d}{dt} \langle\, A \,\rangle \;=\; \frac{i}{\hbar} \langle\, \psi |\; H\, A \,-\, A\, H \;| \psi \,\rangle \,+\, \langle\, \psi \,|\; \frac{\partial}{\partial t} A \;| \psi \,\rangle \;. \tag{2.82}$$

Theorem 2.5 (Time evolution of expectation value)

The time evolution of the expectation value is given by

- *the commutator of the observable with the Hamiltonian*
- *the time evolution of the observable*

$$\tfrac{d}{dt} \langle\, A \,\rangle \;=\; \tfrac{i}{\hbar} \langle\, [\, H \,,\, A\,] \,\rangle \,+\, \left\langle\, \tfrac{\partial}{\partial t} A \,\right\rangle$$

Remark I: Classical analogy

There is a classical analogy to Theorem 2.5. In classical (Hamiltonian) mechanics observables are represented by functions $f(p_i, q_i, t)$ on phase space, where p_i, q_i are the canonical coordinates. Such functions obey the following relation:

$$\frac{d}{dt} f(p_i, q_i, t) \;=\; \{\, f \,,\, H \,\} \,+\, \frac{\partial}{\partial t} f(p_i, q_i, t) \,, \tag{2.83}$$

---

[13] There exists, however, the notion of a time-translation operator, which we will encounter later on.

where H is Hamilton's principal function and $\{\,.\,,\,.\,\}$ denotes the Poisson-bracket

$$\{\, f\,,\, g\,\} \,=\, \sum_{i=1}^{N} \left[\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} \,-\, \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} \right] \,. \tag{2.84}$$

The transition from the Poisson bracket of classical mechanics to the commutator of quantum mechanics

$$\{\,.\,,\,.\,\} \qquad \Rightarrow \qquad -\frac{i}{\hbar}\,[\,.\,,\,.\,] \tag{2.85}$$

is quite generally valid and is called Dirac's rule. Compare, e.g., the canonical commutation relations (Theorem 2.3) or the Liouville equation (see theorem 9.2) with the corresponding classical relations.

Remark II: Conserved observables

We obtain a special case of Theorem 2.5 by considering observables that do not explicitly depend on time. Then the partial derivative of the operator A with respect to time vanishes and the time evolution of the expectation value is determined just by the commutator of A with the Hamiltonian H, i.e.

$$\frac{d}{dt}\,\langle\, A\,\rangle \,=\, \frac{i}{\hbar}\,\langle\,[\,H\,,\,A\,]\,\rangle\,. \tag{2.86}$$

If the commutator of observable A with Hamiltonian H vanishes its expectation value remains constant in time, or, we may say, the observable is conserved,

$$[\,H\,,\,A\,] \,=\, 0 \qquad \Rightarrow \qquad \frac{d}{dt}\,\langle\, A\,\rangle \,=\, 0 \qquad \Rightarrow \qquad \langle\, A\,\rangle \,=\, \text{const.} \tag{2.87}$$

Consider, for example, a free particle and its momentum P. The Hamiltonian for the free motion is given by

$$H \,=\, \frac{-\hbar^2}{2m}\Delta \,=\, \frac{P^2}{2m}\,, \tag{2.88}$$

which, according to Theorem 2.3, commutes with the momentum operator P. Thus, for free particles, momentum is a conserved quantity.

Let's now return to the general uncertainty relation for observables (Theorem 2.4) and assume the observable A to be explicitly independent of time[14]. Choosing $B = H$ we get

$$\Delta A\;\Delta H \,\geq\, \frac{1}{2}|\,\langle[\,A\,,\,H\,]\rangle\,| \overset{\text{Eq. (2.86)}}{=} \frac{\hbar}{2}\left|\frac{d}{dt}\,\langle\, A\,\rangle\right|\,. \tag{2.89}$$

We also formalize our statement from the beginning, that Δt is the time variation for a substantial change of the observable A, by defining

---

[14] This turns out to be a reasonable assumption as all practicable observables, e.g. position, momentum, angular momentum, satisfy this condition.

Definition 2.12 *Δt is the time passing, while the expectation value of the observable changes by one unit of ΔA*

$$\Delta A \,=\, \left|\tfrac{d}{dt}\,\langle\, A \,\rangle\right|\,\Delta t\,.$$

Subsequently we identify ΔH with ΔE, the uncertainty of energy, to obtain

$$\Delta E\,\Delta t \,\geq\, \frac{\hbar}{2}\,. \tag{2.90}$$

So we find, whenever an observable changes rapidly with time ($\frac{d}{dt}\,\langle\, A \,\rangle$ big) Δt will be small resulting in a high energy uncertainty, and vice versa. If the system remains stable (stationary), i.e. $\frac{d}{dt}\,\langle\, A \,\rangle \,=\, 0$, then clearly there is no energy uncertainty at all (the system has a "sharp" energy). Later on, we again look at this statement from the perspective of eigenvalue equations (Chapter 4).

2.7 Ehrenfest Theorem and Virial Theorem

We want to present two theorems that link quantum mechanics to classical mechanics.

2.7.1 Ehrenfest Theorem

Classical Newtonian mechanics follows from quantum mechanics in the sense of Ehrenfest's theorem, which we will derive in the following. We start by considering Theorem 2.5, the theorem for the time evolution of the expectation value of an operator. As operators we choose the position and momentum whose partial time-derivatives are vanishing. Thus we are left with the calculation of the corresponding commutators with the Hamiltonian of a particle in an exterior potential $V(\vec{x})$

$$H \,=\, \frac{-\hbar^2}{2m}\Delta \,+\, V(\vec{x}) \,=\, \frac{\vec{P}^2}{2m} \,+\, V(\vec{x})\,, \tag{2.91}$$

$$[\,H\,,\,X_{\rm i}\,] \,=\, \frac{1}{2m}\,\big[\,P_{\rm i}^2\,,\,X_{\rm i}\,\big] \,=\, \frac{P_{\rm i}}{2m}\underbrace{[\,P_{\rm i}\,,\,X_{\rm i}\,]}_{-i\hbar} + \underbrace{[\,P_{\rm i}\,,\,X_{\rm i}\,]}_{-i\hbar}\frac{P_{\rm i}}{2m} \,=\, -i\hbar\frac{P_{\rm i}}{m}\,, \tag{2.92}$$

$$[\,H\,,\,P_{\rm i}\,] \,=\, [\,V(x_{\rm i})\,,\,-i\hbar\nabla_{\rm i}\,] \,=\, -i\hbar V(x_{\rm i})\nabla_{\rm i} \,+\, i\hbar\underbrace{\nabla_{\rm i}V(x_{\rm i})}_{(\nabla V)+V\nabla} \,=\, i\hbar\,(\nabla_{\rm i}V(x_{\rm i}))\,. \tag{2.93}$$

In Eq. (2.92) we have used property (2.69) of a commutator and the canonical commutation relations (Theorem 2.3).

We now consider Theorem 2.5 and introduce a (conservative) force $\vec{F} \,=\, -\vec{\nabla}V(\vec{x})$ to get

$$\frac{d}{dt}\,\langle\,\vec{X}\,\rangle \,=\, \frac{1}{m}\,\langle\,\vec{P}\,\rangle \tag{2.94}$$

$$\frac{d}{dt}\,\langle\,\vec{P}\,\rangle \,=\, -\,\langle\,\vec{\nabla}V\,\rangle \,=\, \langle\,\vec{F}\,\rangle\,. \tag{2.95}$$

When combining both equations, Eq. (2.94) and Eq. (2.95), we can formulate the following theorem:

Theorem 2.6 (Ehrenfest theorem)
The classical equations of motion are valid for the mean values of the operators.

$$m\tfrac{d^2}{dt^2}\,\langle\,\vec{X}\,\rangle \;=\; -\,\langle\,\vec{\nabla}V\,\rangle \;=\; \langle\,\vec{F}\,\rangle$$

Remark I: Planck's constant $\hbar$ does not occur in the Ehrenfest theorem, which is already a sign for the classical nature of the statement.

Remark II: For the classical equation of motion to be valid also for the mean value $\langle\,\vec{x}\,\rangle$, we have to assume that

$$\langle\,\vec{F}(\vec{x})\,\rangle \;=\; \vec{F}(\langle\,\vec{x}\,\rangle)\,. \tag{2.96}$$

This condition is satisfied if the second and all higher derivatives of the force vanish, i.e. $\nabla^{\mathrm{n}}\vec{F} \,=\, 0,\;\; n \,=\, 2, 3, ...$; It is the case, e.g., for the free motion, the harmonic oscillator, and approximately for all slow changing forces $\vec{F}$.

2.7.2 Virial Theorem

The virial theorem is an important theorem in statistical mechanics, which relates the mean value of the kinetic energy of N particles to the one of the gradient of their potential

$$2\,\langle\,T\,\rangle_{\text{time average}} \;=\; \left\langle \sum_{i=1}^{N} \vec{\nabla}_{\mathrm{i}}V\cdot\vec{x}_{\mathrm{i}} \right\rangle_{\text{time average}}\,, \tag{2.97}$$

where $\vec{x}_{\mathrm{i}},\;\nabla_{\mathrm{i}}$ denote the position vector and the gradient of the i-th particle respectively. We can derive such a relation also within quantum mechanics.

We again start from Theorem 2.5 and choose as operator XP. Since the partial derivative with respect to time of XP vanishes we only have to consider the commutator with the Hamiltonian H (given by Eq. (2.91))

$$\begin{aligned} [\,H\,,\,XP\,] &\overset{\text{Eq. (2.69)}}{=} X\,[\,H\,,\,P\,] \,+\, [\,H\,,\,X\,]\,P \,= \\ \text{Eq. (2.92), (2.93)} \;\;\Rightarrow\quad &= X\,i\hbar\nabla\,V(x) \,-\, i\hbar\frac{P^2}{m} \,= \\ &= \frac{\hbar}{i}\left(2T \,-\, X\frac{d}{dx}V(x)\right)\,, \end{aligned} \tag{2.98}$$

where we denoted the kinetic energy by T and replaced the nabla operator by the total derivative. Thus we conclude that

$$\frac{d}{dt}\langle XP\rangle \,=\, 2\langle T\rangle - \langle X\frac{d}{dx}V(x)\rangle\,. \tag{2.99}$$

Finally we assume stationary states which satisfy[15] $\frac{d}{dt}\langle XP\rangle = 0$ and end up with the virial theorem.

Theorem 2.7 (Virial theorem) $\quad 2\langle T\rangle \,=\, \langle X\,\frac{d}{dx}\,V(x)\rangle$

Example: Harmonic oscillator
For the (one-dimensional) harmonic oscillator the potential energy is given by

$$V \,=\, \frac{m\omega^2x^2}{2} \qquad\Rightarrow\qquad \frac{d}{dx}V \,=\, m\omega^2 x\,. \tag{2.100}$$

Then the virial theorem tells us that

$$2\langle T\rangle \,=\, 2\langle\,\frac{m\omega^2x^2}{2}\,\rangle \qquad\Rightarrow\qquad \langle T\rangle \,=\, \langle V\rangle\,. \tag{2.101}$$

2.8 Time Evolution of the Wave Packet

In this section we want to investigate how we can relate the wave functions with the moving particles they represent and what conceptual insights arise from considering the time evolution.

2.8.1 Motion of Plane Waves and Wave Packets

Plane waves:
Let's now consider the motion associated with a plane wave (remember Eq. (2.1))

$$\psi\,(t,x) \,=\, e^{i(kx-\omega t)} \,=\, \exp\left(\frac{i}{\hbar}\,(px-Et)\right), \tag{2.102}$$

which is a solution of the free Schrödinger equation and we restrict the wave (and thus the particle) to a finite area to overcome difficulties with the normalization.

A fixed point on the wave (e.g., the maximum) corresponds to a fixed value of the argument in the exponential, i.e. $x \pm vt$ = constant. Let us calculate this velocity v,

---

[15] This condition can also be regarded as a form of Hamilton's principle, since the product of position and momentum has the dimension of an action, whose variation is required to vanish.

which is the *phase velocity* $\frac{\omega}{k}$, the propagation speed of the points of constant phase of the plane wave, see Fig. 2.3

$$v_{\text{phase}} := \frac{\omega}{k} = \frac{\hbar\omega}{\hbar k} = \frac{E}{p} = \frac{p^2}{2m}\frac{1}{p} = \frac{p}{2m} = \frac{1}{2}\,v_{\text{class}}\,. \tag{2.103}$$

We see that phase velocity of a plane wave travels at half the speed of the particle it should represent. Thus identifying the phase velocity with the velocity of the (classical) particle is not very satisfactory. However, the general solution of the time-dependent Schrödinger equation is a superposition of plane waves, a wave packet, and there the situation changes.

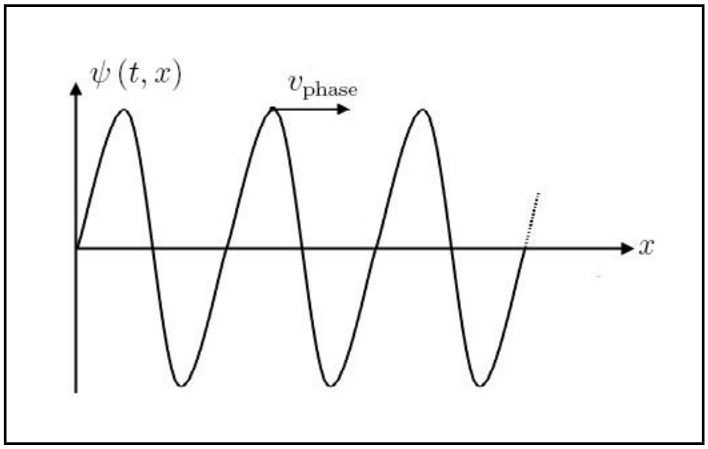

Figure 2.3: Phase velocity of a plane wave: points of constant phase, e.g., the wave crest moves with constant velocity in the x-direction.

Wave packets:
We have already encountered the wave packet in this chapter (recall Eq. (2.2)) as a solution of the Schrödinger equation

$$\psi\,(t,x) = \int\limits_{-\infty}^{\infty} \frac{dk}{\sqrt{2\pi}}\,\tilde{\psi}\,(k)\,e^{i(kx-\omega t)}\,. \tag{2.104}$$

Mathematically, it is the *Fourier transform* of a distribution $\tilde{\psi}\,(k)$ in k-space, modulated with an oscillation $\exp\,(-i\omega t)$. Thus it is just the superposition of plane waves, corresponding to different values of k, weighted with a function $\tilde{\psi}\,(k)\,/\sqrt{2\pi}$. Then the *inverse Fourier transform* at $t = 0$ has to be of the form

$$\tilde{\psi}(k) = \int\limits_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}}\,\psi(0,x)\,e^{-ikx}\,. \tag{2.105}$$

Generally, we expect dispersion, meaning that waves of different frequencies $\omega(k)$ travel at different speeds, i.e., they have different phase velocities. The *dispersion relation*, the relation between ω and k is given by

$$\omega(k) = \frac{\hbar k^2}{2m} \,. \tag{2.106}$$

Thus a phase velocity cannot represent the motion of the whole wave packet and we need to find another type of velocity. It is the *group velocity* v_{group}, as illustrated in Fig. 2.4.

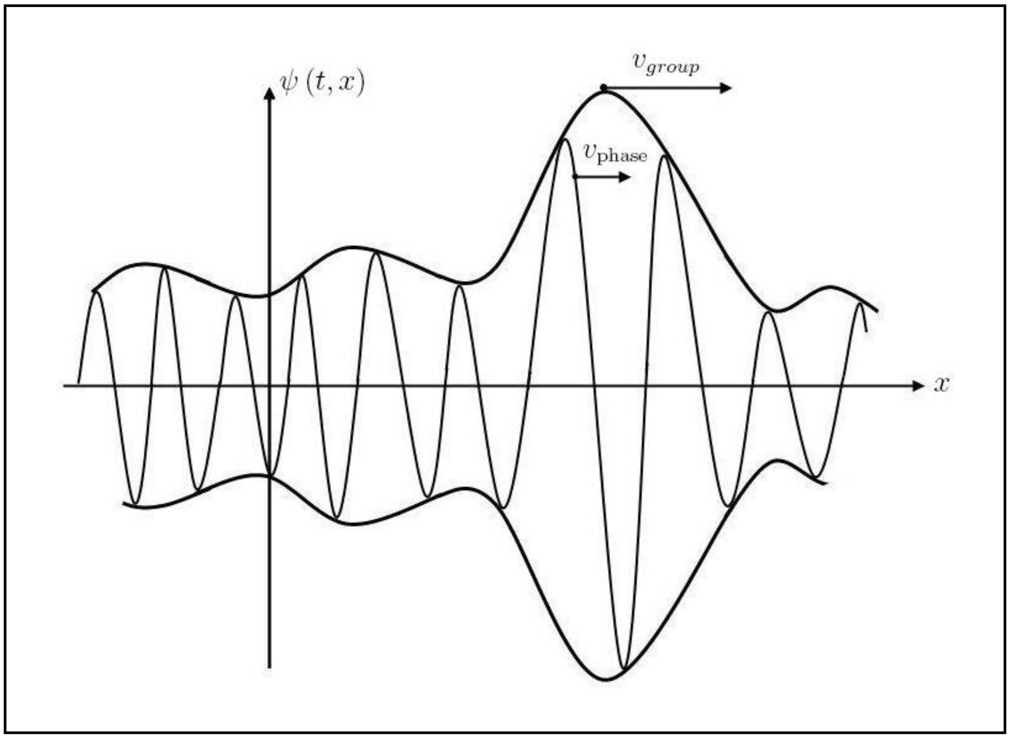

Figure 2.4: Wave packet: the phase velocities of different constituent plane waves suffer from dispersion whereas the group velocity accurately represents the motion of the enveloping function, i.e. the velocity of the peak.

Theorem 2.8 (Group velocity of wave packets)
Quantum mechanical wave packets $\psi(t,x)$, whose Fourier transform $\tilde{\psi}(k)$ is localized around a certain value k_0, move with the group velocity

$$v_{\text{group}} = \frac{d\omega}{dk}\Big|_{k=k_0} = 2\, v_{\text{phase}}$$

Proof:
The requirement of localization in the k-space can easily be fulfilled by a Gaussian distribution, like e.g.

$$\tilde{\psi}(k) \,=\, \text{const.} \,\times\, \exp\,(-\frac{\sigma^2(\mathrm{k}-\mathrm{k}_0)^2}{2})\,. \tag{2.107}$$

The dispersion relation, Eq. (2.106) we can expand as a Taylor series at k_0

$$\omega(k) \,=\, \omega(k_0) \,+\, \frac{d\omega}{dk}\bigg|_{k=k_0} (k-k_0) \,+\, \cdots \,\cong\, \omega_0 \,+\, \omega_0'(k-k_0)\,, \tag{2.108}$$

where we have used the notation $\omega_0 = \omega(k_0)$, $\omega_0' = \frac{d\omega}{dk}\big|_{k=k_0}$, and we may just keep the linear term due to the localized form of the packet in k-space, Eq. (2.107). With a change of variables $k \to k - k_0 = \bar{k}$ we can write the Fourier integral from Eq. (2.104) as

$$\psi\,(t,x) \,\cong\, \int\limits_{-\infty}^{\infty} \frac{d\bar{k}}{\sqrt{2\pi}}\,\tilde{\psi}\left(\bar{k}+k_0\right) \exp\left[i\left((\bar{k}+k_0)x - (\omega_0 + \omega_0'\bar{k})t\right)\right]. \tag{2.109}$$

By examining the form of Eq. (2.109) at different times, e.g. at $t = 0$ and $t > 0$, we get

$$\begin{aligned} \psi\,(t=0,x) \;&=\; \int \frac{d\bar{k}}{\sqrt{2\pi}}\,\tilde{\psi}\left(\bar{k}+k_0\right) \exp\left(i(\bar{k}\,+\,k_0)x\right), && (2.110)\\ \psi\,(t>0,x) \;&\cong\; e^{i(k_0\omega_0'-\omega_0)t} \int \frac{d\bar{k}}{\sqrt{2\pi}}\,\tilde{\psi}\left(\bar{k}+k_0\right) \exp\left(i(\bar{k}+k_0)(x-\omega_0't)\right). && (2.111) \end{aligned}$$

Except for the shift $x \to x - \omega_0't$ the integrals in (2.110) and (2.111) are identical, thus

$$\psi\,(t,x) \,\cong\, e^{i(k_0\omega_0'-\omega_0)t}\,\psi\,(t, x-\omega_0't)\;. \tag{2.112}$$

We conclude that after some time the wave packet propagated away from position x with the velocity $\omega_0' = v_{\text{group}}$. The phasefactor $\exp\left(i(k_0\omega_0' - \omega_0)t\right)$ in front of the integral can be ignored, since it will vanish in $|\psi|^2$. Finally, we simply calculate v_{group} using the dispersion relation (2.106) and find

$$v_{\text{group}} \,=\, \omega_0' \,=\, \frac{d\omega(k)}{dk} \,=\, \frac{d}{dk}\frac{\hbar k^2}{2m} \,=\, \frac{\hbar k}{m} \,=\, \frac{p}{m} \,=\, 2\,v_{\text{phase}} \qquad \text{q.e.d.} \tag{2.113}$$

2.8.2 Spreading of the Free Wave Packet

After having studied the motion of the wave packet, we now take a closer look at the localization of the particle associated with the wave packet. We again start from the general wave packet, from Eq. (2.104), but we use momentum and energy instead of wave number and frequency as variables

$$\psi\,(t,x) \,=\, \int\limits_{-\infty}^{\infty} \frac{dp}{\sqrt{2\pi\hbar}}\,\tilde{\psi}\,(p)\,\exp\left(\frac{i}{\hbar}\,(px-Et)\right). \tag{2.114}$$

We choose for the Fourier transform $\tilde{\psi}\,(p)$ a Gaussian distribution (on p-space)

$$\tilde{\psi}(p) \;=\; \underbrace{\left(\frac{\sigma}{\sqrt{\pi}\hbar}\right)^{1/2}}_{N} \exp\left(-\frac{\sigma^2(p-p_0)^2}{2\hbar^2}\right), \tag{2.115}$$

which ensures that, at an initial time $t = 0$, the wave packet is well localized (in x-space), i.e. σ is small[16]. We have also assumed that the Gaussian is normalized to one, which uniquely determines the normalization constant $N = \sqrt{\frac{\sigma}{\sqrt{\pi}\hbar}}$.

We then calculate the wave packet (Eq. 2.114)

$$\psi\,(t,x) \;=\; \frac{N}{\sqrt{2\pi\hbar}} \int\limits_{-\infty}^{\infty} dp\;\exp\left(-\frac{\sigma^2(p-p_0)^2}{2\hbar^2}\right)\,\exp\left(\frac{i}{\hbar}\,(px-Et)\right). \tag{2.116}$$

Using $E = \frac{p^2}{2m}$ and the abbreviations

$$a := \frac{\sigma^2}{2\hbar^2} + i\frac{t}{2m\hbar}, \qquad b := \frac{\sigma^2 p_0}{2\hbar^2} + i\frac{x}{2\hbar}, \qquad c := \frac{\sigma^2 p_0^2}{2\hbar^2} \tag{2.117}$$

we get

$$\psi\,(t,x) \;=\; \frac{N}{\sqrt{2\pi\hbar}} \int\limits_{-\infty}^{\infty} dp\;\exp\left(-a\left(p-\frac{b}{a}\right)^2 + \frac{b^2}{a} - c\right). \tag{2.118}$$

With a change of variables, $y = p - \frac{b}{a}$, and using the

$$\text{Gauss formula} \qquad \int\limits_{-\infty}^{\infty} dy\, e^{-\alpha y^2} \;=\; \sqrt{\frac{\pi}{\alpha}}, \tag{2.119}$$

we arrive at a simple form for the Gaussian wave packet

$$\psi\,(t,x) \;=\; \frac{N}{\sqrt{2a\hbar}} \exp\left(\frac{b^2}{a} - c\right). \tag{2.120}$$

We are now interested in the probability density $\rho(t,x) = |\psi(t,x)|^2$ of this wave packet

$$|\psi\,(t,x)|^2 \;=\; \frac{N^2}{2\hbar}\,\frac{1}{|a|}\,\exp\left(2\mathrm{Re}\left(\frac{\mathrm{b}^2}{\mathrm{a}} - \mathrm{c}\right)\right). \tag{2.121}$$

Thus we calculate the constituent parts from the abbreviations a, b and c step by step

$$|a|^2 \;=\; a^* a \;=\; (\mathrm{Re(a)})^2 \;+\; (\mathrm{Im(a)})^2 \;=\; \left(\frac{\sigma^2}{2\hbar^2}\right)^2 + \left(\frac{\mathrm{t}}{2\mathrm{m}\hbar}\right)^2 \;=\; \left(\frac{\sigma^2}{2\hbar^2}\right)^2 \left[1 + \Delta^2\right], \tag{2.122}$$

---

[16] The width of the p-space distribution on the other hand is proportional to $1/\sigma$, which means that the momentum spread is high at $t = 0$.

where $\Delta := \frac{\hbar}{m\sigma^2}t$, and using $v = \frac{p_0}{m} = v_{\text{group}}$ we finally have

$$2\,\mathrm{Re}\left(\frac{\mathrm{b}^2}{\mathrm{a}} - \mathrm{c}\right) = 2\,\mathrm{Re}\left(\frac{\mathrm{b}^2\mathrm{a}^* - \mathrm{c}|\mathrm{a}|^2}{|\mathrm{a}|^2}\right) = -\frac{(\mathrm{x} - \mathrm{vt})^2}{\sigma^2(1+\Delta^2)}\,. \tag{2.123}$$

Altogether we find the following form for the **probability density of the Gaussian wave packet as a function of time**

$$|\psi\,(t,x)|^2 = \frac{1}{\sqrt{\pi}\,\sigma\sqrt{1+\Delta^2}}\;\exp\left(-\frac{(x-vt)^2}{\sigma^2(1+\Delta^2)}\right). \tag{2.124}$$

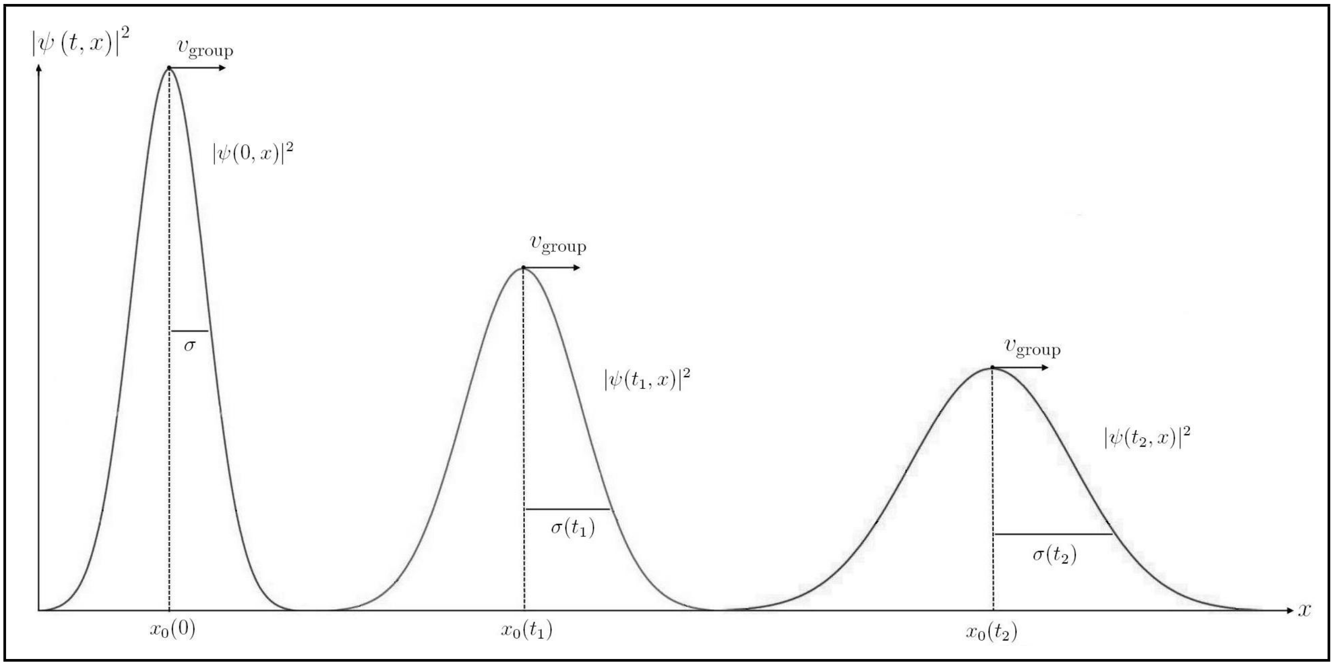

Figure 2.5: Spreading of the wave packet: As time passes, the envelope of the wave packet moves in the x-direction with velocity v_{group} and it gets more and more delocalized. Here $\sigma(t) := \sigma\sqrt{1+\Delta^2(t)}$

Result:

- The wave packet stays a Gaussian for all times, but its width, measured by $\sigma\sqrt{1+\Delta^2}$, increases with time, since $\Delta \propto t$. Thus the localizability of the particle represented by the wave packet decreases over time, which is termed the **spreading of the wave packet**, see Fig. 2.5.
- The maximum of the wave packet moves with $v_{\text{group}} = \frac{\hbar k}{m} = \frac{p_0}{m} = v_{\text{class}}$.
- The constituent plane waves of the superposition propagate with $v_{\text{phase}} = \frac{1}{2}v_{\text{group}}$.

Chapter 3

Mathematical Formalism of Quantum Mechanics

3.1 Hilbert Space

To gain a deeper understanding of quantum mechanics, we will need a more solid mathematical basis for our discussion. This we achieve by studying more thoroughly the structure of the space that underlies our physical objects, which as so often, is a vector space, the *Hilbert space*.

Definition 3.1 *A **Hilbert space** is a finite- or infinite-dimensional, linear vector space with scalar product in* $\mathbb{C}$*.*

The Hilbert space provides, so to speak, the playground for our analysis. In analogy to a classical phase space, the elements of the vector space, the vectors, are our possible physical states. But the physical quantities we want to measure, the observables, are now operators acting on the vectors. As mentioned in Definition 3.1, Hilbert spaces can be finite- or infinite-dimensional, as opposed to classical phase spaces (which are always $6n$-dimensional, where n is the particle number). We shall therefore investigate those two cases separately.

3.1.1 Finite–Dimensional Hilbert Spaces

In finite dimensions the vectors of a Hilbert space, denoted by $\mathcal{H}$, and the corresponding scalar products differ from the standard Euklidean case only by the choice of complex quantities $\mathbb{C}$ instead of real ones $\mathbb{R}$. It means that for vectors $x, y \in \mathcal{H}$

$$x = \begin{pmatrix} x^1 \\ \vdots \\ x^{\mathrm{n}} \end{pmatrix}, \; y = \begin{pmatrix} y^1 \\ \vdots \\ y^{\mathrm{n}} \end{pmatrix}, \qquad x^{\mathrm{i}}, \, y^{\mathrm{i}}, \in \mathbb{C} \, , \tag{3.1}$$

where $\mathcal{H}$ represents the n-dimensional Hilbert space under consideration, the scalar product can be written as

$$x\,y \,=\, x_{\rm i}\,y^{\rm i} \,=\, \langle\, x\,|\,y\,\rangle \,=\, \big(x_1, \cdots, x_{\rm n}\big)\begin{pmatrix} y^1 \\ \vdots \\ y^{\rm n} \end{pmatrix} \,=\, \sum_{i=1}^{n} x^{{\rm i}*} y^{\rm i} \,\in \mathbb{C}\,. \tag{3.2}$$

We have tried here to incorporate many different notations that are commonly used for scalar products, including the Einstein summation convention, which simply means that whenever an upper and a lower index are identical, the product is summed over. Whenever a covector[1], whose components[2] $x_{\rm i} \,=\, x^{{\rm i}*}$ are the complex conjugates of the corresponding vector components, is acting on a vector, with components $y^{\rm i}$, from the left side it yields a complex number. From the above form (3.2) of the scalar product we can immediately infer the following property

$$x\,y \,=\, (y\,x)^*\,. \tag{3.3}$$

Thus the norm $\|x\|$ of a vector is guaranteed to be real and positive

$$\|x\| \,=\, \sqrt{x\,x} \,=\, (x_{\rm i}\,x^{\rm i})^{1/2} \,=\, \left(\sum_{i=1}^{n} x^{{\rm i}*}\,x^{\rm i}\right)^{1/2} \,\in\, \mathbb{R}_{0+}\,. \tag{3.4}$$

Finally, the operators on this Hilbert space map one vector into another, i.e., they are linear transformations on the vector space, that can be represented by matrices.

$$x \,=\, A\,y \quad\Leftrightarrow\quad \begin{pmatrix} x^1 \\ \vdots \\ x^{\rm n} \end{pmatrix} = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} y^1 \\ \vdots \\ y^{\rm n} \end{pmatrix} \quad\Leftrightarrow\quad x^{\rm i} \,=\, A^{\rm i}{}_{\rm j}\,y^{\rm j}\,. \tag{3.5}$$

3.1.2 Infinite–Dimensional Hilbert Spaces

In infinite dimensions the vector space is generalized to a function space of complex valued functions, which now take the role of the state vectors[3]. The scalar product is then again defined as in Sec. 2.3.2

$$\langle\psi\,|\phi\rangle \,:=\, \int\limits_{-\infty}^{\infty} d^3x\,\psi^*(\vec{x})\phi(\vec{x})\,, \tag{3.6}$$

---

[1] A covector is a vector of the dual vectorspace which is denoted by a row- instead of a column-vector and by lower (covariant) instead of upper (contravariant) indices for its components. The dual vectorspace is the space of linear functionals over the vectorspace, which means that in the sense of the scalar product, every vector can be mapped to a (complex) number by the action of a covector.

[2] We assume here that the basis of the covectorspace is the dual basis to our original (vectorspace) basis.

[3] They still are vectors in the abstract sense that they are elements of a vector space, but it might be misleading to view them as some sort of "arrows"

where, analogously to the finite-dimensional case, $\langle\psi|$ is a covecter (or linear functional) acting on the vector $|\phi\rangle$, which we will discuss in Sec. 3.2. The scalar product has the properties stated in Eq. (2.29) – (2.32), from which we can see that the norm of the (state) vectors, which we assume to be square integrable (and in addition normalized to one), satisfies

$$\|\psi\|^2 \,=\, \langle\,\psi\,|\,\psi\,\rangle \,=\, \int dx\,|\psi(x)|^2 \,=\, 1 \,<\, \infty\,. \tag{3.7}$$

The operators on this Hilbert space then map one state into the other

$$|\,\phi\rangle \,=\, A\,|\,\psi\,\rangle\;. \tag{3.8}$$

Another important property of the Hilbert space in infinite dimensions is its *completeness*, which we will define in the following.

Hilbert space is a *complete* function space with scalar product in $\mathbb{C}$.

Definition 3.2 *A (function) space is called* ***complete*** *if every Cauchy-sequence has a limit in the space.*

Definition 3.3 *A sequence* $\{\psi_{\rm i}\}$ *is a* ***Cauchy sequence*** *if* $\forall\,\epsilon\in\mathbb{R}_+$ $\exists\,N\in\mathbb{N}$ *such that for all natural numbers* $n, m > N$*:* $\|\psi_{\rm n}\,-\,\psi_{\rm m}\| < \epsilon\,.$

This means that every sequence whose elements get even closer to each other as the sequence progresses – i.e. a Cauchy sequence – has a limit in the space, i.e., the space includes all the limits of its converging sequences and is therefore called complete. Such a property can be visualized as a space not missing any points. It guarantees that every function can be expanded with respect to the complete orthonormal basis chosen. Take for example the basis of plane waves

$$\left\{\frac{1}{\sqrt{2\pi}}e^{ikx}\right\}\,, \tag{3.9}$$

then every function $f(x)$ can[4] be expanded as

$$f(x) \,=\, \frac{1}{\sqrt{2\pi}}\int dk\,\tilde{f}(k)\,e^{ikx} \quad\text{and}\quad \tilde{f}(k) \,=\, \frac{1}{\sqrt{2\pi}}\int dx\,f(x)\,e^{-ikx}\,, \tag{3.10}$$

which is exactly the Fourier transformation of the function.

---

[4] if the following integral exists

3.2 Dirac Notation

In 1930 Paul Adrian Maurice Dirac introduced in his famous book "The principles of Quantum Mechanics" the so-called "bra-ket" notation[5] which has proven very useful, easy to handle, and became therefore the standard notation in quantum mechanics. Let's discuss it in more detail.

We have already explicitly formulated the scalar product of vectors in Hilbert spaces, see Eq. (3.2) and Eq. (3.6), and we used already the notation of "bra" $\langle\,.\,|$ and "ket" $|\,.\,\rangle$. These notions can now be used independently of each other (and of the scalar product) as vectors and covectors of a Hilbert space and its dual space. We will therefore denote the vectors of a Hilbert space by the ket

$$\text{"ket"} \qquad |\,\psi\,\rangle \,\in\, \mathcal{H}\,. \tag{3.11}$$

Since the Hilbert space $\mathcal{H}$ is a vector space it has a dual vector space $\mathcal{H}^*$, which is also called the space of linear functionals over the vector space. This means that the covectors are maps from the vector space into the associated field ($\mathbb{C}$ in this case), which here is exactly provided by the "bra". So the "bra"-vectors are the elements of the dual Hilbert space[6]

$$\text{"bra"} \qquad \langle\,\phi\,| \,\in\, \mathcal{H}^*\,. \tag{3.12}$$

The (anti-)isomorphism of the Hilbert space and its dual space guarantees, that we can write down the scalar product as "bra" acting on "ket" as we are used to. Also the notation is quite unambiguous, since to every vector $|\,\psi\,\rangle$ there is exactly one dual vector $\langle\,\psi\,|$ and the bidual vector is again $|\,\psi\,\rangle$[7]. So the covectors have a one-to-one correspondence to vectors, which are our physical states. Thus we can interpret the scalar product, the "bra-ket"

$$\langle\,\phi\,\,|\,\psi\,\rangle \;=\; \int dx\,\phi^*(x)\,\psi(x) \tag{3.13}$$

as a ***transition amplitude*** from the physical state $|\,\psi\,\rangle$ to $\langle\,\phi\,|$. Technically, we turn a vector into a covector by hermitian conjugation

$$|\,\psi\,\rangle^\dagger \;=\; \langle\,\psi\,|\;. \tag{3.14}$$

Thus, naturally, the operators on the Hilbert space are represented on the dual space by their adjoint operator (for hermitian operators these are identical)

$$A\,|\,\psi\,\rangle \quad\to\quad \langle\,\psi\,|\,A^\dagger\,. \tag{3.15}$$

---

[5] Also Dirac's delta-function was introduced by him in the same book.

[6] This formulation is a little bit sloppy, but it suffices for this course. The interested reader might look up Riesz's representation theorem, which gives the mathematically exact justification.

[7] The last property is called reflexivity of the Hilbert space.

If two operators are acting on a vector, their action on the dual vector is reversed

$$A\,B\,|\,\psi\,\rangle \quad \rightarrow \quad \langle\,\psi\,|\,B^{\dagger}\,A^{\dagger}\,. \tag{3.16}$$

3.3 Projection Operators

The insights from the last section now allow us to try out several combinations of vectors, covectors and operators, for example,

$$A\,B\,|\,\psi\,\rangle\,\langle\,\phi\,|\,C\,|\,\Phi\,\rangle\;. \tag{3.17}$$

Expression (3.17) can be interpreted in different ways, either the operators B and A act successively on the vector $|\,\psi\,\rangle$ multiplied by the scalar $\langle\,\phi\,|\,C\,|\,\Phi\,\rangle$, or the vector $|\,\Phi\,\rangle$ is acted upon by the operator $A\,B\,|\,\psi\,\rangle\,\langle\,\phi\,|\,C$. In this last case, however, we see that we can construct operators with the combination of ket and bra, called the *exterior product*

$$\text{operator:} \quad D\,:=\,|\,\psi\,\rangle\,\langle\,\phi\,| \qquad \text{adjoint operator:} \quad D^{\dagger}\,:=\,|\,\phi\,\rangle\,\langle\,\psi\,|\;. \tag{3.18}$$

If the vectors are now chosen to be their dual vectors respectively, we get an important class of operators, the *projection operators*

$$P\,:=\,|\,\psi\,\rangle\,\langle\,\psi\,|\;, \tag{3.19}$$

with the property

$$P^{\,2}\,=\,|\,\psi\,\rangle\underbrace{\langle\,\psi\,|\,\psi\,\rangle}_{1}\langle\,\psi\,|\,=\,|\,\psi\,\rangle\,\langle\,\psi\,|\,=\,P\,. \tag{3.20}$$

3.3.1 Projectors for Discrete Spectra

The projection operators are a very important tool to expand a vector in a complete orthonormal basis $|\,\psi_{\mathrm{n}}\,\rangle$. We can express each vector of the Hilbert space as a linear combination of the basis vectors with complex numbers c_{n}

$$|\,\psi\,\rangle\,=\,\sum_{n} c_{\mathrm{n}}\,|\,\psi_{\mathrm{n}}\,\rangle\;. \tag{3.21}$$

By applying $\langle\,\psi_{\mathrm{m}}\,|$ on both sides of Eq. (3.21) we get

$$\langle\,\psi_{\mathrm{m}}\,|\,\psi\,\rangle\,=\,\sum_{n} c_{\mathrm{n}}\underbrace{\langle\,\psi_{\mathrm{m}}\,|\,\psi_{\mathrm{n}}\,\rangle}_{\delta_{mn}}\,. \tag{3.22}$$

Thus the coefficients c_{n} are given by

$$c_{\mathrm{n}}\,=\,\langle\,\psi_{\mathrm{n}}\,|\,\psi\,\rangle\;, \tag{3.23}$$

i.e. the transition amplitudes of state $|\,\psi\,\rangle$ to states $|\,\psi_\mathrm{n}\,\rangle$. If we now insert Eq. (3.23) into Eq. (3.21)

$$|\,\psi\,\rangle \;=\; \sum_n \underbrace{|\,\psi_\mathrm{n}\,\rangle\,\langle\,\psi_\mathrm{n}|}_{P_\mathrm{n}}\;\psi\,\rangle\;, \tag{3.24}$$

we see that for a complete set of orthonormal basis vectors the orthogonal projectors satisfy the following *completeness relation*

$$\sum_n P_\mathrm{n} \;=\; \sum_n\,|\,\psi_\mathrm{n}\,\rangle\,\langle\,\psi_\mathrm{n}| \;=\; \mathbb{1}\,. \tag{3.25}$$

A projection operator P_n acting on an arbitrary state $|\,\psi\,\rangle$ will thus project the state to the state $|\,\psi_\mathrm{n}\,\rangle$ with a probability of $|\,\langle\,\psi_\mathrm{n}\,|\,\psi\,\rangle\,|^2$. Summarizing, the P_n satisfy

$$P_\mathrm{n}P_\mathrm{m} \;=\; \delta_{nm} \qquad \text{and} \qquad P_\mathrm{n}^{\,2} \;=\; P_\mathrm{n}\,. \tag{3.26}$$

Physically, this represents the class of projective measurements such as the measurement of the polarization of light.

Example: Polarization filter

Consider a photon, linearly polarized along the 45°-plane (with respect to the horizontal plane). We can then describe its polarization by a state vector

$$|\,\psi\,\rangle \;=\; \frac{1}{\sqrt{2}}\,(\,|\,H\,\rangle \,+\, |\,V\,\rangle\,)\;, \tag{3.27}$$

where $|\,H\,\rangle$ and $|\,V\,\rangle$ are the basis vectors of a 2–dimensional Hilbert space corresponding to horizontal and vertical polarization respectively. If we perform a measurement of the polarization by sending the photon through a polarization filter, e.g. in horizontal orientation, we get the measurement outcome by calculating the expectation value of the horizontal projector $|\,H\,\rangle\,\langle\,H\,|$. Lets first calculate the projection onto $|\,H\,\rangle$

$$|\,H\,\rangle\,\langle\,H\;|\,\psi\,\rangle \;=\; \frac{1}{\sqrt{2}}\left(|\,H\,\rangle\underbrace{\langle\,H\;|\,H\,\rangle}_{1} + |\,H\,\rangle\underbrace{\langle\,H\;|\,V\,\rangle}_{0}\right) \;=\; \frac{1}{\sqrt{2}}\,|\,H\,\rangle\;, \tag{3.28}$$

then we apply $\langle\,\psi\,|$ onto the left side to obtain the expectation value

$$\langle\,\psi\,|\,H\,\rangle\,\langle\,H\;|\,\psi\,\rangle \;=\; \frac{1}{2}\left(\underbrace{\langle\,H\,|\,H\,\rangle}_{1} + \underbrace{\langle\,V\,|\,H\,\rangle}_{0}\right) \;=\; \frac{1}{2}\,. \tag{3.29}$$

It's interesting to note that the expectation value of the projector is exactly the squared transition amplitude $\langle\,H\;|\,\psi\,\rangle$ – the transition probability. We conclude that the probability for the photon to pass the polarization filter is $\frac{1}{2}$.

3.3.2 Projectors for Continuous Spectra

Up to now we have only considered operators with discrete spectra, but we also have to explore projectors on, e.g., position eigenfunctions. In order to do so, we first change our notation. When working with the vectors of a system – usually, a complete orthonormal system (CONS) – the Dirac formalism allows a simplification in the notation, we may use just the label identifying the vector in the CONS, e.g.,

$$|\,\psi_{\mathrm{n}}\,\rangle \quad \rightarrow \quad |\,n\,\rangle \;, \qquad |\,\psi_{\xi}\,\rangle \quad \rightarrow \quad |\,\xi\,\rangle \;. \tag{3.30}$$

This labeling can be applied for vectors corresponding to a discrete spectrum n as well as to a continuous spectrum ξ, as exemplified in Eq. (3.30).

Position eigenvectors:
Let us assume vector $|\,\xi\,\rangle$ is a position eigenvector, this means

$$X\;|\,\xi\,\rangle \;=\; \xi\;|\,\xi\,\rangle \;, \tag{3.31}$$

where X is the position operator and ξ denotes the eigenvalue corresponding to the eigenvector $|\,\xi\,\rangle$. We will discuss spectra of operators and eigenvalue equations in more detail in Sec. 3.4. Projection operators are constructed in the same way as before by exterior multiplication but the corresponding properties have to be modified. The sum in the completeness relation needs to be replaced by an integral

$$\int\! d\xi\;|\,\xi\,\rangle\langle\,\xi\,| \;=\; \mathbb{1}\,, \tag{3.32}$$

and the orthogonality relation involves the Dirac delta function

$$\langle\,\xi'\,|\,\xi\,\rangle \;=\; \delta(\xi' - \xi)\,. \tag{3.33}$$

Now we can cast a further view on the wave functions of a Hilbert space by defining

$$\psi(x) \;:=\; \langle\,x\,|\,\psi\,\rangle \quad , \quad \psi^*(x) \;=\; \langle\,\psi\,|\,x\,\rangle \tag{3.34}$$

and analogously for the basis vectors of any discrete CONS

$$\psi_{\mathrm{n}}(x) \;:=\; \langle\,x\,|\,n\,\rangle \quad , \quad \psi^*_{\mathrm{n}}(x) \;=\; \langle\,n\,|\,x\,\rangle \;. \tag{3.35}$$

Eqs. (3.34) and (3.35) thus provide the basis-dependent notation – the wave function – of the abstract vector. That means, the abstract vector in the x-representation or the vector with respect to the chosen $|\,x\rangle$ basis. While the ket $|\,\psi\,\rangle$ denotes the basis-independent representation of the vector. The action of operators on the wave function is, however, independent of the basis as the operators only act on the ket. Therefore, equations such as Eq. (3.8) actually should be read as

$$A\,\psi(x) \;=\; \langle\,x|\;A\;|\,\psi\,\rangle \;=\; \langle\,x\,|\,\phi\rangle \;, \tag{3.36}$$

but keeping in mind that the operators act only on the ket we omit the bra $\langle\,x|$.

Let's consider now as wavefunction, the eigenfunction $\psi_\xi(x)$ of position operator X

$$\psi_\xi(x) \,=\, \langle\, x \,|\, \xi \,\rangle \,=\, \delta(x-\xi)\,, \tag{3.37}$$

which is, as expected for a position eigenfunction, perfectly localized. Technically, we need to keep in mind, that such an object is not square-integrable and thus not an element of the Hilbert space[8], but there are methods to deal with that problem, which shouldn't bother us at the moment.

Finally we can insert a CONS into Eq. (3.37) to gain the completeness relation of the wave functions in x-representation.

$$\begin{aligned} \langle\, x \,|\, x' \,\rangle &= \langle\, x \,|\, \mathbb{1} \,|\, x' \,\rangle \,=\, \langle\, x \,|\, \sum_n |\, n \,\rangle \langle\, n \,|\, |\, x' \,\rangle \,=\, \sum_n \langle\, x \,|\, n \,\rangle \langle\, n \,|\, x' \,\rangle \,= \\ &= \sum_n \psi_{\mathrm{n}}(x)\; \psi^*_{\mathrm{n}}(x') \,=\, \delta(x-x')\,. \end{aligned} \tag{3.38}$$

Momentum eigenvectors:
Using again the notation of Eq. (3.30) we now write $|\, p \,\rangle$ instead of $|\, \psi_{\mathrm{p}} \,\rangle$ in the eigenvalue equation for the momentum operator

$$P\,|\, p \rangle \,=\, p\,|\, p \rangle \;. \tag{3.39}$$

To be more precise we should write

$$P\,\psi_{\mathrm{p}} \,=\, \langle\, x \,|\; P\,|\, p \rangle \,=\, \langle\, x \,|\; p\,|\, p \rangle \;, \tag{3.40}$$

then we can insert the x-representation of the momentum operator – the quantum mechanical correspondence (2.8) – to calculate ψ_{p}

$$-i\hbar\nabla\,\psi_{\mathrm{p}} \,=\, p\,\psi_{\mathrm{p}}\,. \tag{3.41}$$

Here ψ_{p} depends on x, $\psi_{\mathrm{p}} \,=\, \psi_{\mathrm{p}}(x)$, and we have to solve an ordinary differential equation

$$\begin{aligned} -i\hbar\frac{d}{dx}\,\psi_{\mathrm{p}} \,=\, p\,\psi_{\mathrm{p}} \quad &\Rightarrow \quad \int\frac{d\psi_{\mathrm{p}}}{\psi_{\mathrm{p}}} \,=\, \frac{i}{\hbar}\,p\,dx \\ \ln\psi_{\mathrm{p}} \,=\, \frac{i}{\hbar}\,p\,x \,+\, \text{const.} \quad &\Rightarrow \quad \psi_{\mathrm{p}} \,=\, \text{const.} \times e^{\,ipx/\hbar}\,. \end{aligned} \tag{3.42}$$

The normalization of the wave function determines the constant explicitly and we can write the momentum eigenstate in the x-representation – the momentum eigenfunction – as

$$\langle\, x \,|\, p \rangle \,=\, \psi_{\mathrm{p}}(x) \,=\, \frac{1}{\sqrt{2\pi\hbar}}\, e^{\,ipx/\hbar}\,. \tag{3.43}$$

---

[8] This awkwardness can be overcome by redefining the structure of the underlying Hilbert space, which is then called the rigged Hilbert space (or Gelfand triple), but this considerations need not concern us here.

Using the completeness relation for position states, Eq (3.32), and the one for momentum eigenstates

$$\int dp \, |\, p\,\rangle \langle\, p\,| \,=\, \mathbb{1}\,, \tag{3.44}$$

we can prove the orthogonality property of the momentum and position eigenstates

$$\begin{aligned} \langle\, p' \,|\, p\,\rangle &= \langle\, p' \,|\, \mathbb{1} \,|\, p\,\rangle \,=\, \langle\, p' \,| \int dx \, |\, x\,\rangle \langle\, x\,| \,|\, p\,\rangle \,=\, \int dx \, \langle\, p' \,|\, x\,\rangle \langle\, x \,|\, p\,\rangle \,= \\ &= \int dx \; \psi^*_{\mathrm{p}'}(x)\; \psi_{\mathrm{p}}(x) \,=\, \frac{1}{2\pi\hbar} \int dx \; \exp\left(\frac{-ip'x}{\hbar}\right) \; \exp\left(\frac{ipx}{\hbar}\right) \,= \\ &= \frac{1}{2\pi\hbar} \int dx \; \exp\left(\frac{i}{\hbar}(p \,-\, p')x\right) \,=\, \delta(p - p')\,, \end{aligned} \tag{3.45}$$

$$\begin{aligned} \langle\, x' \,|\, x\,\rangle &= \langle\, x' \,|\, \mathbb{1} \,|\, x\,\rangle \,=\, \langle\, x' \,| \int dp \, |\, p\,\rangle \langle\, p\,| \,|\, x\,\rangle \,=\, \int dp \, \langle\, x' \,|\, p\,\rangle \langle\, p \,|\, x\,\rangle \,= \\ &= \int dp \; \psi_{\mathrm{p}}(x')\; \psi^*_{\mathrm{p}}(x) \,=\, \frac{1}{2\pi\hbar} \int dp \; \exp\left(\frac{ipx'}{\hbar}\right) \; \exp\left(\frac{-ipx}{\hbar}\right) \,= \\ &= \frac{1}{2\pi\hbar} \int dp \; \exp\left(\frac{i}{\hbar}(x' \,-\, x)p\right) \,=\, \delta(x' - x)\,. \end{aligned} \tag{3.46}$$

Expansion into a CONS:
Complete sets of orthonormal vectors are needed to expand a given vector of the Hilbert space. We start by expanding the ket

$$|\,\psi\,\rangle \,=\, \int dp \, |\, p\,\rangle \langle\, p \,|\, \psi\,\rangle \,=\, \int dp \, \tilde{\psi}(p) \, |\, p\,\rangle \;, \tag{3.47}$$

where we have used $\tilde{\psi}(p) \,=\, \langle\, p \,|\, \psi\,\rangle$ in total analogy to Eq. (3.34). The notation of the tilde for the wave function $\tilde{\psi}(p)$ explains itself once we calculate $\psi(x)$

$$\psi(x) \,=\, \langle\, x \,|\, \psi\,\rangle \,=\, \int dp \, \tilde{\psi}(p) \; \langle\, x \,|\, p\,\rangle \,=\, \int dp \, \tilde{\psi}(p) \, \psi_{\mathrm{p}}(x) \,=\, \frac{1}{\sqrt{2\pi\hbar}} \int dp \, \tilde{\psi}(p) \, e^{\,ipx/\hbar}\,. \tag{3.48}$$

We recognize Eq. (3.48) as the Fourier transformation, i.e., the wave functions of position and of momentum space are related to each other by a Fourier transformation. It often helps to simplify complicated calculations by transforming between those spaces.

3.4 Eigenvectors and Spectral Theorem

3.4.1 Eigenvalue Equation

Eigenvalue equations play an important role in quantum mechanics. Remember that possible measurement outcomes are given by the eigenvalues of hermitian operators. This means, performing a measurement of an observable leaves[9] the system in an eigenstate of

---

[9] The system stays in that particular state after the measurement, unless the state is changed by other influences (e.g. interactions, decoherence).

the corresponding operator. An arbitrary *eigenvalue equation* is of the form

$$A \,|\, j \,\rangle \,=\, \lambda_{\mathrm{j}} \,|\, j \,\rangle \;, \tag{3.49}$$

where A is a linear (hermitian if we consider observables) operator with eigenvalues λ_{j} and corresponding eigenvectors (eigenstates) $|\, j \,\rangle$.

Let's consider a finite-dimensional Hilbert space, e.g. 2–dimensional, and an operator A acting on it. Since we are in 2 dimensions we can represent A by a 2×2 matrix and use the following identity

$$(A \,-\, \lambda_{\mathrm{j}}\mathbb{1})\,|\, j \,\rangle \,=\, 0\,, \tag{3.50}$$

which is satisfied if the determinant of the operator acting on $|\, j \,\rangle$ vanishes

$$\det\,(A \,-\, \lambda_{\mathrm{j}}\mathbb{1}) \,=\, 0\,. \tag{3.51}$$

The determinant of Eq. (3.51) provides the so-called characteristic polynomial whose roots yield the possible eigenvalues of A. Making the ansatz

$$|\, i \,\rangle \,=\, \begin{pmatrix} a \\ b \end{pmatrix} \;, \qquad |\, j \,\rangle \,=\, \begin{pmatrix} c \\ d \end{pmatrix} \;, \qquad a, b, c, d \in \mathbb{C} \tag{3.52}$$

for the two (orthogonal) eigenvectors $|\, i \,\rangle$ and $|\, j \,\rangle$ and inserting them into the eigenvalue equation Eq. (3.49) gives two matrix equations, each for the pairs a, b and c, d. Two additional equations can be gained from the normalization of the vectors, which determines all constants completely

$$\langle\, i \,|\, i \,\rangle \,=\, \langle\, j \,|\, j \,\rangle \,=\, 1 \tag{3.53}$$

$$\Rightarrow \qquad |a|^2 \,+\, |b|^2 \,=\, 1 \;, \qquad |c|^2 \,+\, |d|^2 \,=\, 1\,. \tag{3.54}$$

3.4.2 Spectral Theorem

The spectral theorem gives us a precise formulation of the relation between an operator and its eigenvalues.

Definition 3.4 *A linear operator A is called normal, if $A^\dagger A \,=\, A\,A^\dagger$.*

Theorem 3.1 (Spectral theorem)
Every normal operator A can be expanded in its spectrum $\{\lambda_{\mathrm{i}}\}$ by projection operators

$$A \,=\, \sum_j \lambda_{\mathrm{j}}\,|\, j \,\rangle\,\langle\, j \,|$$

The theorem can also be made more specific for certain subclasses of operators, like hermitian or bounded/unbounded operators, but our formulation will suffice for this course. An important feature of the theorem, however, is its applicability to functions of operators

$$f(A) = \sum_j f(\lambda_{\mathrm{j}})\,|\,j\,\rangle\,\langle\,j\,|\;. \tag{3.55}$$

Example:
Consider a two–dimensional Hilbert space with an operator[10] we will call σ_{x} represented by the matrix

$$\sigma_{\mathrm{x}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.56}$$

We will proceed as mentioned in Section 3.4.1 to calculate the eigenvalues and eigenvectors of this matrix:

$$\det(\sigma_{\mathrm{x}} - \lambda \mathbb{1}) = \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1 = 0 \quad \Rightarrow \quad \lambda_{1,2} = \pm 1 \tag{3.57}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \lambda_1 \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \quad \Rightarrow \quad a = b \tag{3.58}$$

$$|a|^2 + |b|^2 = 1 \quad \Rightarrow \quad a = b = \frac{1}{\sqrt{2}}, \tag{3.59}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} c \\ d \end{pmatrix} = \lambda_2 \begin{pmatrix} c \\ d \end{pmatrix} = -\begin{pmatrix} c \\ d \end{pmatrix} \quad \Rightarrow \quad c = -d \tag{3.60}$$

$$|c|^2 + |d|^2 = 1 \quad \Rightarrow \quad c = -d = \frac{1}{\sqrt{2}}. \tag{3.61}$$

Thus we find for the eigenvectors and for the corresponding eigenvalue equations

$$|+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \sigma_{\mathrm{x}}\,|+\rangle = +\,|+\rangle\;, \tag{3.62}$$

$$|-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \qquad \sigma_{\mathrm{x}}\,|-\rangle = -\,|-\rangle\;. \tag{3.63}$$

To construct the projectors onto $|+\rangle$ and $|-\rangle$ we first take a look at the exterior product for arbitrary components. We remember from Eq. (3.2) that the components of the covector are the complex conjugates of the ordinary vector components and that the scalar product is constructed by multiplying a row-vector with column-vector. Here the order is reversed and we get a matrix instead of a scalar

$$|\,\psi\,\rangle\,\langle\,\phi\,| = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}\begin{pmatrix} \phi_1^* & \phi_2^* \end{pmatrix} = \begin{pmatrix} \psi_1\phi_1^* & \psi_1\phi_2^* \\ \psi_2\phi_1^* & \psi_2\phi_2^* \end{pmatrix}. \tag{3.64}$$

---

[10] This operator is one of the Pauli matrices which occur as spin observable.

Specifically, for the projection operators P_+ and P_- we find

$$P_+ \,=\, |+\rangle\langle+| \,=\, \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}1 & 1\end{pmatrix} \,=\, \frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}, \tag{3.65}$$

$$P_- \,=\, |-\rangle\langle-| \,=\, \frac{1}{2}\begin{pmatrix}1\\-1\end{pmatrix}\begin{pmatrix}1 & -1\end{pmatrix} \,=\, \frac{1}{2}\begin{pmatrix}1 & -1\\-1 & 1\end{pmatrix}. \tag{3.66}$$

The completeness and orthogonality relations then read

$$P_+ \,+\, P_- \,=\, \mathbb{1}\,, \qquad P_+\,P_- \,=\, 0\,, \tag{3.67}$$

and the spectral decomposition of σ_x (according to Theorem 3.1) is given by

$$\sigma_\mathrm{x} \,=\, \lambda_1\,P_+ \,+\, \lambda_2\,P_- \,=\, P_+ \,-\, P_-\,. \tag{3.68}$$

We can also calculate functions of σ_x with help of the spectral theorem, like e.g. the exponential of the matrix

$$\begin{aligned} \exp(-i\frac{\alpha}{2}\sigma_\mathrm{x}) \,&=\, \exp(-i\frac{\alpha}{2}(+1))\,P_+ \,+\, \exp(-i\frac{\alpha}{2}(-1))\,P_- \,= \\ =\, \frac{1}{2}\begin{pmatrix} e^{-i\frac{\alpha}{2}} + e^{i\frac{\alpha}{2}} & e^{-i\frac{\alpha}{2}} - e^{i\frac{\alpha}{2}} \\ e^{-i\frac{\alpha}{2}} - e^{i\frac{\alpha}{2}} & e^{-i\frac{\alpha}{2}} + e^{i\frac{\alpha}{2}} \end{pmatrix} \,&=\, \begin{pmatrix} \cos\left(\frac{\alpha}{2}\right) & -i\,\sin(\frac{\alpha}{2}) \\ -i\,\sin(\frac{\alpha}{2}) & \cos\left(\frac{\alpha}{2}\right) \end{pmatrix}. \end{aligned} \tag{3.69}$$

It corresponds to the rotation of a particle with spin $\frac{1}{2}$ around the x-axis by an angle α.

3.5 Résumé: Axioms and Physical Interpretation

1. The quantum state of a system is described by the **wave function** $\psi(t,x)$. The **probability** of finding the particle in the interval $[x, x{+}dx]$ is supplied by $|\psi(t,x)|^2 dx$.

2. The **time evolution** of the quantum state is given by the

   Schrödinger equation $\quad i\hbar\frac{\partial}{\partial t}\psi(t,x) \,=\, H\psi(t,x)\,,$

   where H denotes the Hamilton operator: $H = -\frac{\hbar^2}{2m}\Delta + V(x)\,.$

3. Physical quantities – **observables** – like e.g. E, P, or X, are represented by hermitian operators $A^\dagger = A$, having real eigenvalues. The eigenvalues a_n, determined by the

   eigenvalue equation $\quad A\,|\,n\,\rangle \,=\, a_\mathrm{n}\,|\,n\,\rangle\,,$

   correspond to the possible measurement outcomes.

4. Generally, a **quantum state** can be written as an

   expansion into a CONS $$|\,\psi\,\rangle \;=\; \sum_n c_{\rm n}\,|\,n\,\rangle\;,$$

   where the coefficients $c_{\rm n} \;=\; \langle\, n\,|\,\psi\,\rangle$ express the projections of the state $|\,\psi\,\rangle$ onto the eigenstates $|\,n\,\rangle$.

5. Experimentally, the **expectation value** of an observable is obtained by measuring a large number of identically prepared systems, each measurement provides – at random ! – a certain value $a_{\rm n}$, which will be averaged to a mean value.

   Theoretically, the mean value $\langle A\,\rangle_\psi$ of an observable A in the quantum state $|\,\psi\,\rangle$ is defined by

   expectation value $$\langle\;A\;\rangle_\psi \;=\; \int\limits_{-\infty}^{\infty} dx\,\psi^*(t,x)\,A\,\psi(t,x)\,,$$

   and expressed in terms of eigenvalues:

   expectation value $$\langle A\,\rangle_\psi \;=\; \langle\psi|\,A\,|\psi\rangle \;=\; \sum_n |c_{\rm n}|^2\,a_{\rm n}\,.$$

6. The **probability** for finding an eigenvalue $a_{\rm n}$ in the quantum state $|\,\psi\,\rangle$ is given by $|c_{\rm n}|^2 \;=\; |\,\langle\,n\,|\,\psi\,\rangle\,|^2$. After the measurement of A providing the value $a_{\rm n}$ the system remains in an eigenstate $|\,n\,\rangle$ – **projection postulate**.

7. The **variances** of the expectation values of operators satisfy the

   uncertainty relation $$\Delta A\,\Delta B \;\geq\; \tfrac{1}{2}|\,\langle\,[\,A\,,\,B\,]\,\rangle\,|\,,$$

   where $(\Delta A)^2 \;=\; \langle\,A^2\,\rangle \,-\, \langle\,A\,\rangle^2\,.$

Chapter 4

Time–Independent Schrödinger Equation

4.1 Stationary States

We consider again the time dependent Schrödinger equation (Prop. 2.1)

$$i\hbar \frac{\partial}{\partial t} \psi\,(t,x) = \left(-\frac{\hbar^2}{2m}\,\Delta \,+\, V(x)\right) \psi\,(t,x) = H\,\psi\,(t,x)\ , \tag{4.1}$$

where the potential in the Hamiltonian is assumed to be time independent $V \,=\, V(x)$. We calculate the solutions of this equation by using the *method of separation of variables*, i.e. we make the following ansatz for the solution $\psi(t,x)$:

$$\psi(t,x) \,=\, \psi(x)\, f(t) \tag{4.2}$$

and insert it into the time dependent Schrödinger equation, Eq. (4.1),

$$\begin{aligned} i\hbar\,\psi(x)\,\frac{\partial\, f(t)}{\partial t} &= -\frac{\hbar^2}{2m}\,\frac{\partial^2\psi(x)}{\partial x^2}\, f(t) \,+\, V(x)\,\psi(x) f(t) \qquad \Big|\ \cdot\ \frac{1}{\psi(x)\, f(t)} \\ i\hbar\,\frac{1}{f(t)}\,\frac{df(t)}{dt} &= -\frac{\hbar^2}{2m}\,\frac{1}{\psi(x)}\,\frac{d^2\psi(x)}{dx^2} \,+\, V(x)\,. \end{aligned} \tag{4.3}$$

Since now the left hand side in Eq. (4.3) is only dependent on t and the right hand side only on x, both sides must be equal to a constant, which we will call E, and we can thus solve each side independently. The left side yields

$$\begin{aligned} i\hbar\,\frac{1}{f(t)}\,\frac{df(t)}{dt} \,=\, E \quad &\Rightarrow \quad \frac{df}{f} \,=\, -\frac{i}{\hbar}\,E\,dt \\ \Rightarrow \quad \ln(f) \,=\, -\frac{i}{\hbar}\,E\,t \,+\, \text{const.} \quad &\Rightarrow \quad f \,=\, \text{const.}\, e^{-i\,E\,t/\hbar}\ . \end{aligned} \tag{4.4}$$

The constant in Eq. (4.4) will later on be absorbed into $\psi(x)$.

Then multiplying the right side of Eq. (4.3) with $\psi(x)$ we get

$$-\frac{\hbar^2}{2m}\frac{1}{\psi(x)}\frac{d^2\psi(x)}{dx^2} + V(x) \,=\, E \quad \Rightarrow \quad \underbrace{-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\,\psi(x)}_{H\,\psi(x)} \,=\, E\,\psi(x)\,. \tag{4.5}$$

The operators on the left express the Hamiltonian H acting on $\psi(x)$, which represents the *time independent Schrödinger equation.*

Theorem 4.1 (Time-independent Schrödinger equation)

$$H\,\psi(x) \,=\, E\,\psi(x)$$

$$\textit{where} \quad H \,=\, -\tfrac{\hbar^2}{2m}\Delta \,+\, V(x) \quad \textit{is the Hamiltonian}$$

Definition 4.1 *A state is called* ***stationary****, if it is represented by the wave function*

$$\psi(t,x) \,=\, \psi(x)\,e^{-i\,E\,t/\hbar}\,.$$

For such states the probability density is time independent

$$|\psi(t,x)|^2 \,=\, \psi^*(x)\,\psi(x)\,\underbrace{e^{i\,E\,t/\hbar}\,e^{-i\,E\,t/\hbar}}_{1} \,=\, |\psi(x)|^2\,. \tag{4.6}$$

The expectation values of observables $A(X,P)$ are time independent as well

$$\begin{aligned} \langle\, A(X,P)\,\rangle \;&=\; \int dx\,\psi^*(x)\,e^{i\,E\,t/\hbar}\,A(x,-i\hbar\frac{\partial}{\partial x})\,\psi(x)\,e^{-i\,E\,t/\hbar} \\ &=\; \int dx\,\psi^*(x)\,A(x,-i\hbar\frac{\partial}{\partial x})\,\psi(x)\,. \end{aligned} \tag{4.7}$$

Remark I: As a consequence, the eigenvalues of the Hamiltonian, which are the possible energy levels of the system, are clearly time independent.

To see it, just take $H(X,P)$ instead of $A(X,P)$ in Eq. (4.7) and use the time-independent Schrödinger equation (Theorem 4.1)

$$\langle\, H(X,P)\,\rangle \,=\, \int dx\,\psi^*(x)\,H\,\psi(x) \,=\, \int dx\,\psi^*(x)\,E\,\psi(x) \,=\, E\,\underbrace{\int dx\,\psi^*(x)\,\psi(x)}_{<\infty}\,. \tag{4.8}$$

Remark II: The normalization of the wavefunction will restrict the possible values of the constant E, the energy of the system, in the Schrödinger equation.

Two more interesting features about stationary states and the corresponding energies will be formulated here in the form of two lemmata, whose proofs we will leave as exercises.

Lemma 4.1 *For normalizable solutions $\psi(x)$ of the Schrödinger equation the energy E must be real, $E \in \mathbb{R}$.*

Lemma 4.2 *Solutions $\psi(x)$ of the time-independent Schrödinger equation can always be chosen to be real.*

Definition 4.2 *The **parity operator** $\mathcal{P}$ acting on a function $f(x)$ changes the sign of its argument:*

$$\mathcal{P}\, f(x) \,=\, f(-x)\,.$$

We conclude that even and odd functions are eigenfunctions of the parity operator

$$\mathcal{P}\,\psi_{\text{even}} \,=\, +\,\psi_{\text{even}} \qquad\qquad \mathcal{P}\,\psi_{\text{odd}} \,=\, -\,\psi_{\text{odd}}\,, \tag{4.9}$$

which we will use in the following theorem that will be helpful later on.

Theorem 4.2
For a symmetric potential $V(x) = V(-x)$ a basis of states can be chosen, that consists entirely of even and odd functions.

$$\psi_{\text{even}}(x) \,=\, \psi(x) \,+\, \psi(-x) \qquad\qquad \psi_{\text{odd}}(x) \,=\, \psi(x) \,-\, \psi(-x)$$

The proof for this theorem will be left as an exercise too.

4.2 Schrödinger Equation as Eigenvalue Equation

A subject concerning the time-independent Schrödinger equation we have not yet touched is its interpretation as an eigenvalue equation. Clearly, from its form we see that stationary states $|\,\psi\,\rangle$ are eigenvectors/eigenfunctions of the Hamiltonian H with eigenvalues E

$$H\,|\,\psi\,\rangle \,=\, E\,|\,\psi\,\rangle\;. \tag{4.10}$$

It implies the exact determination of the energy E. A stationary state has a precisely defined energy. Calculating the expectation value of the Hamiltonian for a stationary system just gives

$$\langle\,H\,\rangle \,=\, \langle\,\psi\,|\,H\,|\,\psi\,\rangle \,=\, \langle\,\psi\,|\,E\,|\,\psi\,\rangle \,=\, E\,\langle\,\psi\,|\,\psi\,\rangle \,=\, E\,. \tag{4.11}$$

Consequently, there is no energy uncertainty ΔE for these states

$$\Delta E \,=\, \Delta H \,=\, \sqrt{\langle\,H^2\,\rangle \,-\, \langle\,H\,\rangle^2} \,=\, \sqrt{E^2 \,-\, E^2} \,=\, 0\,. \tag{4.12}$$

Generally eigenvalue equations for linear operators take the form

$$A\,|\,\phi\,\rangle \,=\, a\,|\,\phi\,\rangle\;, \tag{4.13}$$

where a is an eigenvalue of the linear operator A with corresponding eigenvector $|\,\phi\,\rangle$. For hermitian operators there exist important statements about their eigenvalues and eigenfunctions.

Theorem 4.3
The eigenvalues of hermitian operators are real and the eigenvectors corresponding to different eigenvalues are orthogonal.

The proof is easy and again left as an exercise. The above theorem is vitally important for the spectrum $\{E_{\rm n}\}$ of the Hamiltonian, which is thereby guaranteed to be real

$$H\,|\,\psi_{\rm n}\,\rangle \,=\, E_{\rm n}\,|\,\psi_{\rm n}\,\rangle\;. \tag{4.14}$$

Using our notation $|\,\psi_{\rm n}\,\rangle \,\equiv\, |\,n\,\rangle$ the orthogonality and completeness relations (remember equations (3.25) and (3.26)) can be written as

$$\langle\,n\,|\,m\,\rangle \,=\, \delta_{\rm nm} \qquad\qquad \sum_n |\,n\,\rangle\,\langle\,n\,| \,=\, \mathbb{1}\,. \tag{4.15}$$

4.3 Expansion into Stationary States

Using the spectral theorem (Theorem 3.1) we can then expand a given state into a complete orthonormal system of energy eigenstates $|\, n\,\rangle$ exactly as outlined in Section 3.3.1

$$|\,\psi\,\rangle \;=\; \sum_n c_\mathrm{n}\,|\,n\,\rangle \qquad\qquad c_\mathrm{n} \;=\; \langle\, n\,|\,\psi\,\rangle\ . \tag{4.16}$$

By inserting a continous CONS of position eigenstates (Eq. (3.32)) into the transition amplidute the expansion coefficients c_n can be rewritten as

$$c_\mathrm{n} \;=\; \langle\, n\,|\,\psi\,\rangle \;=\; \int\! dx\ \langle\, n\,|\,x\,\rangle\,\langle\, x\,|\,\psi\,\rangle \;=\; \int\! dx\ \psi_\mathrm{n}^*(x)\ \psi(x)\,. \tag{4.17}$$

We can now extend the expansion from the time independent case to the time dependent one. We just remember the time dependent Schrödinger equation

$$i\hbar\,\frac{\partial}{\partial t}\,\psi\,(t,x) \;=\; H\,\psi\,(t,x)\ , \tag{4.18}$$

with a particular solution

$$\psi_\mathrm{n}(t,x) \;=\; \psi_\mathrm{n}(x)\,e^{-i\,E_\mathrm{n}\,t/\hbar}\,. \tag{4.19}$$

The general solution is then a superposition of particular solutions

$$\psi(t,x) \;=\; \sum_n c_\mathrm{n}\ \psi_\mathrm{n}(x)\,e^{-i\,E_\mathrm{n}\,t/\hbar}\,. \tag{4.20}$$

The expansion coefficients can easily be computed by setting $t = 0$ and taking the scalar product with $\psi_\mathrm{m}(x)$

$$\begin{aligned}\int\! dx\ \psi_\mathrm{m}^*(x)\ \psi(0,x) \;&=\; \int\! dx\ \psi_\mathrm{m}^*(x)\ \sum_n c_\mathrm{n}\ \psi_\mathrm{n}(x)\ \underbrace{e^{-i\,E_\mathrm{n}\,0/\hbar}}_{1}\\ \langle\,\psi_\mathrm{m}\,|\,\psi(t=0)\,\rangle \;&=\; \sum_n c_\mathrm{n}\ \underbrace{\int\! dx\ \psi_\mathrm{m}^*(x)\ \psi_\mathrm{n}(x)}_{\delta_\mathrm{mn}}\ .\end{aligned}$$

Thus the expansion coefficients are given by

$$c_\mathrm{n} \;=\; \langle\,\psi_\mathrm{n}\,|\,\psi(t=0)\,\rangle\ . \tag{4.21}$$

Physical interpretation of the expansion coefficients:
Let's consider an observable A with eigenstates ψ_n and eigenvalues a_n

$$A\;|\,\psi_\mathrm{n}\,\rangle \;=\; a_\mathrm{n}\;|\,\psi_\mathrm{n}\,\rangle\ . \tag{4.22}$$

If a system is in an eigenstate of this observable the expectation value (in this state) is equal to the corresponding eigenvalue

$$\langle\, A\,\rangle \;=\; \langle\,\psi_\mathrm{n}\,|\ A\;|\,\psi_\mathrm{n}\,\rangle \;=\; a_\mathrm{n}\;\langle\,\psi_\mathrm{n}\,|\;\psi_\mathrm{n}\,\rangle \;=\; a_\mathrm{n}\,. \tag{4.23}$$

Thus a measurement of the observable always produces the result a_{n} which implies that the uncertainty of the observable vanishes for this state $\Delta A = 0$. Furthermore the measurement leaves the state unchanged, the system remains in the eigenstate

$$|\,\psi_{\mathrm{n}}\,\rangle \quad \xrightarrow{A} \quad |\,\psi_{\mathrm{n}}\,\rangle\,. \tag{4.24}$$

If the system, however, is in a general state $|\,\psi\,\rangle$, which is a superposition of eigenstates, the expectation value is given by the sum of all eigenvalues, weighted with the modulus squared of the expansion coefficients

$$\begin{aligned}\langle\, A\,\rangle \;&=\; \langle\,\psi\,|\,A\,|\,\psi\,\rangle \,=\, \sum_{n}\sum_{m}\langle\, c_{\mathrm{m}}\psi_{\mathrm{m}}\,|\,A\,|\,c_{\mathrm{n}}\psi_{\mathrm{n}}\,\rangle \\ &=\; \sum_{n}\sum_{m} c_{\mathrm{m}}^{*}\,c_{\mathrm{n}}\,a_{\mathrm{n}}\,\underbrace{\langle\,\psi_{\mathrm{m}}\,|\,\psi_{\mathrm{n}}\,\rangle}_{\delta_{\mathrm{mn}}} \,=\, \sum_{n}\,|c_{\mathrm{n}}|^{2}\,a_{\mathrm{n}}\,.\end{aligned} \tag{4.25}$$

The expansion coefficients $c_{\mathrm{n}} = \langle\,\psi_{\mathrm{n}}\,|\,\psi\,\rangle$ can thus be regarded as a probability amplitude for the transition from a state ψ to an eigenstate ψ_{n} when the corresponding observable is measured. The actual transition probability is given by its modulus squared $|c_{\mathrm{n}}|^{2}$ – the *probability for measuring the result* a_{n} – which also obeys

$$\sum_{n}\,|c_{\mathrm{n}}|^{2} \,=\, 1\,. \tag{4.26}$$

So a measurement of an observable in a general state changes the state to one of the eigenstates of the observable. This process is often called the *reduction* or *collaps of the wave function*

$$|\,\psi\,\rangle \quad \xrightarrow{A} \quad |\,\psi_{\mathrm{n}}\,\rangle\,. \tag{4.27}$$

4.4 Infinite Potential Well

Our goal in the next sections is to calculate the energy eigenvalues and eigenfunctions for several Hamiltonians, i.e. for several potentials. Let us begin with the infinite potential well, represented by the potential $V(x)$, as illustrated in Fig. 4.1, such that

$$V(x) \,=\, \begin{cases} 0 & \text{for } x \in [\,0\,,\,L\,] \\ \infty & \text{else} \end{cases} \tag{4.28}$$

This means that the quantum object is limited to a certain region between $x = 0$ and $x = L$ where it moves freely but cannot ever leave. Thus mathematically we have

$$\psi(x) \,=\, 0 \quad \text{for} \quad x \notin [\,0\,,\,L\,]\,. \tag{4.29}$$

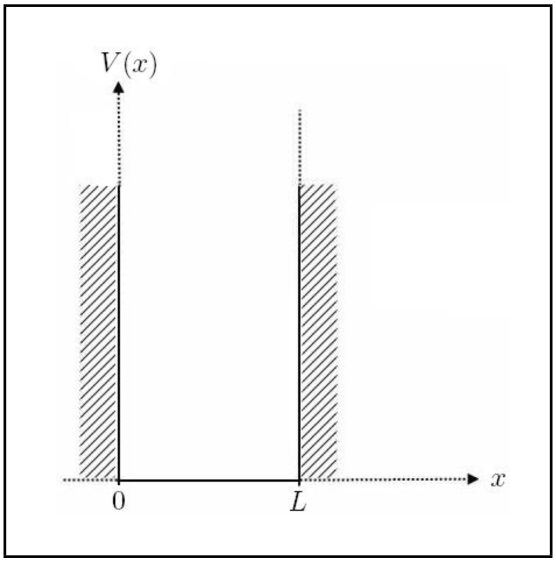

Figure 4.1: Infinite potential well: The potential is infinite outside the interval $[\,0\,,\,L\,]$, inside it vanishes. Therefore the only physically allowed region for a particle is inside the interval.

Furthermore, for the wave function to be continuous we have to require that it vanishes at the boundaries

$$\psi(0) \,=\, \psi(L) \,=\, 0\,. \tag{4.30}$$

The only region were particles are allowed is inside the well, where they behave like free particles, i.e. they are not exposed to a potential. Therefore we need to solve the free (time-independent) Schrödinger equation with the boundary conditions from Eq. (4.30)

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\,\psi(x) \,=\, E\,\psi(x)\,. \tag{4.31}$$

With the abbreviation

$$k^2 \,=\, \frac{2mE}{\hbar^2}\,, \qquad\qquad k \,=\, \frac{\sqrt{2mE}}{\hbar} \tag{4.32}$$

the free Schrödinger equation takes the following form

$$\frac{d^2}{dx^2}\,\psi(x) \,=\, -k^2\,\psi(x)\,, \tag{4.33}$$

where the general solution is well known, and given by

$$\psi(x) \,=\, a\,\sin(kx) \,+\, b\,\cos\,(kx)\,. \tag{4.34}$$

Here a and b are some constants that are yet to be determined by the boundary conditions, starting with $\psi(0) \,=\, 0$

$$0 \,=\, \psi(0) \,=\, a\,\underbrace{\sin(0)}_{0} + b\,\cos(0) \qquad \Rightarrow \qquad b \,=\, 0\,. \tag{4.35}$$

Exploiting the second boundary condition $\psi(L) = 0$, leads to discrete values of k

$$0 = \psi(L) = a \sin(kL) \qquad \Rightarrow \qquad kL = n\pi \qquad \Rightarrow \qquad k = \frac{n\pi}{L}\,, \tag{4.36}$$

where $n = 1, 2, 3, \ldots$ can be any natural number. Inserting our result into Eq. (4.32) and solving it with respect to E we see that the *energy is quantized*. Labeling the several energy levels by n we find

$$E_{\mathrm{n}} = \frac{n^2 \pi^2 \hbar^2}{2mL^2}\,. \tag{4.37}$$

Finally, the value of the constant a follows from the normalization of the wave function

$$\int_0^L dx\, |\psi|^2 = 1 \quad \Rightarrow \quad |a|^2 \int_0^L dx\; \sin^2(\frac{n\pi}{L}x) = 1 \quad \Rightarrow \quad |a|^2 = \frac{2}{L}\,. \tag{4.38}$$

Thus the bound states of the infinite potential well, which form a CONS, are then given by

$$\psi_{\mathrm{n}}(x) = \sqrt{\frac{2}{L}}\, \sin(\frac{n\pi}{L}x)\,. \tag{4.39}$$

For $n = 1$ we get the ground state energy and wave function E_1, ψ_1 of the infinite potential well, the higher states with $n > 1$ are called *excited states*.

4.5 Finite Potential Well

We now study a similar problem as in Section 4.4, but with the change that the potential walls are no longer infinitely high. Classically, a particle is trapped within the box, if its energy is lower than the height of the walls, i.e., it has zero probability of being found outside the box. We will see here that, quantum mechanically, the situation is different.

The time-independent Schrödinger equation is again our starting point where we now insert the following potential $V(x)$ into our Hamiltonian

$$V(x) = \begin{cases} -V_0 & \text{for } |x| \leq L \\ 0 & \text{for } |x| > L \end{cases} \tag{4.40}$$

For the possible energy range $E > -V_0$ we consider separately the two energy regions, $-V_0 < E < 0$ for the *bound states* and $E > 0$ for the *scattered states*. We also split the whole x-range into the three regions I, II, and III, where we solve the equations separately.

4.5.1 Bound States

Region I: $x < -L\,, \quad V(x) = 0$

Here we have again the free Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\,\psi(x) \,=\, E\,\psi(x)\,, \tag{4.41}$$

which we rewrite by substituting $\kappa = \frac{1}{\hbar}\sqrt{-2mE}$, where $\kappa > 0$ because $E < 0$,

$$\frac{d^2}{dx^2}\,\psi(x) \,=\, \kappa^2\,\psi(x)\,. \tag{4.42}$$

We already know that the general solution of Eq. (4.42) is given by

$$\psi(x) \,=\, A\,e^{-\kappa\,x} \,+\, B\,e^{\kappa\,x}\,, \tag{4.43}$$

where A and B are constants, yet to be determined. Since we are in the region where $x < -L < 0$ the exponent of the first term would ever increase for $x \to -\infty$. In order to keep the wave function normalizable we must demand that the constant A be identically zero, and we get as solution for region I

$$\boxed{\psi(x) \,=\, B\,e^{\kappa\,x}\,.} \tag{4.44}$$

Region II: $-L \leq x \leq L\,, \quad V(x) = -V_0$

In this region acts the potential and we have

$$\text{Schrödinger equation:} \qquad \left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} \,-\, V_0\right)\psi(x) \,=\, E\,\psi(x)\,, \tag{4.45}$$

which, by introducing a new constant q, becomes the equation

$$\frac{d^2}{dx^2}\,\psi(x) \,=\, -\,q^2\,\psi(x) \qquad\qquad q \,=\, \tfrac{1}{\hbar}\sqrt{2m(E \,+\, V_0)} \,>\, 0\,, \tag{4.46}$$

with the general solution

$$\boxed{\psi(x) \,=\, \bar{C}\,e^{-i\,q\,x} \,+\, \bar{D}\,e^{i\,q\,x} \,=\, C\,\sin(qx) \,+\, D\,\cos(qx)\,.} \tag{4.47}$$

Again, $C = i(\bar{D} - \bar{C})$ and $D = \bar{C} + \bar{D}$ are some constants.

Region Ⅲ: $x > L\,, \quad V(x) = 0$

Here we have the same case as in region I with the Schrödinger equation (4.41) and general solution

$$\psi(x) \,=\, F\, e^{-\kappa\, x} \,+\, G\, e^{\kappa\, x}\,. \tag{4.48}$$

But now, in order to keep the wave function normalizable we have to set $G = 0$ otherwise the corresponding exponent would ever increase for increasing x. We thus get as solution of region Ⅲ

$$\psi(x) \,=\, F\, e^{-\kappa\, x}\,. \tag{4.49}$$

Summary: Let's summarize the solutions for the energy range $-V_0 < E < 0$. We have

$$\psi(x) \,=\, \begin{cases} B\, e^{\kappa\, x} & \text{in region I} \\ C\, \sin(qx) \,+\, D\, \cos(qx) & \text{in region II} \\ F\, e^{-\kappa\, x} & \text{in region III} \end{cases} \tag{4.50}$$

Remark: The motion of a classical particle with energy $E < 0$ is strictly confined to region Ⅱ. A quantum mechanical particle, however, can penetrate into the classically forbidden regions I and Ⅲ, i.e. the probability density is non-vanishing, $|\psi(x)|^2 \neq 0$. How far the particle can penetrate depends on the respective energy, it can reach a depth of about

$$\Delta x \,\propto\, \frac{1}{\kappa} \,=\, \frac{\hbar}{\sqrt{-2mE}} \,\longrightarrow\, 0 \quad \text{for} \quad |E| \,\longrightarrow\, \infty\,, \tag{4.51}$$

which vanishes for large energies in deep potentials. Accordingly, there exists a momentum uncertainty which a classical particle would need to overcome the potential barrier

$$\Delta p \,\propto\, \frac{\hbar}{\Delta x} \,=\, \sqrt{-2mE}\,. \tag{4.52}$$

If we now remember Theorem 4.2 we can conclude that for the symmetric potential defined in Eq. (4.40) there is a family of even and odd solutions, which we will call $\psi^{(+)}(x)$ and $\psi^{(-)}(x)$, sketched in Fig. 4.2

$$\psi^{(+)}(x) \,=\, \begin{cases} B\, e^{\kappa\, x} & \text{I} \\ D\, \cos(qx) & \text{II} \\ B\, e^{-\kappa\, x} & \text{III} \end{cases} \qquad \psi^{(-)}(x) \,=\, \begin{cases} -B\, e^{\kappa\, x} & \text{I} \\ C\, \sin(qx) & \text{II} \\ B\, e^{-\kappa\, x} & \text{III} \end{cases} \tag{4.53}$$

At the boundaries of the potential well the functions that are solutions in their respective areas need to merge smoothly into each other. Mathematically this means, that the total wave function needs to be smooth, i.e. the values as well as the first derivatives of the respective partly solutions must match at $\pm\, L$.

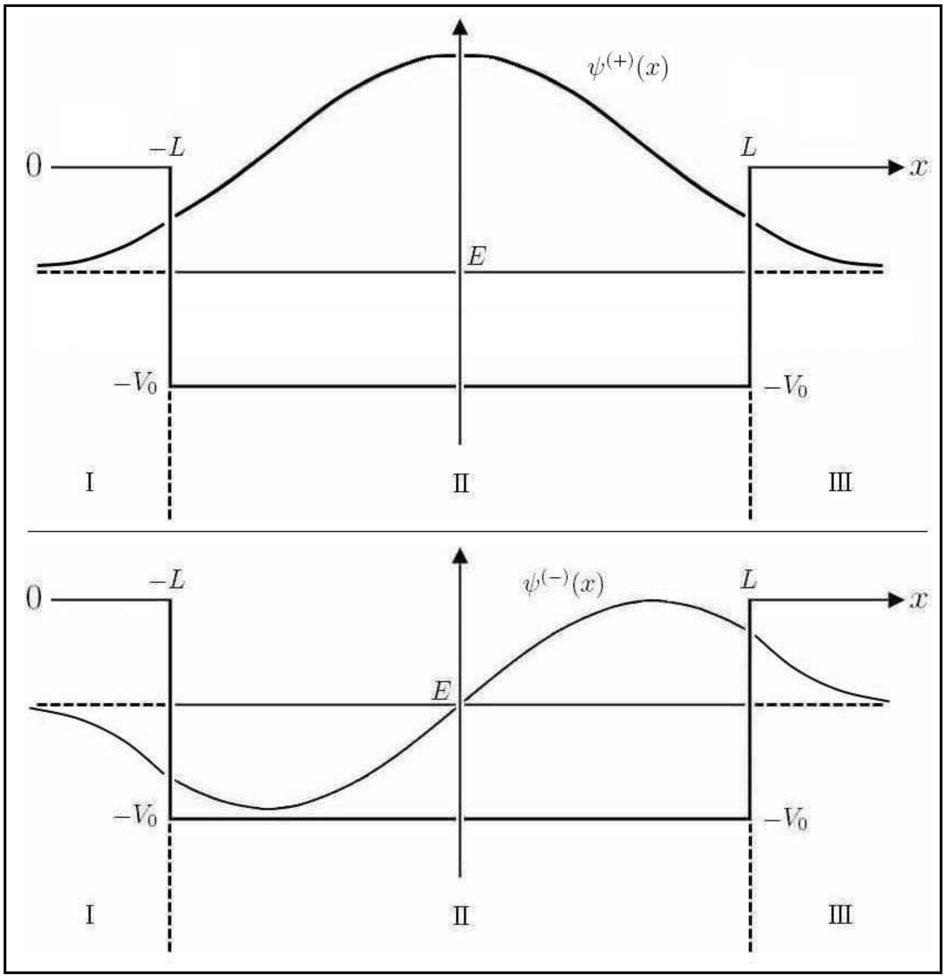

Figure 4.2: Bound state solutions: The bound states can be split into even and odd solutions, both solutions together with their first derivatives must be continuous with respect to x.

We can summarize these two requirements into the statement, that the *logarithmic derivative* of the wave function must be continuous

$$\text{logarithmic derivative:} \quad \frac{d}{dx}\,\ln(\psi(x)) \,=\, \frac{\psi'(x)}{\psi(x)} \quad \text{continuous}\,. \tag{4.54}$$

For the even solutions[1] this translates to

$$\left.\frac{\psi^{(+)\prime}(x)}{\psi^{(+)}(x)}\right|_{x\to L} \;\Rightarrow\; \frac{-\,D\,q\,\sin(q\,L)}{D\,\cos(q\,L)} \,=\, \frac{-\,B\,\kappa\,e^{-\kappa\,L}}{B\,e^{-\kappa\,L}}\,. \tag{4.55}$$

Eq. (4.55) can then easily be rewritten as

$$q\,\tan(q\,L) \,=\, \kappa\,. \tag{4.56}$$

---

[1] It does not matter here whether one chooses the boundary between regions I and II or II and III, the result is the same.

Since κ and q depend on the energy E via

$$\kappa \,=\, \frac{1}{\hbar}\sqrt{-\,2\,m\,E} \quad , \quad q \,=\, \frac{1}{\hbar}\sqrt{2\,m(E\,-\,V_0)} \tag{4.57}$$

Eq. (4.56) becomes a formula for the permitted energies, a *quantization condition*.

The analogous calculation as in Eq. (4.55) can be done for the odd solutions, which then gives the corresponding condition for the odd states

$$q\,\cot(q\,L) \,=\, -\,\kappa\,. \tag{4.58}$$

Graphical solution: The two equations (4.56) and (4.58) are so called *transcendental equations*, which means that they can only be written in implicit form. Solutions to this group of equations can be found numerically or graphically, the latter of which we will do here, but not analytically. However before we do so, we will introduce new (dimensionless) variables z, z_0 to simplify the calculation

$$z \,:=\, q\,L \quad , \quad z_0 \,:=\, \frac{L}{\hbar}\sqrt{2\,m\,V_0}\,. \tag{4.59}$$

To relate our old variables κ and q to the new ones, we first look at

$$\kappa^2 \,+\, q^2 \,=\, \frac{-2\,m\,E}{\hbar^2} \,+\, \frac{2\,m\,(E\,+\,V_0)}{\hbar^2} \,=\, \frac{2\,m\,V_0}{\hbar^2}\,, \tag{4.60}$$

which we multiply with L^2 to get

$$\Rightarrow \quad \kappa^2\,L^2 \,+\, \underbrace{q^2\,L^2}_{z^2} \,=\, z_0^2 \quad \Rightarrow \quad \frac{\kappa^2\,L^2}{q^2\,L^2} \,=\, \frac{z_0^2\,-\,z^2}{z^2} \,=\, \left(\frac{z_0}{z}\right)^2 \,-\, 1$$

$$\Rightarrow \quad \frac{\kappa}{q} \,=\, \sqrt{\left(\frac{z_0}{z}\right)^2 \,-\, 1} \quad \Rightarrow \quad \text{insert in Eq. (4.56)} \tag{4.61}$$

$$\tan z \,=\, \sqrt{\left(\frac{z_0}{z}\right)^2 \,-\, 1}\,. \tag{4.62}$$

We can now study this graphically by plotting both the left hand and the right hand function for given values of z_0, e.g. for L, m and V_0 chosen such that[2], see Fig. 4.3,

$$z_0 \,=\, \frac{L}{\hbar}\sqrt{2\,m\,V_0} \,=\, 8\,. \tag{4.63}$$

---

[2] Keep in mind that z_0 is dimensionless.

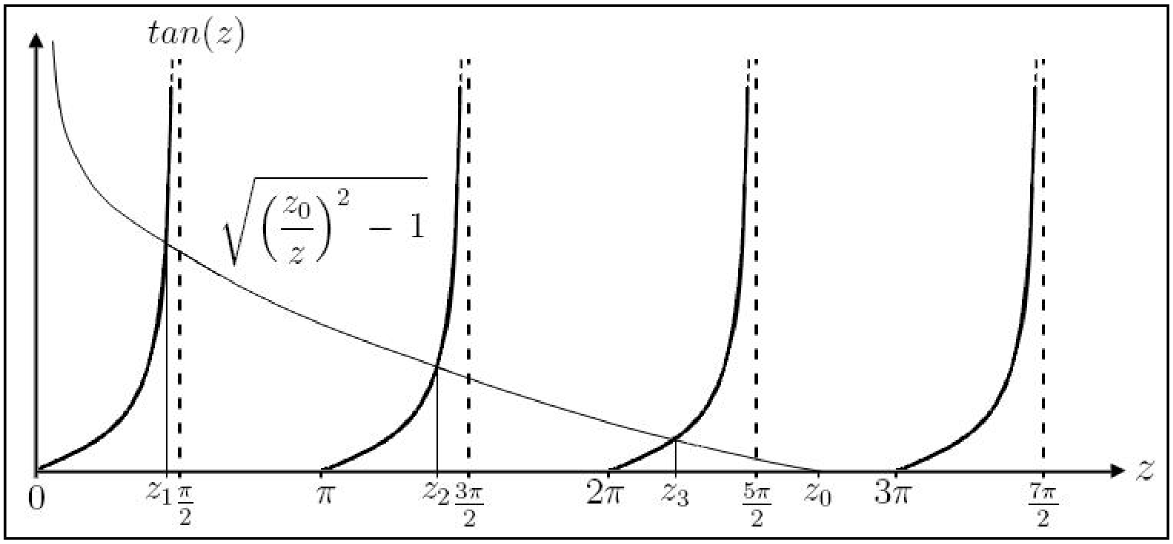

Figure 4.3: Graphic solution of transcendental equation (4.62): The equation cannot be solved analytically, thus both sides are plotted for the given parameter $z_0 = 8$. The intersections then lead to the allowed values of z.

For the chosen parameter we have three solutions $z_1 = 0,8 \times \frac{\pi}{2}$, $z_2 = 2,6 \times \frac{\pi}{2}$ and $z_3 = 4,25 \times \frac{\pi}{2}$ such that $E = E_{\mathrm{n}},\ n = 1,2,3$, from

$$z_{\mathrm{n}} = q_{\mathrm{n}} L = \frac{1}{\hbar} \sqrt{2m(E_{\mathrm{n}} + V_0)} . \tag{4.64}$$

For increasing parameters L and V_0 the value of z_0 grows as well and we obtain more bound states. We could now repeat the same procedure for the odd states by replacing $\tan z$ with $-\cot z$ (recall Eq. (4.58)), which is the same function but shifted by $\frac{\pi}{2}$, so we will skip this calculation and use qualitative statements about the energies of the odd solutions later on.

Special cases: Let us now study some limits of the graphical solutions, where we can find analytical approximations to our problem.

Case I: Broad & deep potential well: For big values for L and V_0 the quantity z_0 also reaches higher values and the intersections with the tangens-curves get even closer to the singularities $(2k+1)\frac{\pi}{2}$, where k is natural number

$$z^2_{\mathrm{2k+1}} \approx (2k + 1)\frac{\pi}{2} , \tag{4.65}$$

which turns into

$$z^2_{\mathrm{2k+1}} = (q_{\mathrm{2k+1}} L)^2 \quad \longrightarrow \quad \frac{L^2}{\hbar^2} 2m(E_{\mathrm{2k+1}} + V_0) \approx \frac{(2k + 1)^2 \pi^2}{4} . \tag{4.66}$$

Thus we can express the corresponding energies as

$$E_{\mathrm{2k+1}} + V_0 \approx \frac{(2k + 1)^2 \pi^2 \hbar^2}{2m(2L)^2} . \tag{4.67}$$

Since we already concluded, that the feature of the odd solutions is to shift the tangens by $\frac{\pi}{2}$, we can easily deduce that in the limit of very large z_0 the intersections will get close to the singularities of the cotangens, which are given by integer multiples of π

$$\psi^{(-)} : \qquad E_{2\mathrm{k}} + V_0 \approx \frac{(2k)^2 \, \pi^2 \, \hbar^2}{2 \, m \, (2L)^2} \, . \tag{4.68}$$

We can then summarize all solutions to the formula

$$E_{\mathrm{n}} + V_0 \approx \frac{(n)^2 \, \pi^2 \, \hbar^2}{2 \, m \, (2L)^2} \, , \tag{4.69}$$

which is the expression for the infinite potential well (see Eq. (4.37)), if we keep in mind that the energy scales we used differ by V_0 and the width of the potential well is here $2L$ instead of L.

Case II: Narrow & small potential well: We now consider small values for L and V_0, which reduces z_0 as well and we finally push its value under $\frac{\pi}{2}$, see Fig. 4.4.

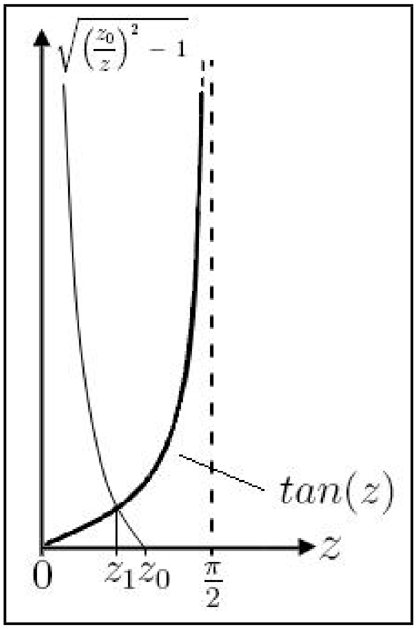

Figure 4.4: Graphic solution of small potential well: However small the parameter z_0 gets, an intersection of the functions always remains, thus always allowing at least one bound state.

Remark: The finite potential well in one dimension always has a bound state, but this is not generally the case for all dimensions. The 3–dimensional Schrödinger equation does not automatically allow for a bound state in such a problem. Though the equation can be rewritten in terms of the radius r, and thus get reduced to a 1–dimensional problem, the radius however is defined on $\mathbb{R}^+$ in contrast to $x \in \mathbb{R}$. The ground state then is an odd wave function (see Fig. 4.5) which is the equivalent of the first excited state in our problem, and therefore can get "pushed" out if the parameters are too small.

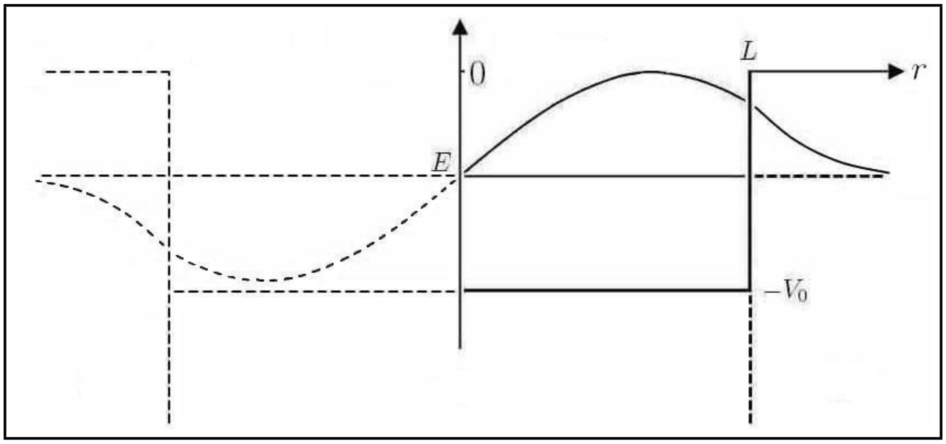

Figure 4.5: Potential well in three dimensions: The 3–dimensional problem can be reduced to a 1–dimensional problem in r, but the reduced wave function of the ground state then is an odd function, because r is only defined on $\mathbb{R}^+$.

Let us now return again to the even and odd wavefunctions, Eq. (4.53), where we still have to determine the constants B, C and D. We first use the continuity at $x = L$ to get the equations

$$\psi^{(+)}: \quad D\cos(q_n L) = B\,e^{-\kappa_n L} \quad , \quad \psi^{(-)}: \quad C\sin(q_n L) = B\,e^{-\kappa_n L}\,, \tag{4.70}$$

providing us with

$$D = B\,\frac{e^{-\kappa_n L}}{\cos(q_n L)} \quad , \quad C = B\,\frac{e^{-\kappa_n L}}{\sin(q_n L)}\,. \tag{4.71}$$

Finally we obtain the constant B from the normalization

$$B_n = \frac{\cos(q_n L)}{\sqrt{1 + \kappa_n L}}\,e^{-\kappa_n L} \quad \Rightarrow \quad C_n = D_n = \frac{1}{\sqrt{1 + \kappa_n L}}\,. \tag{4.72}$$

Physical interpretation of the finite potential well: An application for the finite potential well is the model for free electrons in metal, used in solid state physics. There the atoms of the metal crystal "share" the electrons which are thus free to move inside the metal, but face a potential barrier, which keeps them inside. Thus in a first approximation, the finite (square) potential well is a good model for the situation, see Fig. 4.6.

To release one electron from the metal, the energy W must be invested. This is the work function, we encountered in Section 1.2.3, which we can calculate[3] with the formula

$$W = V_0 - E_n\,, \tag{4.73}$$

---

[3] Keep in mind, that we rescaled the energy here in contrast to our previous calculations, the potential V_0 as well as the energies of the bound states are positive here.

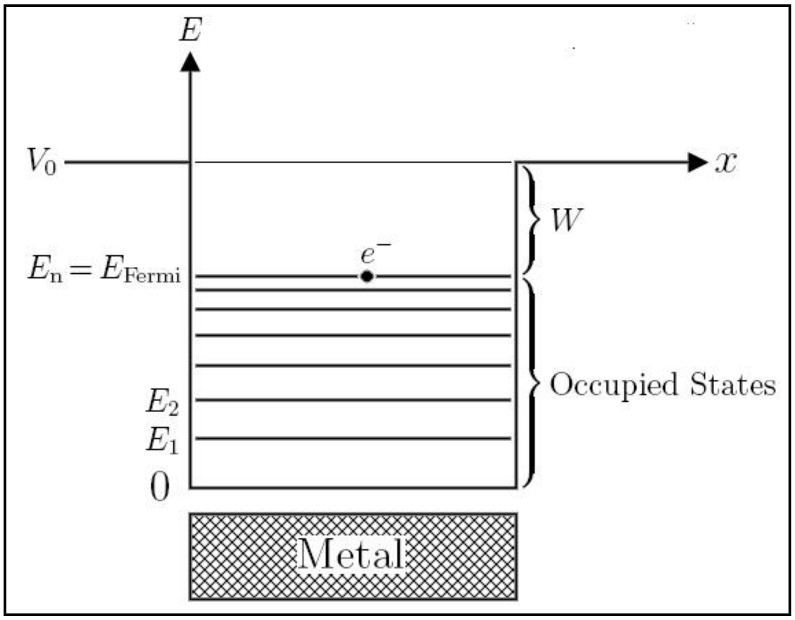

Figure 4.6: Potential model for electrons in metal: The electrons are free inside the metal, occupying the states up to a certain energy level E_n. In order to separate an electron from the metal, one needs to overcome the energy difference W, the work function.

where E_n is the highest[4] occupied energy level.

4.5.2 Scattered States

We will now investigate the Schrödinger equation, including the potential from Eq. (4.40) for positive energies $E > 0$. We will assume that an initial plane wave travels from $x = -\infty$ to our potential and study the possible states that are not bound but scattered, i.e. transmitted or reflected by the potential, see Fig. 4.7. As before we will analyze our problem for the regions I, II and III separately before combining our results.

Region I: $x < -L\,, \quad V(x) = 0$

In this region we have to solve the free Schrödinger equation

$$\frac{d^2\psi}{dx^2} = -\frac{2\,m\,E}{\hbar^2}\,\psi = -\,k^2\,\psi \quad \text{where} \quad k^2 = \frac{2\,m\,E}{\hbar^2}\,. \tag{4.74}$$

---

[4] The electrons in the metal are *fermions*, particles with half integer spin / antisymmetric wave function. Therefore they are subject to the *Pauli Exclusion Principle*, stating that no more than one of them can occupy the same quantum state (see Section 7.3). Thus at temperature $T = 0$ the states of the potential well are filled one by one from the bottom of the well to a certain energy level, named the *Fermi Energy*, creating the so called Fermi Sea.

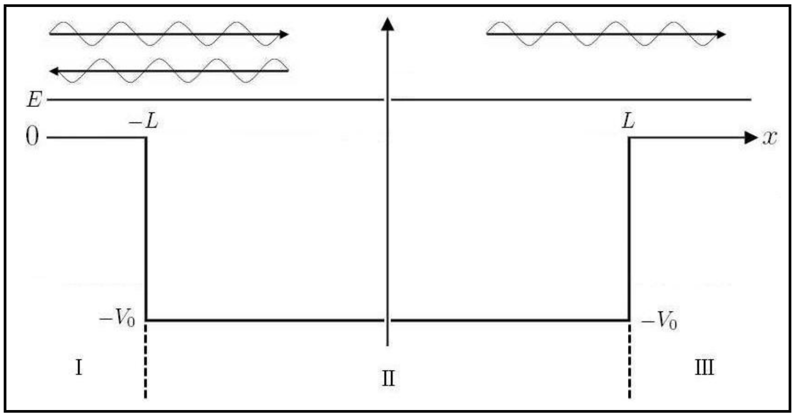

Figure 4.7: Scattered states of the finite potential well: The quantum mechanical potential well can either reflect or transmit an incoming state with energy $E > 0$, whereas classically only the transmission is possible.

Like in Eq. (4.46) we can write down the general solution to this equation as

$$\psi(x) \,=\, \underbrace{A\,e^{i\,k\,x}}_{\text{incoming}} \,+\, \underbrace{B\,e^{-i\,k\,x}}_{\text{reflected}} \quad \text{with} \quad k \,=\, \frac{1}{\hbar}\sqrt{2\,m\,E} \,>\, 0\,. \tag{4.75}$$

Region II: $\quad -L \,\leq\, x \,\leq\, L\,, \quad V(x) \,=\, -V_0$

In total analogy to the case for bound states, Eq. (4.45), we get in region II

$$\psi(x) \,=\, C\,\sin(qx) \,+\, D\,\cos(qx) \qquad q \,=\, \tfrac{1}{\hbar}\sqrt{2m(E \,+\, V_0)} \,>\, 0\,. \tag{4.76}$$

Region III: $\quad x \,>\, L\,, \quad V(x) \,=\, 0$

Here we have exactly the situation as in region I, and can thus easily write down the solution as in Eq. (4.75) with new constants F and G, but we set $G \equiv 0$ since we assume no reflection at infinity (the associated solution would represent a wave traveling from right to left).

$$\psi(x) \,=\, \underbrace{F\,e^{i\,k\,x}}_{\text{transmitted}}\,. \tag{4.77}$$

Summarizing our solutions we have

$$\psi(x) \,=\, \begin{cases} A\,e^{i\,k\,x} \,+\, B\,e^{-i\,k\,x} & \text{I} \\ C\,\sin(qx) \,+\, D\,\cos(qx) & \text{II} \\ F\,e^{i\,k\,x} & \text{III} \end{cases} \tag{4.78}$$

where the constants A, B and F now characterize the incoming, reflected and transmitted parts of our solution respectively. We then regard the boundary conditions, i.e. the continuity of the wave function and its first derivative at the edges of the potential wall

$$\psi(x) \xrightarrow[x \to -L]{} \quad A\,e^{-ikL} + B\,e^{ikL} \quad = -\,C\,\sin(qL) + D\,\cos(qL) \tag{4.79}$$

$$\psi'(x) \xrightarrow[x \to -L]{} \quad ik\,\left(A\,e^{-ikL} - B\,e^{ikL}\right) \quad = q\,\left(C\,\cos(qL) + D\,\sin(qL)\right) \tag{4.80}$$

$$\psi(x) \xrightarrow[x \to +L]{} \quad C\,\sin(qL) + D\,\cos(qL) \quad = F\,e^{ikL} \tag{4.81}$$

$$\psi'(x) \xrightarrow[x \to +L]{} \quad q\,\left(C\,\cos(qL) - D\,\sin(qL)\right) \quad = ik\,F\,e^{ikL}\ . \tag{4.82}$$

Together with the normalization condition we thus have 5 equations for our 5 variables A, B, C, D and F. To solve this system of equations we start by calculating

$$\text{Eq. (4.81)} \cdot \sin(\text{qL}) \quad + \quad \text{Eq. (4.82)} \cdot \frac{1}{\text{q}}\cos(\text{qL}) \quad \Rightarrow$$

$$\underbrace{\left(\sin^2(qL) + \cos^2(qL)\right)}_{1}\,C \;=\; C = \left(\sin(qL) + i\,\frac{k}{q}\,\cos(qL)\right)\,F\,e^{ikL} \tag{4.83}$$

$$\text{Eq. (4.81)} \cdot \cos(\text{qL}) \quad + \quad \text{Eq. (4.82)} \cdot \frac{1}{\text{q}}\sin(\text{qL}) \quad \Rightarrow$$

$$\underbrace{\left(\cos^2(qL) + \sin^2(qL)\right)}_{1}\,D \;=\; D = \left(\cos(qL) - i\,\frac{k}{q}\,\sin(qL)\right)\,F\,e^{ikL}\ . \tag{4.84}$$

Then we insert the results for C and D into Eq. (4.79) and Eq. (4.80) to get A and B in dependence of F

$$\begin{aligned} A\,e^{-ikL} + B\,e^{ikL} \;&=\; -\,\text{Eq. (4.83)} \cdot \sin(\text{qL}) + \text{Eq. (4.84)} \cdot \cos(\text{qL}) \\ &=\; \left(-\sin^2(qL) - 2\,i\,\frac{k}{q}\,\sin(qL)\,\cos(qL) + \cos^2(qL)\right)\,F\,e^{ikL}\,. \end{aligned} \tag{4.85}$$

Using the following identities we can then rewrite Eq. (4.85)

$$\cos^2 x - \sin^2 x = \cos(2x) \quad , \quad 2\,\sin x\,\cos x = \sin(2x) \tag{4.86}$$

$$\Rightarrow \quad A\,e^{-2ikL} + B = \left(\cos(2qL) - i\,\frac{k}{q}\,\sin(2qL)\right)\,F\ , \tag{4.87}$$

and applying the same procedure for Eq. (4.80) gives

$$\begin{aligned} i\,k\,\left(A\,e^{-ikL} - B\,e^{ikL}\right) &= q\,\left(\text{Eq. (4.83)} \cdot \cos(\text{qL}) + \text{Eq. (4.84)} \cdot \sin(\text{qL})\right) \\ &= \left(i\,\frac{k}{q}\,\cos^2(qL) + 2\,\sin(qL)\,\cos(qL) - i\,\frac{k}{q}\,\sin^2(qL) +\right)\,F\,e^{ikL} \end{aligned} \tag{4.88}$$

$$\overset{\text{Eq. (4.86)}}{\Rightarrow} \quad A\,e^{-2ikL} - B = -\,i\,\frac{q}{k}\,\left(\sin(2qL) + i\,\frac{k}{q}\,\cos(2qL)\right)\,F\ . \tag{4.89}$$

At last we subtract Eq. (4.88) from Eq. (4.87) to get the coefficient B, which characterizes the reflection from the potential well

$$\text{Eq. (4.87)} \; - \; \text{Eq. (4.88)} \; = \; 2\,\mathrm{B} \; = \; \mathrm{i} \left(\frac{\mathrm{q}}{\mathrm{k}} \, - \, \frac{\mathrm{k}}{\mathrm{q}} \right) \sin(2\mathrm{qL})\,\mathrm{F} \tag{4.90}$$

$$\Rightarrow \quad \frac{B}{F} \; = \; i\,\frac{q^2 \, - \, k^2}{2kq} \; \sin(2qL) \quad \text{reflexion/transmission amplitude} \; . \tag{4.91}$$

This can be understood as a balance between the reflected and the transmitted part of the wave function. To get the probability for the reflection or transmission we have to normalize each part by the amplitude of the incoming wave and to take the modulus squared of each expression. We also want to express the quantities q and k by the more familiar constants m, $\hbar$ and V_0, using Eq. (4.75) and Eq. (4.76)

$$(q^2 \, - \, k^2)^2 \; = \; \frac{1}{\hbar^4}\,(2m\,(E \, + \, V_0) \, - \, 2mE)^2 \; = \; \frac{(2m)^2 V_0^2}{\hbar^4} \tag{4.92}$$

$$4\,q^2\,k^2 \; = \; 4\,\frac{1}{\hbar^4}\,(2m)^2\,E\,(E \, + \, V_0) \; . \tag{4.93}$$

Thus we find for the ***reflection coefficient*** $R(E)$ describing the probability of reflection

$$R(E) \; = \; \frac{|B|^2}{|A|^2} \; = \; \frac{V_0^2}{4\,E\,(E \, + \, V_0)} \; \sin^2(2qL)\,\frac{|F|^2}{|A|^2} \; \neq \; 0 \; . \tag{4.94}$$

We conclude that there is a nonzero possibility for a reflection of the wave at the potential well, a purely quantum mechanical effect that is not possible classically.

The reflection, of course, is related to the transmission of the wave. Focussing on the transmission, we can calculate the *transmission coefficient* $T(A) = \frac{|F|^2}{|A|^2}$ which rates the transmitted against the incoming intensity. We first calculate the sum of Eq. (4.87) and Eq. (4.88) to get the *transmission amplitude*

$$\text{Eq. (4.87)} \, + \, \text{Eq. (4.88)} \; = \; A\,e^{-2\,i\,kL} \; = \; 2\left(\cos(2qL) \, - \, i\,\sin(2qL)\,\tfrac{1}{2}\left(\frac{q}{k} \, + \, \frac{k}{q}\right)\right) F$$

$$\Rightarrow \quad \frac{F}{A} \; = \; e^{-2\,i\,kL}\left(\cos(2qL) \, - \, i\,\frac{q^2 \, + \, k^2}{2kq}\,\sin(2qL)\right)^{-1} , \tag{4.95}$$

and by taking the modulus squared and inserting the expressions for q and k (Eq. (4.75) and Eq. (4.76)) we obtain the ***transmission coefficient*** $T(E)$

$$T(E) \,=\, \frac{|F|^2}{|A|^2} \,=\, \left(1 \,+\, \frac{V_0^2}{4\,E\,(E\,+\,V_0)}\,\sin^2(2qL) \right)^{-1} \,. \tag{4.96}$$

Of course, both coefficients are related by

$$R(E) \,+\, T(E) \,=\, 1\,. \tag{4.97}$$

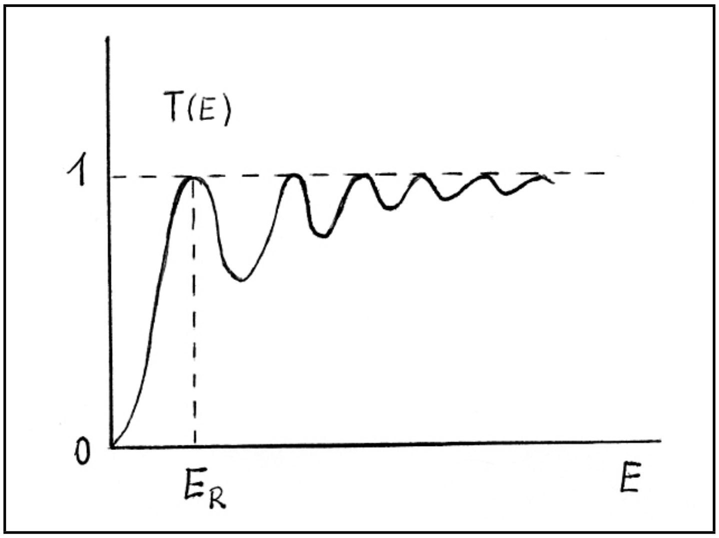

Figure 4.8: Transmission coefficient: The transparency of the well, Eq. (4.96), is plotted as a function of the energy showing the positions E_R of the resonances.

Studying the transmission coefficient we easily see that for some arguments the sinus function will be zero and thus the probability for transmission equal to one, $T\,=\,1$, see Fig. 4.8. To find the corresponding energies we simply look for the roots of the sinus, finding the following condition

$$\sin(2qL) \,=\, 0 \quad \Rightarrow \quad \frac{2L}{\hbar}\,\sqrt{2\,m\,(E\,+\,V_0)} \,=\, n\pi \;, \tag{4.98}$$

where $n\,\in\,\mathbb{N}$ labels the energies. We thus get the positions of the ***resonances*** of a finite potential well, the energies for which it becomes transparent, as

$$E_{\mathrm{n}} \,=\, \frac{n^2\pi^2\hbar}{2m\,(2L)^2} \,-\, V_0\;. \tag{4.99}$$

These resonance positions, interestingly, happen to be precisely at the allowed energy levels of the infinite potential well (see Eq. (4.37) and the remarks accompanying Eq. (4.69)). Accordingly, the minima of transmission can found by using the condition

$$\sin(2qL) \,=\, \pm\,1 \quad \Rightarrow \quad \frac{2L}{\hbar}\,\sqrt{2\,m\,(E\,+\,V_0)} \,=\, (2n\,+\,1)\,\tfrac{\pi}{2}\;. \tag{4.100}$$

We can also conclude, that the resonances become more distinct the deeper the potential.

At last let us study the transmission coefficient in the vicinity of the resonant energies. To this end we go back to the transmission amplitude Eq. (4.95), where we expand the denominator in a Taylor series around the resonant energies, calculating the first two terms of the expansion individually

$$\left(\cos(2qL) \,-\, i\,\frac{q^2 \,+\, k^2}{2kq}\,\sin(2qL)\right)_{E=E_\mathrm{R}} \,=\, \cos(n\pi) \,=\, \pm 1 \,. \tag{4.101}$$

We can easily see that, obeying the condition of Eq. (4.98), all terms containing $\sin(2qL)$ evaluated at $E = E_\mathrm{R}$ vanish, while the cosinus of this argument gives ± 1. Also the second order can then be calculated easily

$$\begin{aligned} \frac{d}{dE}\left(\cos(2qL) \,-\, i\,\frac{q^2 \,+\, k^2}{2kq}\,\sin(2qL)\right)_{E=E_\mathrm{R}} &= \\ = \left(-\,i\,\frac{q^2 \,+\, k^2}{2kq}\,\frac{d(2qL)}{dE}\right)_{E=E_\mathrm{R}} \underbrace{\cos(n\pi)}_{\pm 1} &= \\ = \mp i\left(\frac{q^2 \,+\, k^2}{2kq}\,\frac{d(2qL)}{dE}\right)_{E=E_\mathrm{R}} &=: \,\mp i\,\frac{2}{\Gamma}\,. \end{aligned} \tag{4.102}$$

Neglecting all higher terms the expansion of the denominator then is of the form

$$\pm 1 \mp i\,\frac{2}{\Gamma}\,(E \,-\, E_\mathrm{R}) \,+\, \cdots \,\approx\, \frac{2}{\Gamma}\left(\pm\frac{\Gamma}{2} \mp i\,(E \,-\, E_\mathrm{R})\right)\,. \tag{4.103}$$

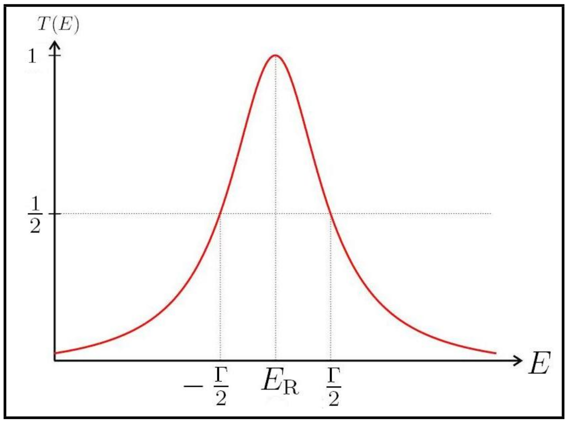

Figure 4.9: Breit–Wigner distribution: The distribution $T(E)$ has height 1 and width Γ at half maximum; it is also known as Cauchy distribution in stochastic physics and as Breit–Wigner- or Lorentz distribution in particle physics; figure from http://de.wikipedia.org/w/index.php?title=Bild:Breitwig_fig.png&filetimestamp=20050622183322

Reinserting expression (4.103) as the denominator of Eq. (4.95) and taking the modulus squared to get the corresponding transmission coefficient we arrive at the expression valid near the resonances

$$T(E) \,=\, \frac{\frac{\Gamma^2}{4}}{(E \,-\, E_{\mathrm{R}})^2 \,+\, \frac{\Gamma^2}{4}} \;. \tag{4.104}$$

This is the well-known ***Breit–Wigner distribution***, also known as Lorentz- or Cauchy distribution, which describes resonance phenomena[5]. The quantity Γ represent the width at half maximum of the distribution and is proportional to the inverse lifetime of the resonance, $\Gamma = \tau^{-1}$, see Fig. 4.9.

As a simple explanation of the resonances of the finite potential well, we may regard them as destructive interference between the waves reflected at $x = L$ and $x = -L$, sketched in Fig. 4.10.

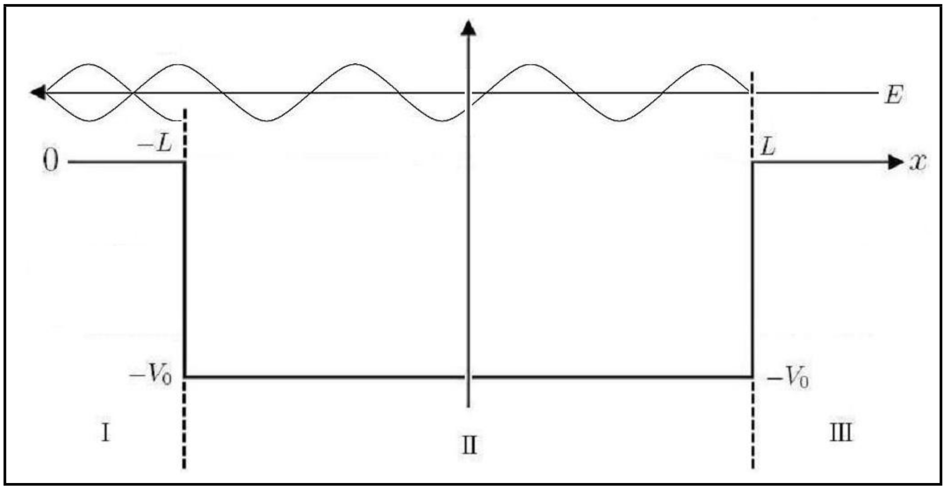

Figure 4.10: Interference of reflected waves: The resonances of the finite potential well can be regarded as interference of the reflected waves, thus permitting free transmission.

Experiment: Ramsauer–Townsend effect

The resonance of transmission can be nicely seen in the scattering of slow electrons in a noble gas (e.g. Ne, Ar, Xe) which has been studied independently by C. Ramsauer and J.S. Townsend in the 1920's. The probability for the electrons to collide with the gas particles, which classically should decrease monotonically for increasing energy, is observed to reach local minima for certain energies. This effect is in total agreement with the quantum mechanical prediction of the transparency of the potential well for the resonance energies.

---

[5] E.g. the forced resonance of a driven oscillator or energy resonances in particle physics describing unstable particles, where the lifetime is proportional to the inverse of the width.

4.6 Tunnel Effect

4.6.1 Finite Potential Barrier

The so-called tunnel effect of quantum mechanics can be derived from a special case of the potential well, by changing $-V_0$ into $+V_0$, thus creating a potential barrier, as seen in Fig. 4.11

$$V(x) = \begin{cases} V_0 & \text{for } |x| \leq L \\ 0 & \text{for } |x| > L \end{cases} \tag{4.105}$$

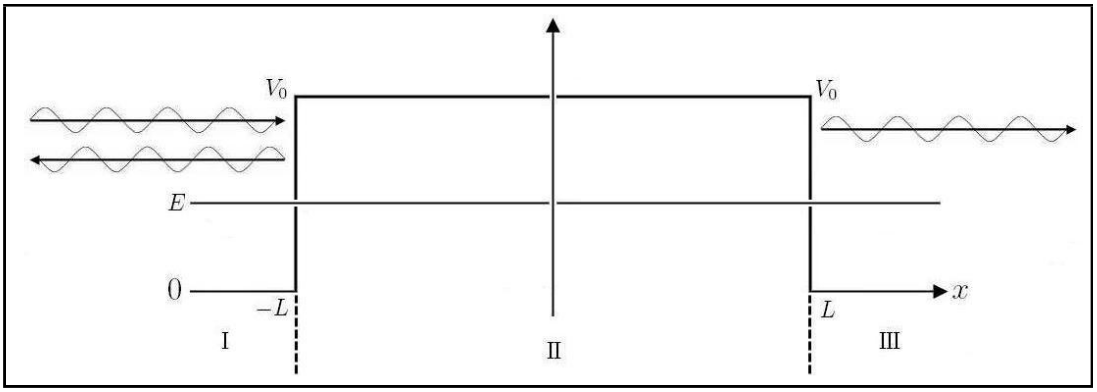

Figure 4.11: Tunnel effect: For a given potential barrier with height V_0 the solutions of the Schrödinger equation with energy $E < V_0$ still have a nonvanishing probability density in region III, which allows them to "tunnel" through the barrier although this would classically be forbidden.

Classically, a particle with less energy than the potential barrier could only be reflected. But in quantum mechanics, due to continuity the wave function decreases exponentially in the forbidden region II, resulting in a nonvanishing probability density in region III. It allows the particle to pass the barrier as if it was through a tunnel, this linguistic illustration gives rise to the name ***tunnel effect***.

Mathematically, we can use the solutions from Section 4.5.2, where we replace the $-V_0$ into $+V_0$ changing the solutions in region II

$$\psi(x) = \begin{cases} A\, e^{i\,k\,x} + B\, e^{-i\,k\,x} & \text{I} \\ C\, e^{-qx} + D\, e^{qx} & \text{II} \\ F\, e^{i\,k\,x} & \text{III} \end{cases} \tag{4.106}$$

where

$$k = \frac{1}{\hbar}\sqrt{2\,m\,E} > 0 \quad , \quad q = \frac{1}{\hbar}\sqrt{2\,m\,(V_0 - E)} > 0\,. \tag{4.107}$$

Since the exponent of the solution in region II is not imaginary anymore, we can replace the sinus and cosinus functions of Section 4.5.2 with their hyperbolic counterparts

$$\sin(2qL) \to \sinh(2qL) \quad , \quad \cos(2qL) \to \cosh(2qL)\,, \tag{4.108}$$

and calculate the missing constants from the continuity requirements as before, which we will however not do explicitly again.

4.6.2 Tunneling – Transmission Coefficient

For the transmission amplitude we get the result

$$\frac{F}{A} = e^{-2\,i\,kL}\left(\cosh(2qL) - i\,\frac{q^2 - k^2}{2kq}\,\sinh(2qL)\right)^{-1}\;, \tag{4.109}$$

which, using the identity

$$\cosh^2(x) - \sinh^2(x) = 1\,, \tag{4.110}$$

gives the transmission coefficient

$$T(E) = \frac{|F|^2}{|A|^2} = \left(1 + \frac{(k^2 + q^2)^2}{4\,k^2\,q^2}\,\sinh^2(2qL)\right)^{-1} \neq 0\;. \tag{4.111}$$

It can be simplified under the condition that $qL \gg 1$, which is a good approximation in most cases. Then

$$qL \gg 1 \quad \Rightarrow \quad \sin(2qL) \approx \tfrac{1}{2}\,e^{2qL}\;, \tag{4.112}$$

and we can rewrite Eq. (4.111) as

$$T \approx (\;1 + \underbrace{\frac{(k^2 + q^2)^2}{4\,k^2\,q^2}\,\frac{e^{4qL}}{4}}_{\gg 1}\;)^{-1} \approx \frac{16\,k^2\,q^2}{(k^2 + q^2)^2}\,e^{-4qL}\;. \tag{4.113}$$

We then use Eq. (4.107) to express the transmission coefficient by the energy and potential strength to get

$$T \approx \frac{16\,E\,(V_0 - E)}{V_0^2}\,e^{-4\,\frac{L}{\hbar}\,\sqrt{2m\,(V_0 - E)}}\,. \tag{4.114}$$

Using $e^x\,e^y = e^{x+y}$ we can write the whole coefficient as an exponential function

$$\Rightarrow \quad T \approx \exp\left(-4\,\frac{L}{\hbar}\,\sqrt{2m\,(V_0 - E)} + \ln\left(\frac{16\,E\,(V_0 - E)}{V_0^2}\right)\right)\;. \tag{4.115}$$

Since we required that qL be much bigger than one and the logarithm increases only very slowly we can conclude that the first term in the exponent outweighs the second one, which we therefore neglect to obtain

$$T \approx e^{-4\,\frac{L}{\hbar}\,\sqrt{2m\,(V_0 - E)}}\,. \tag{4.116}$$

We now have a good approximation for the transmission probability of a single potential step, constant in certain interval and vanishing outside. This potential is of course

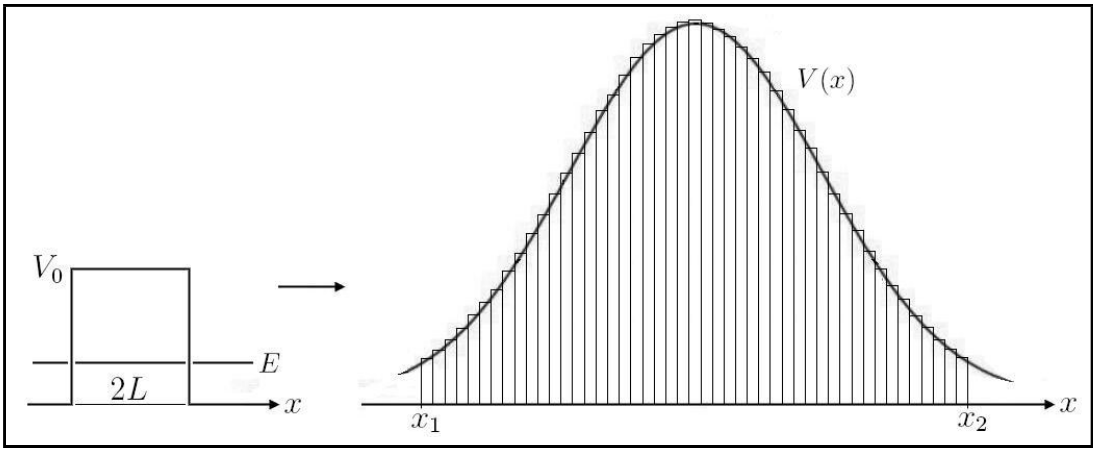

Figure 4.12: Calculation of Gamow factor: The generalization of the transmission coefficient, from a single constant potential barrier with width $2L$, to the incorporation of a function $V(x)$ is done straightforward by integration of infinitesimal potential barriers.

a very crude approximation of real life potentials, which usually are more complicated functions of x. To meet these concerns we can generalize the transmission coefficient of Eq. (4.116) to the so called ***Gamow factor*** by "chopping" a given potential in infinitesimal potential steps with constant values and integrating over a reasonable range $[x_1, x_2]$ for which the potential stays above a certain value, see Fig. 4.12.

$$T \approx e^{-\frac{2}{\hbar}\int\limits_{x_1}^{x_2} dx \sqrt{2m\,(V(x)-E)}} . \tag{4.117}$$

4.6.3 Physical Examples of Tunneling

We will now briefly present some examples, where the tunnel effect explains the observed phenomena.

Tunneling between conductors: Imagine two conducting materials, separated by a thin insulating material, sketched in Fig. 4.13. The tunnel effect then allows the electrons to tunnel through that barrier, thus creating a current. This effect is also observed for superconducting materials, where it is named Josephson effect.

Cold emission: Electrons can be emitted from metals at very low temperatures, even without incident light if an exterior electric field is applied. Assuming that the electrons have a very low energy compared to the potential height, it is not very probable that an electron can tunnel through the potential barrier. Only if an electric field raises the energy of the electron, the transmission coefficient increases and the electron emission can be observed.

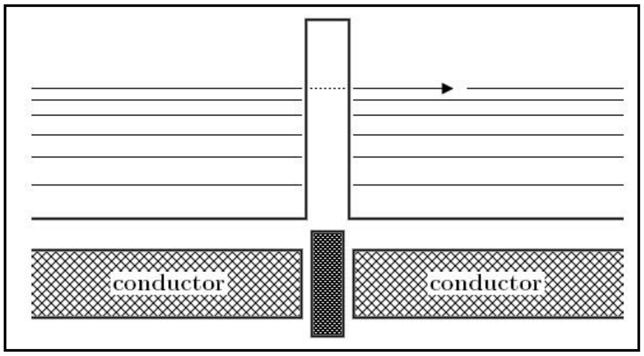

Figure 4.13: Tunneling through insulator: Between two conducting materials, separated by a thin insulating barrier, the tunnel effect creates a current.

Alpha decay: In nuclear physics the tunnel effect plays an important role in the process of alpha decay. In the nucleus two protons and two neutrons can form a alpha-particle, i.e. a ^4He nucleus, which due to the energy gain of the binding process, has the required energy to tunnel through the Coulomb-barrier, see Fig. 4.14. E.g. the alpha decay of polonium to lead

$$^{212}Po \rightarrow \, ^{208}Pb + \alpha \,. \tag{4.118}$$

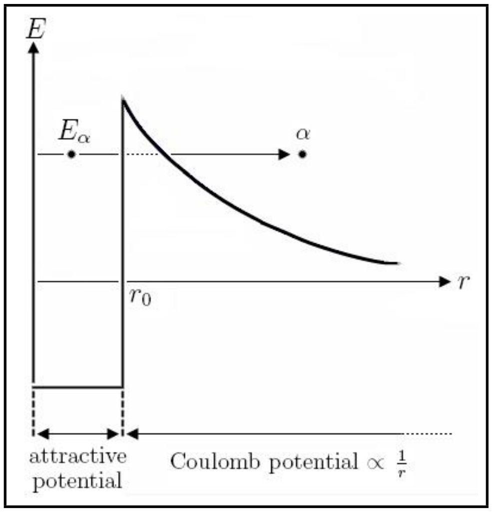

Figure 4.14: Alpha decay: The alpha particle formed in the nucleus has an energy boost due to the mass deficit of the binding process, which allows it to overcome the attractive strong force inside the nucleus and tunnel through the Coulomb barrier.

Chapter 5

Harmonic Oscillator and Coherent States

5.1 Harmonic Oscillator

In this chapter we will study the features of one of the most important potentials in physics, it's the harmonic oscillator potential which is included now in the Hamiltonian

$$V(x) = \frac{m\omega^2}{2}x^2 \,. \tag{5.1}$$

There are two possible ways to solve the corresponding time independent Schrödinger equation, the algebraic method, which will lead us to new important concepts, and the analytic method, which is the straightforward solving of a differential equation.

5.1.1 Algebraic Method

We start again by using the time independent Schrödinger equation, into which we insert the Hamiltonian containing the harmonic oscillator potential (5.1)

$$H\,\psi = \left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{m\omega^2}{2}x^2\right)\psi = E\,\psi \,. \tag{5.2}$$

We rewrite Eq. (5.2) by defining the new operator $\bar{x} := m\omega x$

$$H\,\psi = \frac{1}{2m}\left[\left(\frac{\hbar}{i}\frac{d}{dx}\right)^2 + (m\omega x)^2\right]\psi = \frac{1}{2m}\left[p^2 + \bar{x}^2\right]\psi = E\,\psi \,. \tag{5.3}$$

We will now try to express this equation as the square of some (yet unknown) operator

$$p^2 + \bar{x}^2 \quad \rightarrow \quad (\bar{x} + ip)(\bar{x} - ip) = p^2 + \bar{x}^2 + i(p\bar{x} - \bar{x}p) \,, \tag{5.4}$$

but since x and p do not commute (remember Theorem 2.3), we only will succeed by taking the $x - p$ commutator into account. Eq. (5.4) suggests to factorize our Hamiltonian by defining new operators a and $a^\dagger$ as:

Definition 5.1 $a \; := \; \frac{1}{\sqrt{2m\omega\hbar}}(m\omega x \, + \, ip)$ *annihilation operator*

$a^\dagger \; := \; \frac{1}{\sqrt{2m\omega\hbar}}(m\omega x \, - \, ip)$ *creation operator*

These operators each create/annihilate a quantum of energy $E = \hbar\omega$, a property which gives them their respective names and which we will formalize and prove later on. For now we note that position and momentum operators are expressed by a's and $a^\dagger$'s like

$$x \, = \, \sqrt{\frac{\hbar}{2m\omega}} \, \left(a \, + \, a^\dagger \right) \qquad\qquad p \, = \, -i \sqrt{\frac{m\omega\hbar}{2}} \, \left(a \, - \, a^\dagger \right) \, . \tag{5.5}$$

Let's next calculate the commutator of the creation and annihilation operators. It's quite obvious that they commute with themselves

$$\left[\, a \, , \, a \,\right] \, = \, \left[\, a^\dagger \, , \, a^\dagger \,\right] \, = \, 0 \, . \tag{5.6}$$

To find the commutator of a with $a^\dagger$ we first calculate $aa^\dagger$ as well as $a^\dagger a$

$$a\,a^\dagger \;\; = \;\; \frac{1}{2m\omega\hbar} \left[(m\omega x)^2 \, + \, im\omega \, [\, p \, , \, x \,] \, + \, p^2 \right] \tag{5.7}$$

$$a^\dagger a \;\; = \;\; \frac{1}{2m\omega\hbar} \left[(m\omega x)^2 \, - \, im\omega \, [\, p \, , \, x \,] \, + \, p^2 \right] \, . \tag{5.8}$$

Since we know $[\, p \, , \, x \,] \, = \, -i\hbar$ we easily get

$$\left[\, a \, , \, a^\dagger \,\right] \, = \, 1 \, . \tag{5.9}$$

Considering again the Hamiltonian from Eq. (5.3) we use expressions (5.7) and (5.8) to rewrite it as

$$H \, = \, \frac{1}{2m} \, \left(p^2 \, + \, (m\omega x)^2 \right) \, = \, \frac{\hbar\omega}{2} \, \left(a^\dagger a \, + \, a\,a^\dagger \right) \, . \tag{5.10}$$

Using the commutator (5.9) we can further simplify the Hamiltonian

$$\left[\, a \, , \, a^\dagger \,\right] \, = \, a\,a^\dagger \, - \, a^\dagger a \, = \, 1 \quad \Rightarrow \quad a\,a^\dagger \, = \, a^\dagger a \, + \, 1 \, , \tag{5.11}$$

$$H \, = \, \hbar\omega(a^\dagger a \, + \, \tfrac{1}{2}) \, . \tag{5.12}$$

For the Schrödinger energy eigenvalue equation we then get

$$\hbar\omega(a^\dagger a \, + \, \tfrac{1}{2}) \, \psi \, = \, E \, \psi \, , \tag{5.13}$$

which we rewrite as an eigenvalue equation for the operator $a^\dagger a$

$$a^\dagger a\,\psi \;=\; \left(\frac{E}{\hbar\omega} \,-\, \frac{1}{2}\right)\,\psi\;, \tag{5.14}$$

having the following interpretation:

Definition 5.2 $N \;:=\; a^\dagger a$ *occupation (or particle) number operator*

and which satisfies the commutation relations

$$\left[\,N\,,\,a^\dagger\,\right] \;=\; a^\dagger \qquad\qquad \left[\,N\,,\,a\,\right] \;=\; -a\;. \tag{5.15}$$

Next we are looking for the eigenvalues ν and eigenfunctions ψ_ν of the occupation number operator N, i.e. we are seeking the solutions of equation

$$N\,\psi_\nu \;=\; \nu\,\psi_\nu\;. \tag{5.16}$$

To proceed we form the scalar product with ψ_ν on both sides of Eq. (5.16), use the positive definiteness of the scalar product (Eq. (2.32)) and the definition of the adjoint operator (Definition 2.5)

$$\nu\,\underbrace{\langle\,\psi_\nu\,|\,\psi_\nu\,\rangle}_{\neq\,0} \;=\; \langle\,\psi_\nu\,|\,N\,\psi_\nu\,\rangle \;=\; \left\langle\,\psi_\nu\,\middle|\,a^\dagger a\,\psi_\nu\,\right\rangle \;=\; \langle\,a\,\psi_\nu\,|\,a\,\psi_\nu\,\rangle \;\geq\; 0\;. \tag{5.17}$$

The above inequality vanishes only if the corresponding vector $a\,\psi$ is equal to zero. Since we have for the eigenvalues $\nu \geq 0$, the lowest possible eigenstate ψ_0 corresponds to the eigenvalue $\nu \;=\; 0$

$$\nu \;=\; 0 \qquad \rightarrow \qquad a\,\psi_0 \;=\; 0\;. \tag{5.18}$$

Inserting the definition of the annihilation operator (Definition 5.1) into condition (5.18), i.e. that the ground state is annihilated by the operator a, yields a differential equation for the ground state of the harmonic oscillator

$$\begin{aligned} a\,\psi_0 \;=\; \frac{1}{\sqrt{2m\omega\hbar}}(m\omega x \,+\, i\,\frac{\hbar}{i}\,\frac{d}{dx})\,\psi_0 \;&=\; 0 \\ \Rightarrow \quad \left(\frac{m\omega}{\hbar}x \,+\, \frac{d}{dx}\right)\,\psi_0 \;&=\; 0\;. \end{aligned} \tag{5.19}$$

We can solve this equation by separation of variables

$$\int\frac{d\psi_0}{\psi_0} \;=\; -\int dx\,\frac{m\omega}{\hbar}\,x \quad \Rightarrow \quad \ln\psi_0 \;=\; -\,\frac{m\omega}{2\hbar}\,x^2 \,+\, \ln\mathcal{N}\;, \tag{5.20}$$

where we have written the integration constant as $\ln\mathcal{N}$, which we will fix by the normalization condition

$$\ln\left(\frac{\psi_0}{\mathcal{N}}\right) \;=\; -\,\frac{m\omega}{2\hbar}\,x^2 \quad\Rightarrow\quad \psi_0(x) \;=\; \mathcal{N}\,\exp\left(-\tfrac{m\omega}{2\hbar}x^2\right)\;. \tag{5.21}$$

We see that the ground state of the harmonic oscillator is a Gaussian distribution. The normalization $\int\limits_{-\infty}^{\infty} dx\,|\psi_0(x)|^2 \;=\; 1$ together with formula (2.119) for Gaussian functions determines the normalization constant

$$\mathcal{N}^2 \;=\; \sqrt{\frac{m\omega}{\pi\hbar}} \quad\Rightarrow\quad \mathcal{N} \;=\; \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}}\;. \tag{5.22}$$

We will now give a description of the whole set of eigenfunctions ψ_ν of the operator N based on the action of the creation operator by using the following lemma:

Lemma 5.1 *If ψ_ν is an eigenfunction of N with eigenvalue ν, then $a^\dagger\psi_\nu$ also is an eigenfunction of N with eigenvalue $(\nu+1)$.*

Proof:

$$\begin{aligned} N\,a^\dagger\psi_\nu \;&\overset{\text{Eq. (5.15)}}{=}\; (a^\dagger\,N \,+\, a^\dagger)\,\psi_\nu \;=\; a^\dagger\,(N\,+\,1)\,\psi_\nu \\ &=\; a^\dagger\,(\nu\,+\,1)\,\psi_\nu \;=\; (\nu\,+\,1)\,a^\dagger\,\psi_\nu\;. \qquad \text{q.e.d.} \end{aligned} \tag{5.23}$$

The state $a^\dagger\psi_\nu$ is not yet normalized, which means it is only proportional to $\psi_{\nu+1}$, to find the proportionality constant we use once more the normalization condition

$$\left\langle\, a^\dagger\psi_\nu \,\middle|\, a^\dagger\psi_\nu \,\right\rangle \;=\; \langle\,\psi_\nu\,|\,\underbrace{a\,a^\dagger}_{a^\dagger a\,+\,1}\,\psi_\nu\,\rangle \;=\; \langle\,\psi_\nu\,|\,(N\,+\,1)\,|\,\psi_\nu\,\rangle \;=\; (\nu\,+\,1)\,\underbrace{\langle\,\psi_\nu\,|\,\psi_\nu\,\rangle}_{1} \tag{5.24}$$

$$\Rightarrow\qquad a^\dagger\psi_\nu \;=\; \sqrt{\nu\,+\,1}\;\psi_{\nu+1}\;. \tag{5.25}$$

This means that we can get any excited state ψ_ν of the harmonic oscillator by successively applying creation operators.

In total analogy to Lemma 5.1 we can formulate the following lemma:

Lemma 5.2 *If ψ_ν is an eigenfunction of N with eigenvalue ν, then $a\psi_\nu$ also is an eigenfunction of N with eigenvalue $(\nu-1)$.*

Proof:

$$\begin{aligned} N\,a\psi_\nu &\overset{\text{Eq. (5.15)}}{=} (a\,N\,-\,a)\,\psi_\nu \,=\, a\,(N\,-\,1)\,\psi_\nu \\ &= a\,(\nu\,-\,1)\,\psi_\nu \,=\, (\nu\,-\,1)\,a\,\psi_\nu\,. \qquad \text{q.e.d.} \end{aligned} \tag{5.26}$$

We again get the proportionality constant from the normalization

$$\langle\, a\psi_\nu\,|\,a\psi_\nu\,\rangle \,=\, \langle\,\psi_\nu\,|\,\underbrace{a^\dagger a}_{N}\,\psi_\nu\,\rangle \,=\, \langle\,\psi_\nu\,|\,\underbrace{N\,|\psi_\nu\,\rangle}_{\nu\,|\,\psi_\nu\,\rangle} \,=\, \nu\,\underbrace{\langle\,\psi_\nu\,|\,\psi_\nu\,\rangle}_{1} \tag{5.27}$$

$$\Rightarrow \qquad a\psi_\nu \,=\, \sqrt{\nu}\;\psi_{\nu\,-\,1}\,. \tag{5.28}$$

Summary: Eigenvalue equation for the harmonic oscillator

The time independent Schrödinger equation, the energy eigenvalue equation

$$H\,\psi_{\mathrm{n}} \,=\, E_{\mathrm{n}}\,\psi_{\mathrm{n}}\,, \tag{5.29}$$

with the Hamiltonian $H \,=\, \hbar\omega\,(N\,+\,\frac{1}{2})$ and the occupation number operator $N \,=\, a^\dagger a$ provides the *energy eigenvalues*

$$E_{\mathrm{n}} \,=\, \hbar\omega\,(n\,+\,\tfrac{1}{2})\,. \tag{5.30}$$

The corresponding state vectors, the *energy eigenfunctions*, are given by

$$\psi_{\mathrm{n}}(x) \,=\, \frac{1}{\sqrt{n!}}(a^\dagger)^{\mathrm{n}}\,\psi_0(x) \,=\, \frac{1}{\sqrt{n!}}\,\left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}}\,(a^\dagger)^{\mathrm{n}}\,\exp(-\frac{m\omega}{2\hbar}x^2)\,, \tag{5.31}$$

with the *ground state*

$$\psi_0(x) \,=\, \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}}\,\exp(-\frac{m\omega}{2\hbar}x^2)\,. \tag{5.32}$$

The explicit form of the excited state wave functions will be calculated later on but we can for now reveal that they are proportional to a product of the ground state and a family of functions, the so-called *Hermite polynomials* H_{n}.

The wave functions thus form a ladder of alternating even and odd energy states, see Fig. 5.1, which are each separated by a quantum of energy $\hbar\omega$, i.e. equally spaced. The creation and annihilation operators then "climb" or "descend" this energy ladder step by step, which is why they are also called ladder operators.

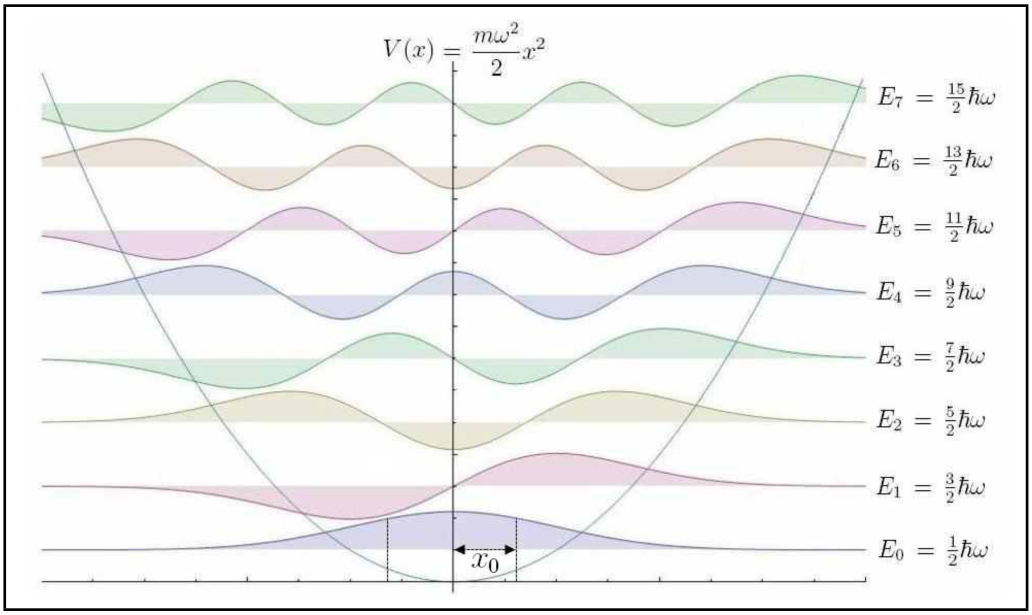

Figure 5.1: Harmonic oscillator: The possible energy states of the harmonic oscillator potential V form a ladder of even and odd wave functions with energy differences of $\hbar\omega$. The ground state is a Gaussian distribution with width $x_0 = \sqrt{\frac{\hbar}{m\omega}}$; picture from http://en.wikipedia.org/wiki/Quantum_mechanical_harmonic_oscillator

5.1.2 Zero Point Energy

We already learned that the lowest possible energy level of the harmonic oscillator is not, as classically expected, zero but $E_0 = \frac{1}{2}\hbar\omega$.

It can be understood in the following way. The ground state is an eigenfunction of the Hamiltonian, containing both kinetic and potential energy contributions, therefore the particle has some kinetic energy in the vicinity of $x = 0$, where the potential energy $V(x \to 0) \to 0$. But this implies according to Heisenberg's uncertainty relation (Eq. (2.80)) that the momentum uncertainty increases pushing the total energy up again until it stabilizes

$$\Delta p \propto \frac{\hbar}{\Delta x} \nearrow \quad \text{for} \quad \Delta x \to 0 \quad \Rightarrow E \neq 0\,. \tag{5.33}$$

We will now illustrate the harmonic oscillator states, especially the ground state and the zero point energy in the light of the uncertainty principle. We start by calculating the position and momentum uncertainties using Definition 2.10

$$\langle\, x \,\rangle \,=\, \langle\, \psi_\mathrm{n} \,|\, x \,|\psi_\mathrm{n} \,\rangle \overset{\text{Eq. (5.5)}}{\propto} \langle\, \psi_\mathrm{n} \,|\, a \,+\, a^\dagger \,|\psi_\mathrm{n} \,\rangle \,\propto\, \underbrace{\langle\, \psi_\mathrm{n} \,|\, \psi_{\mathrm{n}-1} \,\rangle}_{0} + \underbrace{\langle\, \psi_\mathrm{n} \,|\, \psi_{\mathrm{n}+1} \,\rangle}_{0} \,=\, 0 \tag{5.34}$$

$$\langle\, x^2 \,\rangle \,=\, \langle\, \psi_\mathrm{n} \,|\, x^2 \,|\psi_\mathrm{n} \,\rangle \,=\, \frac{\hbar}{2m\omega}\, \langle\, \psi_\mathrm{n} \,|\, a^2 \,+\, aa^\dagger \,+\, a^\dagger a \,+\, (a^\dagger)^2 \,|\, \psi_\mathrm{n} \,\rangle\,. \tag{5.35}$$

Because the the eigenstates ψ_n for different n are orthogonal, the expectation values of a^2 and $(a^\dagger)^2$ vanish identically and we proceed by using Eq. (5.11), where $a^\dagger a = N$.

We get

$$\langle x^2 \rangle = \frac{\hbar}{2m\omega} \, \langle \psi_n \,|\, 2N + 1 \,|\, \psi_n \rangle = \frac{\hbar}{m\omega} (n + \tfrac{1}{2}) = x_0^2 \, (n + \tfrac{1}{2}) , \tag{5.36}$$

where we have introduced a characteristic length of the harmonic oscillator $x_0 = \sqrt{\frac{\hbar}{m\omega}}$ and $n = 0, 1, 2, \cdots$ is a natural number. The position uncertainty is then given by

$$(\Delta x)^2 = \langle x^2 \rangle - \langle x \rangle^2 = x_0^2 \, (n + \tfrac{1}{2}) , \tag{5.37}$$

which vanishes in the classical limit, i.e. $\hbar \to 0$ or $m \to \infty$.

Let us continue with the momentum uncertainty

$$\langle p \rangle = \langle \psi_n \,|\, p \,|\, \psi_n \rangle \overset{\text{Eq. (5.5)}}{\propto} \langle \psi_n \,|\, a - a^\dagger \,|\, \psi_n \rangle \propto \underbrace{\langle \psi_n \,|\, \psi_{n-1} \rangle}_{0} - \underbrace{\langle \psi_n \,|\, \psi_{n+1} \rangle}_{0} = 0 \tag{5.38}$$

$$\begin{aligned} \langle p^2 \rangle &= \langle \psi_n \,|\, p^2 \,|\, \psi_n \rangle = -\frac{m\omega\hbar}{2} \, \langle \psi_n \,|\, a^2 - \underbrace{a a^\dagger - a^\dagger a}_{-(2N+1)} + (a^\dagger)^2 \,|\, \psi_n \rangle \qquad (5.39) \\ &= \frac{\hbar^2}{x_0^2} \, (n + \tfrac{1}{2}) , \end{aligned}$$

where we have again used the characteristic length x_0, the ladder operator commutation relations (Eq. (5.11)) and the orthogonality of the eigenstates. For the momentum uncertainty we then have

$$(\Delta p)^2 = \langle p^2 \rangle - \langle p \rangle^2 = \frac{\hbar^2}{x_0^2} \, (n + \tfrac{1}{2}) . \tag{5.40}$$

Constructing the uncertainty relation from Eq. (5.37) and Eq. (5.40) we finally get

$$\begin{aligned} \Delta x \, \Delta p = \hbar \, (n + \tfrac{1}{2}) \;&\geq\; \tfrac{\hbar}{2} \quad \text{generally} && (5.41) \\ &=\; \tfrac{\hbar}{2} \quad \text{for the ground state} \;\; n = 0 . && (5.42) \end{aligned}$$

The ground state, Eq. (5.32), can be rewritten in terms of the characteristic length x_0

$$\psi_0(x) = \frac{1}{\sqrt{x_0 \sqrt{\pi}}} \, \exp(-\frac{x^2}{2x_0^2}) , \tag{5.43}$$

which is a gaussian wave packet with standard deviation x_0 characterizing the position uncertainty in the ground state, see Fig. 5.1.

Since we have made clear that the zero point energy is in accordance with the uncertainty relation we can even strengthen this correlation in so far as the non-vanishing zero point energy is a direct consequence of the uncertainty principle. Starting from the uncertainty relation (Eq. (2.80)) and keeping in mind that here the expectation values of the position- and momentum operator vanish identically, see Eq. (5.34) and Eq. (5.38), we can rewrite the uncertainty relation as

$$\langle\, x^2 \,\rangle \, \langle\, p^2 \,\rangle \, \geq \, \frac{\hbar^2}{4} \, . \tag{5.44}$$

We now calculate the mean energy of the harmonic oscillator, which is the expectation value of the Hamiltonian from Eq. (5.2)

$$E \, = \, \langle\, H \,\rangle \, = \, \frac{\langle\, p^2 \,\rangle}{2m} \, + \, \frac{m\omega^2}{2} \, \langle\, x^2 \,\rangle \, \overset{\text{Eq. (5.44)}}{\geq} \, \frac{\langle\, p^2 \,\rangle}{2m} \, + \, \frac{m\omega^2}{2} \, \frac{\hbar^2}{4} \, \frac{1}{\langle\, p^2 \,\rangle} \, . \tag{5.45}$$

To find the minimal energy we calculate the variation of Eq. (5.45) with respect to $\langle\, p^2 \,\rangle$, which we then set equal to zero to find the extremal values

$$\frac{1}{2m} \, - \, \frac{m\omega^2}{2} \, \frac{\hbar^2}{4} \, \frac{1}{\langle\, p^2 \,\rangle^2} \, = \, 0 \quad \Rightarrow \quad \langle\, p^2 \,\rangle_{\text{min}} \, = \, \frac{m\omega\hbar}{2} \, . \tag{5.46}$$

Inserting the result of Eq. (5.46) into the mean energy (Eq. (5.45)) we get

$$E \, \geq \, \frac{1}{2m} \, \frac{m\omega\hbar}{2} \, + \, \frac{m\omega^2\hbar^2}{8} \, \frac{2}{m\omega\hbar} \, = \, \frac{\hbar\omega}{4} \, + \, \frac{\hbar\omega}{4} \, = \, \frac{\hbar\omega}{2} \, , \tag{5.47}$$

thus

$$E \, \geq \, \frac{\hbar\omega}{2} \, . \tag{5.48}$$

Theorem 5.1 (Zero point energy)

The zero point energy is the smallest possible energy a physical system can possess, that is consistent with the uncertainty relation. It is the energy of its ground state.

5.1.3 Comparison with the Classical Oscillator

In this section we want to compare the quantum oscillator with predictions we would get from classical physics. There the classical motion is governed by the position function

$$x(t) \, = \, q_0 \, \sin(\omega t) \, , \tag{5.49}$$

and the total energy of the system is constant and given by

$$E \, = \, \frac{m\omega^2}{2} \, q_0^2 \, , \tag{5.50}$$

where in both cases, the angular frequency ω is fixed and q_0 is a given starting value. To make a comparison with quantum mechanics we can construct a classical probability density $W_{\text{class}}(x)$ in the following way

$$W_{\text{class}}(x)\, dx \;=\; \frac{dt}{T}\,, \tag{5.51}$$

where dt is the time period where the classical oscillator can be found in the volume element dx, and $T = \frac{2\pi}{\omega}$ is the full oscillation period. We then calculate the volume element dx by differentiating equation (5.49)

$$dx \;=\; q_0\, \omega\, \cos(\omega t)\, dt \;=\; q_0\, \omega \sqrt{1 \,-\, \sin^2(\omega t)}\; dt \;\overset{\text{Eq. (5.49)}}{=}\; q_0\, \omega \sqrt{1 \,-\, (\tfrac{x}{q_0})^2}\; dt\,. \tag{5.52}$$

Finally inserting Eq. (5.52) into Eq. (5.51) we get the classical probability density of the oscillator as a function of x, see Fig. 5.2 (dashed curve)

$$W_{\text{class}}(x) \;=\; \Big(2\pi\, q_0 \sqrt{1 \,-\, (\tfrac{x}{q_0})^2}\Big)^{-1}\,. \tag{5.53}$$

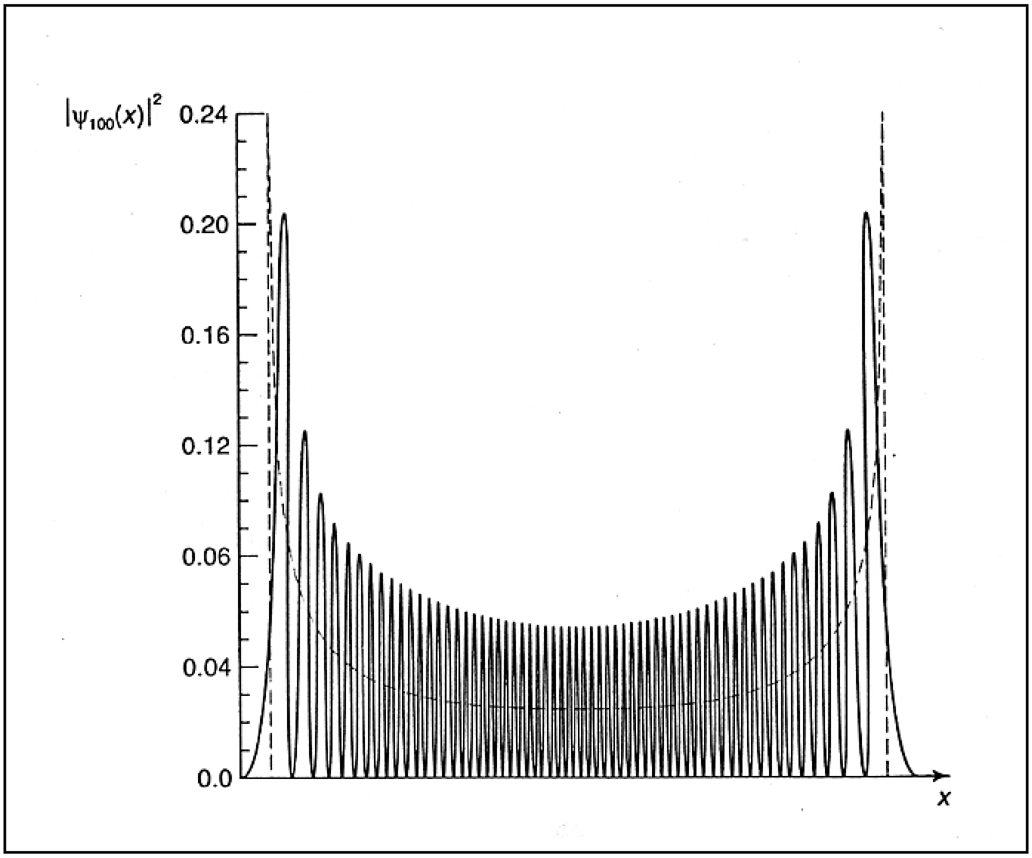

Figure 5.2: Oscillator probabilities: Comparison of the quantum probability (solid curve), $|\,\psi_n\,|^2$ in case of $n = 100$, with a classical oscillator probability (dashed curve), Eq. (5.53).

The starting value q_0 is fixed by the comparison of the energy of the quantum oscillator with the corresponding classical energy

$$E_{\mathrm{n}} \;=\; \frac{2n+1}{2}\,\hbar\omega \quad \longleftrightarrow \quad E_{\mathrm{class}} \;=\; \frac{m\omega^2}{2}\,q_0^2\ . \tag{5.54}$$

It allows us to relate the starting value q_0 of the classical oscillator to the characteristic length x_0 of the quantum mechanical oscillator

$$\Rightarrow \quad q_0 \;=\; \sqrt{\frac{(2n+1)\hbar}{m\omega}} \;=\; \sqrt{2n+1}\,x_0\ . \tag{5.55}$$

In the comparison quantum versus classical oscillator we find the following. For the low-lying states, of course, the two probabilities for finding the particle differ considerably but for the high-lying states the smoothed quantum probability resembles very much the classical one, as can be seen in Fig. 5.2 for the case $n = 100$.

5.1.4 Matrix Representation of the Harmonic Oscillator

In this section we now want to briefly sketch how the harmonic oscillator problem can be written in a matrix formulation, which introduces us to the concept of the *Fock space* and the *occupation number representation*. It is the appropriate formalism for relativistic quantum mechanics, i.e. Quantum Field Theory (QFT).

Starting from the eigenfunctions of the harmonic oscillator $\{\,\psi_{\mathrm{n}}(x)\,\}$, which form a complete orthonormal system of the corresponding Hilbert space, we change our notation and label the states only by their index number n, getting the following vectors[1]

$$\psi_0 \;\rightarrow\; \begin{pmatrix}1\\0\\0\\\vdots\end{pmatrix} = |\,0\,\rangle\ , \quad \psi_1 \;\rightarrow\; \begin{pmatrix}0\\1\\0\\\vdots\end{pmatrix} = |\,1\,\rangle\ , \quad \psi_{\mathrm{n}} \quad \rightarrow \underset{\text{n−th row}}{\nearrow} \begin{pmatrix}\vdots\\0\\1\\0\\\vdots\end{pmatrix} = |\,n\,\rangle\ . \tag{5.56}$$

The vectors $\{\,|\,n\,\rangle\,\}$ now also form a complete orthonormal system, but one of the so called *Fock space* or *occupation number space*[2]. We will not give a mathematical definition of this space but will just note here, that we can apply our rules and calculations as before on the Hilbert space. The ladder operators however have a special role in this space as they allow us to construct any vector from the ground state[3], see Eq. (5.58).

---

[1] Technically the vectors of Eq. (5.56) are infinite-dimensional, which means that the column-matrix notation is problematic, the ket-notation however is fine and we accept the column-vectors as visualization.

[2] In many body quantum mechanics and QFT the occupation number n is no mere label for the physical states, but describes the particle number of the considered state. This means that in QFT particles are described as excitations of a field, composed of a multitude of harmonic oscillators. The creation- and annihilation operators thus create and annihilate particles.

[3] which in QFT is the vacuum state, often also denoted by $|\,\Omega\,\rangle$.

Let us state the most important relations here:

$$\langle\, n \,|\, m \,\rangle \;=\; \delta_{\mathrm{n,m}} \tag{5.57}$$

$$|\, n \,\rangle \;=\; \frac{1}{\sqrt{n!}}\,(a^\dagger)^{\mathrm{n}}\;|\, 0 \,\rangle \tag{5.58}$$

$$a^\dagger\;|\, n \,\rangle \;=\; \sqrt{n\,+\,1}\;|\, n\,+\,1 \,\rangle \tag{5.59}$$

$$a\;|\, n \,\rangle \;=\; \sqrt{n}\;|\, n\,-\,1 \,\rangle \tag{5.60}$$

$$N\;|\, n \,\rangle \;=\; a^\dagger a\;|\, n \,\rangle \;=\; n\;|\, n \,\rangle \tag{5.61}$$

$$H\;|\, n \,\rangle \;=\; \hbar\omega\,(N\,+\,\tfrac{1}{2})\;|\, n \,\rangle \;=\; \hbar\omega\,(n\,+\,\tfrac{1}{2})\;|\, n \,\rangle\;. \tag{5.62}$$

With this knowledge we can write down matrix elements as transition amplitudes. We could, for example, consider Eq. (5.57) as the matrix element from the n-th row and the m-th column

$$\langle\, n \,|\, m \,\rangle \;=\; \delta_{\mathrm{n,m}} \quad\longleftrightarrow\quad \begin{pmatrix} 1 & 0 & \cdots \\ 0 & 1 & \\ \vdots & & \ddots \end{pmatrix}\;. \tag{5.63}$$

In total analogy we can get the matrix representation of the creation operator

$$\langle\, k \,|\; a^\dagger\;|\, n \,\rangle \;=\; \sqrt{n\,+\,1}\;\underbrace{\langle\, k \,|\, n\,+\,1 \,\rangle}_{\delta_{\mathrm{k,n+1}}} \quad\longleftrightarrow\quad a^\dagger \;=\; \begin{pmatrix} 0 & 0 & 0 & \cdots & \\ \sqrt{1} & 0 & 0 & & \\ 0 & \sqrt{2} & 0 & & \\ \vdots & & \sqrt{3} & & \\ & & & \ddots & \end{pmatrix} \tag{5.64}$$

as well as of the annihilation operator, occupation number operator and Hamiltonian

$$\langle\, k \,|\; a\;|\, n \,\rangle \;=\; \sqrt{n}\;\underbrace{\langle\, k \,|\, n\,-\,1 \,\rangle}_{\delta_{\mathrm{k,n-1}}} \quad\longleftrightarrow\quad a \;=\; \begin{pmatrix} 0 & \sqrt{1} & 0 & \cdots & \\ 0 & 0 & \sqrt{2} & & \\ 0 & 0 & 0 & \sqrt{3} & \\ \vdots & & & & \ddots \\ & & & & \end{pmatrix} \tag{5.65}$$

$$\langle\, k \,|\; N\;|\, n \,\rangle \;=\; n\;\underbrace{\langle\, k \,|\, n \,\rangle}_{\delta_{\mathrm{k,n}}} \quad\longleftrightarrow\quad N \;=\; \begin{pmatrix} 0 & 0 & 0 & \cdots & \\ 0 & 1 & 0 & & \\ 0 & 0 & 2 & & \\ \vdots & & & 3 & \\ & & & & \ddots \end{pmatrix} \tag{5.66}$$

$$\langle\, k \,|\; H \;|\, n \,\rangle \;=\; \hbar\omega\,(n \,+\, \tfrac{1}{2})\;\underbrace{\langle\, k \,|\, n \,\rangle}_{\delta_{\mathrm{k,n}}} \qquad \longleftrightarrow \qquad H \;=\; \hbar\omega \begin{pmatrix} \frac{1}{2} & 0 & 0 & \cdots & \\ 0 & \frac{3}{2} & 0 & & \\ 0 & 0 & \frac{5}{2} & & \\ \vdots & & & \frac{7}{2} & \\ & & & & \ddots \end{pmatrix} . \tag{5.67}$$

5.1.5 Analytic Method

In this last section on the harmonic oscillator we want to explore the straightforward method of solving the Schrödinger equation as a differential equation

$$H\,\psi \;=\; \left(-\,\frac{\hbar^2}{2m}\,\frac{d^2}{dx^2} \;+\; \frac{m\omega^2}{2}x^2 \right)\,\psi \;=\; E\,\psi\;. \tag{5.68}$$

Multiplying both sides by $\frac{2}{\hbar\omega}$ and introducing the dimensionless variables $\xi \;=\; \frac{x}{x_0}$ and $K \,=\, \frac{2E}{\hbar\omega}$, where we again have $x_0 = \sqrt{\frac{\hbar}{m\omega}}$, we obtain the differential equation

$$\frac{d^2}{d\xi^2}\,\psi(\xi) \;=\; \big(\xi^2 \,-\, K\big)\;\psi(\xi)\;. \tag{5.69}$$

To get an idea how the solution could look like, we first try to solve this equation for large ξ, which means we can neglect the K-terms

$$\frac{d^2}{d\xi^2}\,\psi(\xi) \;=\; \xi^2\,\psi(\xi)\;. \tag{5.70}$$

For this differential equation we try the ansatz $\psi = \exp(-\frac{\xi^2}{2})$, providing

$$\frac{d}{d\xi}\,\psi(\xi) \;=\; -\xi\,\exp(-\frac{\xi^2}{2}) \quad , \quad \frac{d^2}{d\xi^2}\,\psi(\xi) \;=\; \big(\xi^2 \,-\, 1\big)\;\psi(\xi)\;. \tag{5.71}$$

Since we are in the limit of large ξ, we can drop 1 in the last term on the right hand side of Eq. (5.71) and we have made the correct ansatz (in this approximation). To find the general solution we can use a similar ansatz, differing only insofar as we multiply it with a function of ξ, a method known as variation of the constants

$$\psi(\xi) \;=\; h(\xi)\,\exp(-\frac{\xi^2}{2})\;. \tag{5.72}$$

Differentiating our ansatz (5.72) we obtain

$$\frac{d}{d\xi}\,\psi(\xi) \quad = \quad \left(\frac{d\,h(\xi)}{d\xi} \,-\, \xi\,h(\xi) \right)\,\exp(-\frac{\xi^2}{2}) \tag{5.73}$$

$$\frac{d^2}{d\xi^2}\,\psi(\xi) \quad = \quad \left(\frac{d^2h(\xi)}{d\xi^2} \,-\, 2\,\xi\,\frac{d\,h(\xi)}{d\xi} \,+\, (\xi^2 \,-\, 1)\,h(\xi) \right)\,\exp(-\frac{\xi^2}{2})\;. \tag{5.74}$$

On the other hand we know the second derivative of ψ from the Schrödinger equation (5.69), leading to a differential equation for the function $h(\xi)$

$$\frac{d^2 h(\xi)}{d\xi^2} - 2\,\xi\,\frac{d\,h(\xi)}{d\xi} + (K - 1)\,h(\xi) = 0\,, \tag{5.75}$$

which can be solved by a power series ansatz. We are not computing this here, but direct the interested reader to a lecture and/or book about differential equations[4]. Equation (5.75) is the differential equation for the so called *Hermite polynomials* H_{n}, thus we replace our function by

$$h(\xi) \quad \rightarrow \quad H_{\mathrm{n}}(\xi) \quad , \quad n = 0, 1, 2, \cdots \tag{5.76}$$

where $(K - 1) = 2n$ or $K = \frac{2E}{\hbar\omega} = 2n + 1$ respectively. With this substitution we directly find the energy levels of the harmonic oscillator, as calculated earlier in Eq. (5.30)

$$E_{\mathrm{n}} = \hbar\omega\,(n + \tfrac{1}{2})\,. \tag{5.77}$$

With the help of the Hermite polynomials we can now also give the exact expression for the eigenfunctions of the harmonic oscillator

$$\psi_{\mathrm{n}} = \frac{1}{\sqrt{2^{\mathrm{n}}(n!)}}\left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} H_{\mathrm{n}}\big(\frac{x}{x_0}\big)\,\exp\big(-\frac{x^2}{2x_0^2}\big)\,. \tag{5.78}$$

The Hermite Polynomials can be easiest calculated with the *Formula of Rodrigues*

$$H_{\mathrm{n}}(x) = (-1)^{\mathrm{n}}\,e^{x^2}\left(\frac{d^{\mathrm{n}}}{dx^{\mathrm{n}}}\,e^{-x^2}\right)\,. \tag{5.79}$$

To give some examples we list some of the first Hermite Polynomials

$$\begin{aligned} H_0(x) &= 1\,, \quad H_1(x) = 2x\,, \quad H_2(x) = 4x^2 - 2 \\ H_3(x) &= 8x^3 - 12x\,, \quad H_4(x) = 16x^4 - 48x^2 + 12\,. \end{aligned} \tag{5.80}$$

A last interesting property is the orthogonality relation of the Hermite Polynomials

$$\int_{-\infty}^{\infty} dx\; H_{\mathrm{m}}(x)\;H_{\mathrm{n}}(x)\;e^{-x^2} = \sqrt{\pi}\,2^{\mathrm{n}}\,(n!)\,\delta_{\mathrm{mn}}\,. \tag{5.81}$$

5.2 Coherent States

Coherent states play an important role in quantum optics, especially in laser physics and much work was performed in this field by Roy J. Glauber who was awarded the 2005 Nobel prize for his contribution to the quantum theory of optical coherence. We will try here to give a good overview of coherent states of laser beams. The state describing a laser beam can be briefly characterized as having

---

[4] See for example [14].

1. an indefinite number of photons,
2. but a precisely defined phase,

in contrast to a state with fixed particle number, where the phase is completely random.

There also exists an uncertainty relation describing this contrast, which we will plainly state here but won't prove. It can be formulated for the uncertainties of amplitude and phase of the state, where the inequality reaches a minimum for coherent states, or, as we will do here, for the occupation number N and the phase Φ [5]

$$\Delta N \, \Delta(\sin\Phi) \, \geq \, \frac{1}{2} \cos\Phi \; , \tag{5.82}$$

which, for small Φ, reduces to

$$\Delta N \, \Delta\Phi \, \geq \, \frac{1}{2} \, . \tag{5.83}$$

5.2.1 Definition and Properties of Coherent States

Since laser light has a well-defined amplitude – in contrast to thermal light which is a statistical mixture of photons – we will define coherent states as follows:

Definition 5.3 *A **coherent state** $|\,\alpha\,\rangle$, also called **Glauber state**, is defined as eigenstate of the amplitude operator, the annihilation operator a, with eigenvalues $\alpha \in \mathbb{C}$*

$$a \,|\,\alpha\,\rangle \, = \, \alpha \,|\,\alpha\,\rangle \quad .$$

Since a is a non-hermitian operator the phase $\alpha \, = \, |\alpha| \, e^{i\varphi} \, \in \, \mathbb{C}$ is a complex number and corresponds to the complex wave amplitude in classical optics. Thus coherent states are wave-like states of the electromagnetic oscillator.

Properties of coherent states:

Let us study the properties of coherent states.

Note: The vacuum $|\,0\,\rangle$ is a coherent state with $\alpha = 0$.

Mean energy:

$$\langle\, H \,\rangle \, = \, \langle\, \alpha \,|\; H \;|\, \alpha \,\rangle \, = \, \hbar\omega \; \langle\, \alpha \,|\; a^\dagger a \, + \, \tfrac{1}{2} \;|\, \alpha \,\rangle \, = \, \hbar\omega \, (\, |\alpha|^2 \, + \, \tfrac{1}{2} \,) \, . \tag{5.84}$$

The first term on the right hand side represents the classical wave intensity and the second the vacuum energy.

---

[5] See for example [15].

Next we introduce the *phase shifting operator*

$$U(\theta) \,=\, e^{-i\theta N}\,, \tag{5.85}$$

where N represents the occupation number operator (see Definition 5.2). Then we have

$$U^\dagger(\theta)\, a\, U(\theta) \,=\, a\, e^{-i\theta} \quad \text{or} \quad U(\theta)\, a\, U^\dagger(\theta) \,=\, a\, e^{i\theta}\,, \tag{5.86}$$

i.e. $U(\theta)$ gives the amplitude operator a phase shift θ.

Proof: The l.h.s. $\frac{d}{d\theta} U^\dagger(\theta)\, a\, U(\theta) \,=\, i\, U^\dagger(\theta)\,[N, a]\, U(\theta) \,=\, -i\, U^\dagger(\theta)\, a\, U(\theta)$ and the r.h.s. $\frac{d}{d\theta}\, a\, e^{-i\theta} \,=\, -i\, a\, e^{-i\theta}$ obey the same differential equation.

The phase shifting operator also shifts the phase of the coherent state

$$U(\theta)\ |\,\alpha\,\rangle \,=\, \left|\,\alpha\, e^{-i\theta}\,\right\rangle\ . \tag{5.87}$$

Proof: Multiplying Definition 5.3 with the operator U from the left and inserting the unitarian relation $U^\dagger\, U \,=\, \mathbb{1}$ on the l.h.s we get

$$\begin{aligned} \underbrace{U\, a\, U^\dagger}_{a\, e^{i\theta}}\, U\ |\,\alpha\,\rangle \quad &= \quad \alpha\, U\ |\,\alpha\,\rangle \\ \Rightarrow \quad a\, U\ |\,\alpha\,\rangle \quad &= \quad \alpha\, e^{-i\theta}\, U\ |\,\alpha\,\rangle\ . \end{aligned} \tag{5.88}$$

Denoting $U\,|\,\alpha\,\rangle \,=\, \left|\,\alpha^{'}\,\right\rangle$ we find from Definition 5.3 that $a\left|\,\alpha^{'}\,\right\rangle \,=\, \alpha^{'}\left|\,\alpha^{'}\,\right\rangle$ and thus $\alpha^{'} = \alpha\, e^{-i\theta}$, q.e.d.

Now we want to study the coherent states more accurately. In order to do so we introduce the so called *displacement operator* which generates the coherent states, similar to how the creation operator $a^\dagger$, see Eq. (5.58), creates the occupation number states $|\,n\,\rangle$.

Definition 5.4 *The* ***displacement operator*** $D(\alpha)$ *is defined by*

$$D(\alpha) \,=\, e^{\,\alpha\, a^\dagger \,-\, \alpha^* a}$$

where $\alpha \,=\, |\alpha|\, e^{i\varphi} \,\in\, \mathbb{C}$ *is a complex number and* $a^\dagger$ *and* a *are the creation - and annihilation - operators.*

The displacement operator is an unitary operator, i.e. $D^\dagger D \,=\, \mathbb{1}$ and we can also rewrite it using Eq. (2.72). To use this equation we have to ensure that the commutators $[\,[\,A\,,\,B\,]\,,\,A\,]$ and $[\,[\,A\,,\,B\,]\,,\,B\,]$ vanish, where $A \,=\, \alpha\, a^\dagger$ and $B \,=\, \alpha^* a$. We therefore start by calculating the commutator of A and B

$$[\,A\,,\,B\,] \,=\, \left[\,\alpha\, a^\dagger\,,\,\alpha^* a\,\right] \,=\, \alpha\,\alpha^*\, \underbrace{\left[\,a^\dagger\,,\,a\,\right]}_{-1} \,=\, -\,|\alpha|^2\ . \tag{5.89}$$

Since the result is a real number, it commutes with A and B and we are allowed to use Eq. (2.72) where we can immediately insert the result of Eq. (5.89) to get the displacement operator in the following form:

$$D(\alpha) \,=\, e^{-\frac{1}{2}|\alpha|^2}\, e^{\alpha\, a^\dagger}\, e^{\alpha^* a} \,. \tag{5.90}$$

Properties of the displacement operator:

$$\text{I)} \qquad D^\dagger(\alpha) \,=\, D^{-1}(\alpha) \,=\, D(-\alpha) \quad \text{unitarity} \tag{5.91}$$

$$\text{II)} \qquad D^\dagger(\alpha)\, a\, D(\alpha) \,=\, a \,+\, \alpha \tag{5.92}$$

$$\text{III)} \qquad D^\dagger(\alpha)\, a^\dagger\, D(\alpha) \,=\, a^\dagger \,+\, \alpha^* \tag{5.93}$$

$$\text{IV)} \qquad D(\alpha \,+\, \beta) \,=\, D(\alpha)\, D(\beta)\, e^{-i\,\mathrm{Im}(\alpha\,\beta^*)} \tag{5.94}$$

We will prove two of these statements and leave the other two as an exercise.

Proof: of II)

$$\begin{aligned} D^\dagger(\alpha)\, a\, D(\alpha) \;&=\; e^{\alpha^* a \,-\, \alpha\, a^\dagger}\, a\, e^{\alpha\, a^\dagger \,-\, \alpha^* a} \,=\, a \,+\, \left[\,\alpha^* a \,-\, \alpha\, a^\dagger \,,\, a\,\right] \,= \\ &=\; a \,+\, \alpha^* \underbrace{\left[\,a\,,\,a\,\right]}_{0} - \alpha \underbrace{\left[\,a^\dagger\,,\,a\,\right]}_{-1} \,=\, a \,+\, \alpha \quad \text{q.e.d.} \end{aligned} \tag{5.95}$$

We have used the Baker-Campbell-Hausdorff formula (Eq. (2.71)), where $A = \alpha^* a \,-\, \alpha\, a^\dagger$ and $B = a$. The higher order commutators vanish, since the commutator of A and B is a complex number, that commutes with the other operators.

Proof: of IV)

$$\begin{aligned} D(\alpha \,+\, \beta) \;&=\; e^{\alpha\, a^\dagger \,-\, \alpha^* a \,+\, \beta\, a^\dagger \,-\, \beta^* a} \,=\, e^{\alpha\, a^\dagger \,-\, \alpha^* a}\, e^{\beta\, a^\dagger \,-\, \beta^* a}\, e^{-\frac{1}{2}\left[\,\alpha\, a^\dagger \,-\, \alpha^* a \,,\, \beta\, a^\dagger \,-\, \beta^* a\,\right]} \,= \\ &=\; D(\alpha)\, D(\beta)\, e^{-\frac{1}{2}\,(\alpha\beta^* \,-\, \alpha^*\beta)} \,=\, D(\alpha)\, D(\beta)\, e^{-i\,\mathrm{Im}(\alpha\,\beta^*)} \quad \text{q.e.d.} \end{aligned} \tag{5.96}$$

Here we have used Eq. (2.72), with $A = \alpha\, a^\dagger \,-\, \alpha^* a$ and $B = \beta\, a^\dagger \,-\, \beta^* a$ which is justified because the condition that the higher commutators vanish is satisfied since $[\,A\,,\,B\,] \,\in\, \mathbb{C}$. The result of this last commutator can also be seen easily by noting that the creation- and annihilation-operators commute with themselves but not with each other, giving ±1.

With this knowledge about the displacement operator we can now create coherent states.

Theorem 5.2 *The coherent state $|\,\alpha\,\rangle$ is generated from the vacuum $|\,0\,\rangle$ by the displacement operator $D(\alpha)$*

$$|\,\alpha\,\rangle \;=\; D(\alpha)\;|\,0\,\rangle\;.$$

The "vacuum" $|\,0\,\rangle$ is the ground state with occupation number $n = 0$, which is defined by $a\;|\,0\,\rangle \;=\; 0$, see also Eq. (5.18).

Proof: Applying a negative displacement to $|\,\alpha\,\rangle$ we find from properties (5.91) and (5.92) that

$$\begin{aligned} &a\,D(-\alpha)\;|\,\alpha\,\rangle = D(-\alpha)\;\underbrace{D^\dagger(-\alpha)\,a\,D(-\alpha)}_{a-\alpha}\;|\,\alpha\,\rangle \\ \Rightarrow\quad &a\,D(-\alpha)\;|\,\alpha\,\rangle \;=\; D(-\alpha)(a-\alpha)\;|\,\alpha\,\rangle \;=\; 0\,. \end{aligned} \tag{5.97}$$

The r.h.s. vanishes because of Definition 5.3. This implies that $D(-\alpha)\;|\,\alpha\,\rangle$ is the vacuum state $|\,0\,\rangle$

$$D(-\alpha)\;|\,\alpha\,\rangle \;=\; |\,0\,\rangle \qquad\Rightarrow\qquad |\,\alpha\,\rangle \;=\; D(\alpha)\;|\,0\,\rangle \quad \text{q.e.d.} \tag{5.98}$$

On the other hand, we also could use Theorem 5.2 as a definition of the coherent states then our Definition 5.3 would follow as a theorem.

Proof:

$$\begin{aligned} \underbrace{D(\alpha)\,D^\dagger(\alpha)}_{\mathbb{1}}\;a\;|\,\alpha\,\rangle \;&=\; D(\alpha)\;\underbrace{D^\dagger(\alpha)\,a\,D(\alpha)}_{a+\alpha}\;|\,0\,\rangle = D(\alpha)\;\Bigg(\underbrace{a\;|\,0\,\rangle}_{0} + \alpha\;|\,0\,\rangle\Bigg) \;= \\ &=\; \alpha\,D(\alpha)\;|\,0\,\rangle \;=\; \alpha\;|\,\alpha\,\rangle \\ &\Rightarrow\quad a\;|\,\alpha\,\rangle \;=\; \alpha\;|\,\alpha\,\rangle \quad \text{q.e.d.} \end{aligned} \tag{5.99}$$

The meaning of our Definition 5.3 becomes apparent when describing a coherent photon beam, e.g. a laser beam. The state of the laser remains unchanged if one photon is annihilated, i.e. if it is detected.

Expansion of coherent state in Fock space: The phase $|\,\alpha\,\rangle$ describes the wave aspect of the coherent state. Next we want to study the particle aspect, the states in Fock space. A coherent state contains an indefinite number of photons, which will be evident from the expansion of the coherent state $|\,\alpha\,\rangle$ into the CONS of the occupation number states $\{\,|\,n\,\rangle\,\}$. Let us start by inserting the completeness relation of the occupation number states

$$|\,\alpha\,\rangle \;=\; \sum_n \,|\,n\,\rangle\,\langle\,n\,|\,\alpha\,\rangle \;. \tag{5.100}$$

Then we calculate the transition amplitude $\langle\,n\,|\,\alpha\,\rangle$ by using Definition 5.3, where we multiply the whole eigenvalue equation with $\langle\,n\,|$ from the left

$$\langle\,n\,|\;a\;|\alpha\,\rangle \;=\; \langle\,n\,|\;\alpha\;|\alpha\,\rangle \;. \tag{5.101}$$

Using the adjoint of relation (5.59) we rewrite the left side of Eq. (5.101)

$$a^\dagger\;|\,n\,\rangle \;=\; \sqrt{n\,+\,1}\;|\,n\,+\,1\,\rangle \quad\Rightarrow\quad \langle\,n\,|\;a \;=\; \sqrt{n\,+\,1}\;\langle\,n\,+\,1\,| \tag{5.102}$$

$$\sqrt{n\,+\,1}\;\langle\,n\,+\,1\;|\,\alpha\,\rangle \;=\; \alpha\;\langle\,n\,|\,\alpha\,\rangle\;, \tag{5.103}$$

and replace the occupation number n with $n-1$ to obtain

$$\langle\,n\,|\,\alpha\,\rangle \;=\; \frac{\alpha}{\sqrt{n}}\;\langle\,n\,-\,1\;|\,\alpha\,\rangle\;. \tag{5.104}$$

By iterating the last step, i.e. again replacing n with $n-1$ in Eq. (5.104) and reinserting the result on the right hand side, we get

$$\langle\,n\,|\,\alpha\,\rangle \;=\; \frac{\alpha^2}{\sqrt{n(n\,-\,1)}}\;\langle\,n\,-\,2\;|\,\alpha\,\rangle \;=\;\cdots\;=\; \frac{\alpha^n}{\sqrt{n!}}\;\langle\,0\,|\,\alpha\,\rangle\;, \tag{5.105}$$

which inserted into the expansion of Eq. (5.100) results in

$$|\,\alpha\,\rangle \;=\; \langle\,0\,|\,\alpha\,\rangle\,\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\;|\,n\,\rangle\;. \tag{5.106}$$

The remaining transition amplitude $\langle\,0\,|\,\alpha\,\rangle$ can be calculated in two different ways, first with the normalization condition and second with the displacement operator. For the sake of variety we will use the second method here and leave the other one as an exercise

$$\langle\,0\,|\,\alpha\,\rangle \;=\; \langle\,0\,|\;D(\alpha)\;|\,0\,\rangle \;\overset{\text{Eq. (5.90)}}{=}\; e^{-\frac{1}{2}|\alpha|^2}\;\langle\,0\,|\;e^{\alpha\,a^\dagger}\;e^{\alpha^* a}\;|\,0\,\rangle\;. \tag{5.107}$$

We expand the exponentials of operators in the scalar product into their Taylor series

$$\langle\,0\,|\;e^{\alpha\,a^\dagger}\;e^{\alpha^* a}\;|\,0\,\rangle \;=\; \langle\,0\,|\;(1\,+\,\alpha\,a^\dagger\,+\,\cdots)(1\,+\,\alpha^*\,a\,+\,\cdots)\;|\,0\,\rangle \;=\; 1\;, \tag{5.108}$$

and we let the left bracket act to the left on the bra and right bracket to the right on the ket. In this way we can both times use the vacuum state property $a\;|\,0\,\rangle \;=\; 0$ to eliminate all terms of the expansion except the first. Therefore the transition amplitude is given by

$$\langle\,0\,|\,\alpha\,\rangle \;=\; e^{-\frac{1}{2}|\alpha|^2}\;, \tag{5.109}$$

and we can write down the coherent state in terms of an exact expansion in Fock space

$$|\,\alpha\,\rangle \;=\; e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\;|\,n\,\rangle \;=\; e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{(\alpha\, a^\dagger)^n}{n!}\;|\,0\,\rangle\;. \tag{5.110}$$

Probability distribution of coherent states: We can subsequently analyze the probability distribution of the photons in a coherent state, i.e. the probability of detecting n photons in a coherent state $|\,\alpha\,\rangle$, which is given by

$$P(n) \;=\; |\langle\, n\,|\,\alpha\,\rangle|^2 \;=\; \frac{|\alpha|^{2n}\, e^{-|\alpha|^2}}{n!}\;. \tag{5.111}$$

By noting, that the *mean photon number* is determined by the expectation value of the particle number operator

$$\bar{n} \;=\; \langle\,\alpha\,|\;N\;|\,\alpha\,\rangle \;=\; \langle\,\alpha\,|\;a^\dagger a\;|\,\alpha\,\rangle \;\overset{\text{Def. 5.3}}{=}\; |\alpha|^2\;, \tag{5.112}$$

we can rewrite the probability distribution to get

$$P(n) \;=\; \frac{\bar{n}^n\, e^{-\bar{n}}}{n!}\;, \tag{5.113}$$

which is a *Poissonian distribution.*

Remarks: First, Eq. (5.112) represents the connection between the mean photon number – the particle view – and the complex amplitude squared, the intensity of the wave – the wave view.

Second, classical particles obey the same statistical law, the Poisson formula (5.111), when they are taken at random from a pool with $|\alpha|^2$ on average. Counting therefore the photons in a coherent state, they behave like randomly distributed classical particles, which might not be too surprising since coherent states are wave-like. Nevertheless, it's amusing that the photons in a coherent state behave like the raisins in a Gugelhupf[6], which have been distributed from the cook randomly with $|\alpha|^2$ on average per unit volume. Then the probability of finding n raisins per unit volume in the Gugelhupf follows precisely law (5.111).

Scalar product of two coherent states: From Theorem 4.3 we can infer some important properties of the particle number states, which are the eigenstates of the particle number operator $N = a^\dagger a$. Since this operator is hermitian, the eigenvalues n are real and the eigenstates $|\,n\,\rangle$ are orthogonal. But the coherent states are eigenstates of the annihilation operator a, which is surely not hermitian and as we already know, the eigenvalues α are complex numbers. Therefore we cannot automatically assume the coherent states

---

[6] A traditional Austrian birthday cake.

to be orthogonal, but have to calculate their scalar product, using again Theorem 5.2 and Eq. (5.90)

$$\langle\,\beta\,|\,\alpha\,\rangle \;=\; \langle\,0\,|\; D^\dagger(\beta)\, D(\alpha)\;|\,0\,\rangle \;=\; \langle\,0\,|\; e^{-\beta\, a^\dagger}\, e^{\beta^*\, a}\, e^{\alpha\, a^\dagger}\, e^{-\alpha^* a}\;|\,0\,\rangle\; e^{-\frac{1}{2}(|\alpha|^2 + |\beta|^2)}\;. \tag{5.114}$$

We now use the same trick as in Eq. (5.108), i.e., expand the operators in their Taylor series, to see that the two "outer" operators $e^{-\beta a^\dagger} = (1 - \beta\, a^\dagger + \cdots)$ and $e^{-\alpha^* a} = (1 - \alpha^* a + \cdots)$, acting to left and right respectively, annihilate the vacuum state, save for the term "1" and we can thus ignore them. We then expand the remaining operators as well and consider their action on the vacuum

$$\begin{aligned} &\langle\,0\,|\;(1 \,+\, \beta^* a \,+\, \frac{1}{2!}(\beta^* a)^2 \,+\, \cdots)\,(1 \,+\, \alpha\, a^\dagger \,+\, \frac{1}{2!}(\alpha\, a^\dagger)^2 \,+\, \cdots)\;|\,0\,\rangle \qquad (5.115)\\ &= (\,\cdots \,+\, \langle\,2\,|\,\frac{1}{2!}\,\sqrt{2!}\,(\beta^*)^2 \,+\, \langle\,1\,|\,\beta^* \,+\, \langle\,0\,|\,)\,(\,|\,0\,\rangle \,+\, \alpha\,|\,1\,\rangle \,+\, \frac{1}{2!}\,\sqrt{2!}\,\alpha^2\;|\,2\,\rangle \,+\, \cdots)\;. \end{aligned}$$

Using the orthogonality of the particle number states $\langle\, n\,|\, m\,\rangle \,=\, \delta_{\mathrm{nm}}$ we get

$$\langle\,\beta\,|\,\alpha\,\rangle \;=\; e^{-\frac{1}{2}(|\alpha|^2 + |\beta|^2)}\,(\,1 \,+\, \alpha\,\beta^* \,+\, \frac{1}{2!}\,(\alpha\,\beta^*)^2 \,+\, \cdots\,) \;=\; e^{-\frac{1}{2}(|\alpha|^2 + |\beta|^2) + \alpha\,\beta^*}\;. \tag{5.116}$$

We further simplify the exponent, by noting that

$$|\,\alpha \,-\, \beta\,|^2 \;=\; (\alpha \,-\, \beta)(\alpha^* \,-\, \beta^*) \;= |\alpha|^2 \,+\, |\beta|^2 \,-\alpha\beta^* \,-\, \alpha^*\beta\;, \tag{5.117}$$

and we can finally write down the transition probability in a compact way

$$|\,\langle\,\beta\,|\,\alpha\,\rangle\,|^2 \;=\; e^{-|\,\alpha\,-\,\beta\,|^2}\;. \tag{5.118}$$

That means, the coherent states indeed are *not orthogonal* and their transition probability only vanishes in the limit of large differences $|\,\alpha \,-\, \beta\,| \gg 1$.

Completeness of coherent states: Although the coherent states are not orthogonal, it is possible to expand coherent states in terms of a complete set of states. The *completeness relation* for the coherent states reads

$$\frac{1}{\pi}\int d^2\alpha\;\;|\,\alpha\,\rangle\,\langle\,\alpha\,| \;=\; \mathbb{1}\;. \tag{5.119}$$

In fact, the coherent states are "overcomplete", which means that, as a consequence of their nonorthogonality, any coherent state can be expanded in terms of all the other coherent states. So the coherent states are not linearly independent

$$|\,\beta\,\rangle \;=\; \frac{1}{\pi}\int d^2\alpha\;\;|\,\alpha\,\rangle\,\langle\,\alpha\,|\,\beta\,\rangle \;=\; \frac{1}{\pi}\int d^2\alpha\;\;|\,\alpha\,\rangle\; e^{-\frac{1}{2}(|\alpha|^2 + |\beta|^2) + \alpha\,\beta^*}\;. \tag{5.120}$$

Proof of completeness:

$$\frac{1}{\pi}\int d^2\alpha\ \,|\,\alpha\,\rangle\,\langle\,\alpha\,|\ \overset{Eq.\ (5.110)}{=}\ \frac{1}{\pi}\sum_{n,m}\frac{1}{\sqrt{n!\,m!}}\ |\,n\,\rangle\,\langle\,m\,|\int d^2\alpha\ e^{-|\alpha|^2}\,\alpha^{\mathrm{n}}(\alpha^*)^{\mathrm{m}}\,. \tag{5.121}$$

The integral on the right hand side of Eq. (5.121) can be solved using polar coordinates $\alpha = |\alpha|e^{i\varphi} = r\,e^{i\varphi}$

$$\int d^2\alpha\ e^{-|\alpha|^2}\,\alpha^{\mathrm{n}}(\alpha^*)^{\mathrm{m}}\ =\ \int\limits_0^\infty r\,dr\,e^{-r^2}\,r^{n+m}\underbrace{\int\limits_0^{2\pi} d\varphi\,e^{i(n-m)\varphi}}_{2\pi\,\delta_{\mathrm{nm}}}\ =\ 2\pi\int\limits_0^\infty r\,dr\,e^{-r^2}\,r^{2n}\,. \tag{5.122}$$

Substituting $r^2 = t$, $2rdr = dt$ to rewrite the integral

$$2\int\limits_0^\infty r\,dr\,e^{-r^2}\,r^{2n}\ =\ \int\limits_0^\infty e^{-t}\,t^n\,dt\ =\ \Gamma(n\,+\,1)\ =\ n!\,, \tag{5.123}$$

we recover the exact definition of the Gamma function, see [13, 14]. With this we can finally write down the completeness relation

$$\frac{1}{\pi}\int d^2\alpha\ \,|\,\alpha\,\rangle\,\langle\,\alpha\,|\ =\ \frac{1}{\pi}\sum_n\frac{1}{n!}\ |\,n\,\rangle\,\langle\,n\,|\ \pi\,n!\ =\ \sum_n\,|\,n\,\rangle\,\langle\,n\,|\ =\ \mathbb{1}\quad \text{q.e.d.} \tag{5.124}$$

5.2.2 Coordinate Representation of Coherent States

Recalling the coherent state expansion from Eq. (5.110)

$$|\,\alpha\,\rangle\ =\ e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\ |\,n\,\rangle\ =\ e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{(\alpha\,a^\dagger)^n}{n!}\ |\,0\,\rangle\ , \tag{5.125}$$

and the coordinate representation of the harmonic oscillator states (Eq. (5.78))

$$\psi_{\mathrm{n}}(x)\ =\ \langle\,x\,|\,n\,\rangle\ \ =\ \ \psi_{\mathrm{n}}(x)\ =\ \frac{1}{\sqrt{2^{\mathrm{n}}(n!)\sqrt{\pi}x_0}}\,H_{\mathrm{n}}(\xi)\ e^{-\frac{\xi^2}{2}} \tag{5.126}$$

$$\psi_0(x)\ =\ \langle\,x\,|\,0\,\rangle\ \ =\ \ \psi_0(x)\ =\ \frac{1}{\sqrt{\sqrt{\pi}x_0}}\ e^{-\frac{\xi^2}{2}}\ , \tag{5.127}$$

where $\xi = \frac{x}{x_0}$ and $x_0 = \sqrt{\frac{\hbar}{m\omega}}$, we can easily give the coordinate representation of the coherent states

$$\langle\,x\,|\,\alpha\,\rangle\ =\ \phi_\alpha(x)\ =\ e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\,\psi_{\mathrm{n}}(x)\ =\ e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{(\alpha\,a^\dagger)^n}{n!}\,\psi_0(x)\ . \tag{5.128}$$

Time evolution: Let's first take a look at the time evolution of the harmonic oscillator states and at the end we recall Theorem 4.1 and Definition 4.1. Using the harmonic oscillator energy (Eq. (5.77)) we easily get

$$\psi_{\mathrm{n}}(t,x) \;=\; \psi_{\mathrm{n}}(x)\, e^{-\frac{i}{\hbar}E_{\mathrm{n}}t} \;=\; \psi_{\mathrm{n}}(x)\, e^{-in\,\omega t}\, e^{-\frac{i\omega t}{2}} \;. \tag{5.129}$$

We are then in the position to write down the time evolution of the coherent state

$$\phi_{\alpha}(t,x) \;=\; e^{-\frac{1}{2}|\alpha|^2}\, e^{-\frac{i\omega t}{2}} \sum_{n=0}^{\infty} \frac{(\alpha\, e^{-i\omega t})^n}{\sqrt{n!}}\, \psi_{\mathrm{n}}(x) \;. \tag{5.130}$$

With a little notational trick, i.e. making the label α time dependent $\alpha(t) = \alpha\, e^{-i\omega t}$, we can bring Eq. (5.130) into a more familiar form

$$\phi_{\alpha}(t,x) \;=\; \phi_{\alpha(\mathrm{t})}(x)\, e^{-\frac{i\omega t}{2}} \;, \tag{5.131}$$

which we can identify as a solution of the time dependent Schrödinger equation.

Expectation value of x for coherent states: Subsequently, we want to compare the motion of coherent states to that of the quantum mechanical (and classical) harmonic oscillator, which we will do by studying the expectation value of the position operator. For the quantum mechanical case we have already calculated the mean position, which does not oscillate (see Eq. (5.34))

$$\langle\, x \,\rangle_{\mathrm{oscillator}} \;=\; 0 \,. \tag{5.132}$$

Recalling the expression of x in terms of creation and annihilation operators, Eq. (5.5), and the eigenvalue equation of the annihilation operator, Definition 5.3, we can easily compute the expectation value of the position for the coherent states

$$\begin{aligned} \langle\, x \,\rangle_{\mathrm{coherent}} \;&=\; \left\langle\, \phi_{\alpha(\mathrm{t})} \,\middle|\, x \,\middle|\, \phi_{\alpha(\mathrm{t})} \,\right\rangle \;=\; \frac{x_0}{\sqrt{2}} \left\langle\, \phi_{\alpha(\mathrm{t})} \,\middle|\, a \,+\, a^{\dagger} \,\middle|\, \phi_{\alpha(\mathrm{t})} \,\right\rangle \;= \\ &=\; \frac{x_0}{\sqrt{2}}\, (\alpha(t) \,+\, \alpha^*(t)) \;=\; \sqrt{2}\, x_0\, \mathrm{Re}(\alpha(t)) \;=\; \sqrt{2}\, x_0\, |\alpha| \cos(\omega \mathrm{t} \,-\, \varphi) \;, \end{aligned} \tag{5.133}$$

where we used $\alpha(t) = \alpha\, e^{-i\omega t} = |\alpha|\, e^{-i(\omega t - \varphi)}$. To summarize the calculation, we conclude that the coherent state, unlike the quantum mechanical harmonic oscillator, *does oscillate*, similar to its classical analogue

$$\langle\, x \,\rangle_{\mathrm{coherent}} \;=\; \sqrt{2}\, x_0\, |\alpha| \cos(\omega \mathrm{t} \,-\, \varphi) \;. \tag{5.134}$$

Coordinate representation in terms of displacement operator: In this section we want to express and study the coordinate representation of coherent states via the displacement operator D

$$\phi_\alpha(x) \,=\, \langle\, x \,|\, \alpha \,\rangle \,=\, \langle\, x \,|\; D(\alpha)\; |\, 0 \,\rangle \,=\, \langle\, x \,|\; e^{\alpha\, a^\dagger \,-\, \alpha^* a}\; |\, 0 \,\rangle\ . \tag{5.135}$$

To simplify further calculations we now rewrite the creation and annihilation operators from Definition 5.1 in terms of the dimensionless variable $\xi = \frac{x}{x_0}$,

$$a \,=\, \tfrac{1}{\sqrt{2}}(\xi \,+\, \frac{d}{d\xi}) \quad , \quad a^\dagger \,=\, \tfrac{1}{\sqrt{2}}(\xi \,-\, \frac{d}{d\xi})\ . \tag{5.136}$$

This rewritten operators (5.136) we insert into the exponent of the displacement operator from Eq. (5.135)

$$\begin{aligned} \alpha\, a^\dagger \,-\, \alpha^* a \;&=\; \frac{\alpha}{\sqrt{2}}(\xi \,-\, \frac{d}{d\xi}) \,-\, \frac{\alpha^*}{\sqrt{2}}(\xi \,+\, \frac{d}{d\xi}) \,= \\ &=\; \frac{\xi}{\sqrt{2}}(\underbrace{\alpha \,-\, \alpha^*}_{2i\,\mathrm{Im}(\alpha)}) \,-\, \frac{1}{\sqrt{2}}\frac{d}{d\xi}(\underbrace{\alpha \,+\, \alpha^*}_{2\,\mathrm{Re}(\alpha)}) \,= \\ &=\; \sqrt{2}\, i\, \mathrm{Im}(\alpha)\, \xi \,-\, \sqrt{2}\, \mathrm{Re}(\alpha)\, \frac{d}{d\xi}\ , \end{aligned} \tag{5.137}$$

and get for the wavefunction (5.135) the following expression

$$\phi_\alpha(x) \,=\, \langle\, x \,|\, \alpha \,\rangle \,=\, e^{\sqrt{2}\, i\, \mathrm{Im}(\alpha)\, \xi \,-\, \sqrt{2}\, \mathrm{Re}(\alpha)\, \frac{d}{d\xi}}\ \psi_0(\xi)\ , \tag{5.138}$$

where ψ_0 is the harmonic oscillator ground state (Eq. (5.127)).

The time dependent version follows outright

$$\phi_\alpha(t,x) \,=\, \phi_{\alpha(\mathrm{t})}\, e^{-\frac{i\omega t}{2}} \,=\, \langle\, x \,|\, \alpha(t) \,\rangle\, e^{-\frac{i\omega t}{2}} \,=\, e^{-\frac{i\omega t}{2}}\, \langle\, x \,|\; e^{\alpha(t)\, a^\dagger \,-\, \alpha^*(t) a}\; |\, 0 \,\rangle$$

$$\Rightarrow\quad \phi_\alpha(t,x) \,=\, e^{-\frac{i\omega t}{2}}\, e^{\sqrt{2}\, i\, \mathrm{Im}(\alpha(t))\, \xi \,-\, \sqrt{2}\, \mathrm{Re}(\alpha(t))\, \frac{d}{d\xi}}\ \psi_0(\xi)\ . \tag{5.139}$$

On the other hand, as a check, we can also arrive at this result by starting from expansion (5.125), where we interpret the sum over n as the power series for the exponential map, i.e.

$$\phi_\alpha(t,x) \,=\, e^{-\frac{1}{2}|\alpha|^2}\, e^{-\frac{i\omega t}{2}} \sum_{n=0}^{\infty} \frac{(\alpha(t)\, a^\dagger)^n}{n!}\, \psi_0(x) \,=\, e^{-\frac{1}{2}|\alpha|^2}\, e^{-\frac{i\omega t}{2}}\, e^{\alpha(t)\, a^\dagger}\, \psi_0(x)\ . \tag{5.140}$$

To show that this result is the same as Eq. (5.139) we simply calculate

$$\begin{aligned} e^{\alpha(t)\,a^\dagger\,-\,\alpha^*(t)\,a}\,\psi_0(x) \;&\overset{\text{Eq. (2.72)}}{=}\; e^{\alpha(t)\,a^\dagger}\;e^{-\alpha^*(t)\,a}\;e^{\frac{1}{2}\,\alpha\,\alpha^*\,\overbrace{\left[\,a^\dagger\,,\,a\,\right]}^{-1}}\;\psi_0(x)\;= \\ &=\; e^{-\frac{1}{2}|\alpha|^2}\;e^{\alpha(t)\,a^\dagger}\;(1\,-\,\alpha^*(t)\;\overset{\longrightarrow 0}{a}\;+\cdots)\;\psi_0(x)\;= \\ &=\; e^{-\frac{1}{2}|\alpha|^2}\;e^{\alpha(t)\,a^\dagger}\;\psi_0(x)\;, \end{aligned} \tag{5.141}$$

where we once again used the fact, that the action of the annihilation operator on the ground state vanishes. Inserting the result into Eq. (5.140) we get

$$\phi_\alpha(t,x)\;=\;e^{-\frac{i\omega t}{2}}\;e^{\alpha(t)\,a^\dagger\,-\,\alpha^*(t)\,a}\;\psi_0(x)\;. \tag{5.142}$$

Explicit calculation of the wave function: Our aim is now to explicitly compute the wave function from expression (5.138). We first state the result and derive it in the following

$$\begin{aligned} \phi_\alpha(t,x)\;&=\;\frac{1}{\sqrt{x_0\,\sqrt{\pi}}}\;e^{-\frac{i\omega t}{2}}\;e^{\sqrt{2}\,i\,\mathrm{Im}(\alpha(t))\,\xi\,-\,\sqrt{2}\,\mathrm{Re}(\alpha(t))\,\frac{d}{d\xi}}\;e^{-\frac{\xi^2}{2}} \\ &=\;\frac{1}{\sqrt{x_0\,\sqrt{\pi}}}\;e^{-\frac{i\omega t}{2}}\;e^{\sqrt{2}\,\alpha(t)\,\xi\,-\,\frac{\xi^2}{2}\,-\,\mathrm{Re}(\alpha(t))\,\alpha(t)}\;. \end{aligned} \tag{5.143}$$

Proof: Let us now perform the explicit calculation. That means, we apply the differential operator[7] $e^{\mathcal{D}} = e^{\sqrt{2}\,i\,\mathrm{Im}(\alpha(t))\,\xi\,-\,\sqrt{2}\,\mathrm{Re}(\alpha(t))\,\frac{d}{d\xi}}$ to the ground state of the harmonic oscillator $\psi_0(\xi) = \frac{1}{\sqrt{x_0\,\sqrt{\pi}}}e^{-\frac{\xi^2}{2}}$, but we will only consider the linear and quadratic term of the operator expansion, as sketched in Eq. (5.144)

$$e^{\mathcal{D}}\,e^{-\frac{\xi^2}{2}}\;=\;(1\,+\,\mathcal{D}\,+\,\frac{1}{2!}\,\mathcal{D}^2\,+\,\cdots)\,e^{-\frac{\xi^2}{2}}\;, \tag{5.144}$$

$$\begin{aligned} \text{linear term:}\qquad \mathcal{D}\,e^{-\frac{\xi^2}{2}}\;&=\;\Big(\sqrt{2}\,i\,\mathrm{Im}(\alpha(t))\,\xi\,-\,\sqrt{2}\,\mathrm{Re}(\alpha(t)\,\tfrac{d}{d\xi}\Big)\;e^{-\frac{\xi^2}{2}} \\ &=\;\Big(\sqrt{2}\,i\,\mathrm{Im}(\alpha(t))\,\xi\,+\,\sqrt{2}\,\mathrm{Re}(\alpha(t)\,\xi\Big)\;e^{-\frac{\xi^2}{2}} \\ &=\;\sqrt{2}\,\xi\;\underbrace{(\mathrm{Re}(\alpha(t))\,+\,i\,\mathrm{Im}(\alpha(t)))}_{\alpha(t)}\;e^{-\frac{\xi^2}{2}} \\ &=\;\sqrt{2}\,\alpha(t)\,\xi\,e^{-\frac{\xi^2}{2}}\;, \end{aligned} \tag{5.145}$$

---

[7] The operator $\mathcal{D}$, which is a notation to shorten the cumbersome calculations, is not to be confused with the displacement operator D.

quadratic term:

$$\begin{aligned}
\mathcal{D}^2\, e^{-\frac{\xi^2}{2}} &= \sqrt{2}\,\alpha(t)\,\mathcal{D}\,\xi\, e^{-\frac{\xi^2}{2}} \\
&= \sqrt{2}\,\alpha(t)\,\Big(\sqrt{2}\,i\,\mathrm{Im}(\alpha(t))\,\xi \,-\, \sqrt{2}\,\mathrm{Re}(\alpha(t))\,\tfrac{d}{d\xi}\Big)\,\xi\, e^{-\frac{\xi^2}{2}} \\
&= \big(2\,\alpha(t)\,i\,\mathrm{Im}(\alpha(t))\,\xi^2 \,-\, 2\,\alpha(t)\,\mathrm{Re}(\alpha(t)) \,+\, 2\,\alpha(t)\,\mathrm{Re}(\alpha(t))\,\xi^2\,\big)\; e^{-\frac{\xi^2}{2}} \\
&= (2\,\alpha(t)\,\underbrace{(\mathrm{Re}(\alpha(t)) \,+\, i\,\mathrm{Im}(\alpha(t)))}_{\alpha(t)}\,\xi^2 \,-\, 2\,\alpha(t)\,\mathrm{Re}(\alpha(t)))\, e^{-\frac{\xi^2}{2}}\ .
\end{aligned} \tag{5.146}$$

Inserting the results of Eq. (5.145) and Eq. (5.146) into the expansion (5.144) we get

$$e^{\mathcal{D}}\, e^{-\frac{\xi^2}{2}} \,=\, (1 \,+\, \underbrace{\sqrt{2}\,\alpha(t)\xi \,-\, \alpha(t)\,\mathrm{Re}(\alpha(t))}_{\text{linear term}} \,+ \frac{2\,\alpha(t)\,\alpha(t)\,\xi^2}{2!} \,+\, \cdots)\, e^{-\frac{\xi^2}{2}}\ , \tag{5.147}$$

where we can already identify the linear term and the first part of the quadratic term in the end result of the expansion. Finally, rewriting the whole expansion as exponential function and including the factor $e^{-\frac{\xi^2}{2}}$, we arrive at the proposed result of Eq. (5.143).

Probability distribution of the wave packet: As last computation in this section we want to study the probability density of the wave function of the coherent state, thus we take the modulus squared of expression (5.143)

$$\begin{aligned}
|\phi_\alpha(t,x)|^2 &= \phi_\alpha^*\,\phi_\alpha(t,x) \\
&= \frac{1}{x_0\,\sqrt{\pi}}\; e^{-\xi^2 + \sqrt{2}\,\overbrace{(\alpha^* \,+\, \alpha)}^{2\mathrm{Re}(\alpha(t))}\,\xi \,-\, \mathrm{Re}(\alpha(t))\,\overbrace{(\alpha^* \,+\, \alpha)}^{2\mathrm{Re}(\alpha(t))}} \\
&= \frac{1}{x_0\,\sqrt{\pi}}\; e^{-\left(\xi \,-\, \sqrt{2}\,\mathrm{Re}(\alpha(t))\right)^2} \,=\, \frac{1}{x_0\,\sqrt{\pi}}\; e^{-\left(\xi \,-\, \sqrt{2}\,|\alpha|\,\cos(\omega t \,-\, \varphi)\right)^2}\ .
\end{aligned} \tag{5.148}$$

Recalling the result for the expectation value (the mean value) of the position (Eq. (5.134)) and our choice of $\xi = \frac{x}{x_0}$ we finally find the probability density

$$|\phi_\alpha(t,x)|^2 \,=\, \frac{1}{x_0\sqrt{\pi}}\ \exp\left(-\frac{(x \,-\, \langle\, x\,\rangle\,(t))^2}{x_0^2}\right)\ , \tag{5.149}$$

which is a Gaussian distribution with constant width.

Result: The coherent states are oscillating Gaussian wave packets with constant width in a harmonic oscillator potential, i.e., the wave packet of the coherent state is not spreading (because all terms in the expansion are in phase). It is a wave packet with minimal uncertainty. These properties make the coherent states the closest quantum mechanical analogue to the free classical single mode field. For an illustration see Fig. 5.3.

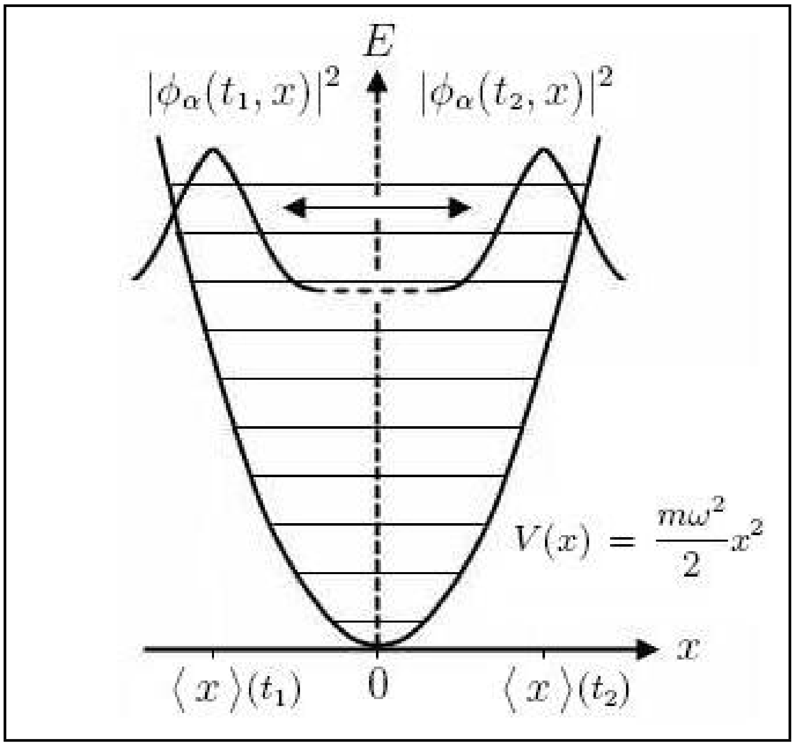

Figure 5.3: Coherent state: The probability density of the coherent state is a Gaussian distribution, whose center oscillates in a harmonic oscillator potential. Due to its being a superposition of harmonic oscillator states, the coherent state energy is not restricted to the energy levels $\hbar\omega(n+\frac{1}{2})$ but can have any value (greater than the zero point energy).

Chapter 6

The 3–dimensional Schrödinger Equation

6.1 Angular Momentum

To study how angular momentum is represented in quantum mechanics we start by reviewing the classical vector of orbital angular momentum

$$\vec{L} \,=\, \vec{x} \,\times\, \vec{p}, \tag{6.1}$$

or in components

$$L_i \,=\, \epsilon_{ijk}\, x_j\, p_k\,. \tag{6.2}$$

To make the transition to quantum mechanics we replace the classical position and momentum variables with their quantum mechanical operator counterparts[1]

$$x_i \,\rightarrow\, X_i \equiv \hat{x}_i \equiv x_i \quad , \quad p_i \,\rightarrow\, P_i \equiv \hat{p}_i \equiv p_i \,=\, -i\hbar\,\nabla_i\,. \tag{6.3}$$

Then the **commutation relations** for angular momentum follow directly

$$[\, L_i\,,\, L_j\,] \;=\; i\hbar\,\epsilon_{ijk}\,L_k \tag{6.4}$$

$$[\, L_i\,,\, x_j\,] \;=\; i\hbar\,\epsilon_{ijk}\,x_k \tag{6.5}$$

$$[\, L_i\,,\, p_j\,] \;=\; i\hbar\,\epsilon_{ijk}\,p_k\,, \tag{6.6}$$

where Eq. (6.4) is known as the *Lie algebra*, in this case of the group $SO(3)$, the rotation group in three dimensions.

---

[1] The notations to distinguish between operators and coordinates (or functions) are numerous, often capital letters, e.g. P instead of p, or hats, e.g. $\hat{x}$ instead of x, are used. Some authors however do not make a notational difference at all, but leave the distinction to the context of application. To simplify the notation we will follow this last rule from now on.

Theorem 6.1 *The orbital angular momentum is the generator of rotations in Hilbert space, i.e. $\exists$ unitary operator U, such that*

$$U(\varphi) = e^{\frac{i}{\hbar}\vec{\varphi}\vec{L}} \doteq \mathbb{1} + \tfrac{i}{\hbar}\vec{\varphi}\vec{L}$$

is generating infinitesimal rotations for small φ.

This means that the unitary operator $U(\varphi)$ rotates a given state vector $\psi(\vec{x}\,)$, which we can easily calculate using our definitions

$$U(\vec{\varphi}\,)\,\psi(\vec{x}\,) = \Big(\mathbb{1} + \tfrac{i}{\hbar}\vec{\varphi}\vec{L}\Big)\,\psi(\vec{x}\,) = \Big(\mathbb{1} + \tfrac{i}{\hbar}\vec{\varphi}\,(\vec{x} \times \tfrac{\hbar}{i}\vec{\nabla})\Big)\,\psi(\vec{x}\,)\,. \tag{6.7}$$

By rearranging the triple product (remember $\vec{a}\cdot(\vec{b}\times\vec{c}) = (\vec{a}\times\vec{b})\cdot\vec{c}$) we then get

$$U(\vec{\varphi}\,)\,\psi(\vec{x}\,) = \Big(\mathbb{1} + (\vec{\varphi}\times\vec{x}\,)\,\vec{\nabla}\Big)\,\psi(\vec{x}\,)\,, \tag{6.8}$$

which basically is the Taylor expansion at the rotated vector $\vec{x}\,' = \vec{x} + (\vec{\varphi}\times\vec{x}\,)$, visualized in Fig. 6.1

$$U(\vec{\varphi}\,)\,\psi(\vec{x}\,) = \psi(\vec{x} + (\vec{\varphi}\times\vec{x}\,)\,) = \psi(\vec{x}\,'\,)\,. \tag{6.9}$$

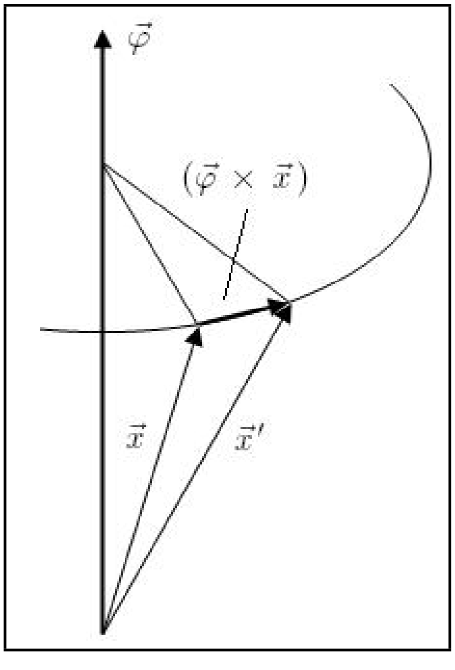

Figure 6.1: Rotation around the axis $\vec{\varphi}$ by an angle $|\vec{\varphi}|$. The vector $\vec{x}$ is rotated into $\vec{x}\,' = \vec{x} + (\vec{\varphi}\times\vec{x}\,)$ by addition of the vector $(\vec{\varphi}\times\vec{x}\,)$.

Operators in the rotated system: Since we now know how the states themselves are affected by rotations, we need to consider how the operators are changed by such operations. Let us therefore consider some (linear) operator A acting on our state $\psi(\vec{x})$ and then apply a unitary (rotation) operator U

$$A\,\psi(\vec{x}\,) \,=\, \phi(\vec{x}\,) \quad \rightarrow \quad U\,A\,\psi(\vec{x}\,) \,=\, U\,\phi(\vec{x}\,)\,. \tag{6.10}$$

We already know from Eq. (6.9) how the state vector is rotated. By further using the unitarity property $U^\dagger\,U = \mathbb{1}$ we get

$$U\,A\,\mathbb{1}\,\psi(\vec{x}\,) \,=\, U\,A\,U^\dagger\,\underbrace{U\,\psi(\vec{x}\,)}_{\psi(\vec{x}\,'\,)} \,=\, U\,\phi(\vec{x}\,) \,=\, \phi(\vec{x}\,'\,)\,, \tag{6.11}$$

and we can at last define the operator in the rotated system as

$$A' \,=\, U\,A\,U^\dagger\,, \tag{6.12}$$

which, by using Theorem 6.1, we can rewrite for infinitesimal rotations as

$$A' \,\doteq\, A \,+\, \tfrac{i}{\hbar}\,\vec{\varphi}\,\Big[\,\vec{L}\,,\,A\,\Big]\,. \tag{6.13}$$

Examples: Let us next study two special cases of operators that take on a simple form under rotations.

I) Let A be a scalar, rotation invariant operator, then the commutator with the angular momentum operator vanishes and the operator remains unchanged

$$\Big[\,\vec{L}\,,\,A\,\Big] \,=\, 0 \quad \rightarrow \quad A \,=\, A'\,. \tag{6.14}$$

Examples for such operators are $\vec{p}^{\,2}$, $\vec{L}^2$ and H.

II) Let A be a vector-valued operator, then the commutator with the angular momentum operator is given by the commutation relations of Eq. (6.4) - Eq. (6.6) and the operator is rotated exactly like a vector in 3 dimensions, see Fig. 6.1,

$$A_i' \,=\, A_i \,+\, \tfrac{i}{\hbar}\,\varphi_j\,[\,L_j\,,\,A_i\,] \,=\, A_i \,+\, \tfrac{i}{\hbar}\,\varphi_j\,i\hbar\,\epsilon_{jik}\,A_k \,=\, A_i \,+\, \epsilon_{ijk}\,\varphi_j\,A_k\,, \tag{6.15}$$

resulting in

$$\vec{A}' \,=\, \vec{A} \,+\, \vec{\varphi}\,\times\,\vec{A}\,. \tag{6.16}$$

Examples for this class of operators are $\vec{L}$, $\vec{x}$ and $\vec{p}$.

6.2 Angular Momentum Eigenvalues

Studying again the ***Lie algebra of the rotation group SO(3)***

$$[\, L_i \,,\, L_j \,] \;=\; i\hbar\, \epsilon_{ijk}\, L_k \,, \tag{6.17}$$

we can conclude that different components of the angular momentum operator, e.g. L_{x} and L_{y}, do not commute with each other, i.e. they are incompatible. By remembering the general uncertainty relation for operators A and B from Theorem 2.4, we can immediately deduce the uncertainty relation for angular momentum

$$\Delta L_{\mathrm{x}} \,\cdot\, \Delta L_{\mathrm{y}} \;\geq\; \tfrac{\hbar}{2}\, |\, \langle\, L_{\mathrm{z}} \,\rangle \,| \,. \tag{6.18}$$

In the context of an experiment this translates to the statement that different components of angular momentum observables can not be measured "simultaneously". In a theoretical framework this is expressed in the fact that these operators do not have common eigenfunctions.

We therefore need to find operators, that do have common eigenfunctions, i.e. which commute with each other. This requirement is fulfilled for $\vec{L}^2$ and any angular momentum component L_i, since

$$\left[\, \vec{L}^2 \,,\, L_i \,\right] \;=\; 0 \quad \text{where} \quad i = 1, 2, 3 \,. \tag{6.19}$$

The next task will be to find the eigenfunctions and eigenvalues of $\vec{L}^2$ and one of the components, w.l.o.g. L_{z}. We start with the ansatz

$$\vec{L}^2 f \;=\; \lambda\, f \quad , \quad L_{\mathrm{z}}\, f \;=\; \mu\, f \,, \tag{6.20}$$

and use the technique of ladder operators which we define as:

Definition 6.1 *The angular momentum* ***ladder operators*** $L_\pm$ *are defined by*

$$L_\pm \;:=\; L_{\mathrm{x}} \,\pm\, i\, L_{\mathrm{y}}$$

where L_+ *is called the raising operator, and*
L_- *is called the lowering operator.*

With this definition and help of Eq. (6.17) one can easily rewrite the ***Lie algebra of the angular momentum operators***

$$[\; L_{\mathrm{z}} \,,\, L_\pm \,] \;=\; \pm\, \hbar\, L_\pm \tag{6.21}$$

$$[\; \vec{L}^2 \,,\, L_\pm \,] \;=\; 0 \tag{6.22}$$

$$[\, L_+ \,,\, L_- \,] \;=\; 2\, \hbar\, L_{\mathrm{z}} \,. \tag{6.23}$$

Lemma 6.1 *If f is an eigenfunction of $\vec{L}^2$ and L_{z} then the function $L_\pm\, f$ is as well.*

Proof: We do the proof in two steps, first proving that $L_\pm f$ is an eigenfunction of $\vec{L}^2$ followed by the proof to be an eigenfunction of L_z:

$$\vec{L}^2\, L_\pm\, f \overset{\text{Eq. (6.22)}}{=} L_\pm\, \vec{L}^2\, f \,=\, \lambda\, L_\pm\, f\,, \tag{6.24}$$

$$\begin{aligned} L_\mathrm{z}\, L_\pm\, f \;&=\; L_\mathrm{z}\, L_\pm\, f \,-\, L_\pm\, L_\mathrm{z}\, f \,+\, L_\pm\, L_\mathrm{z}\, f \,= \\ &=\; [\, L_\mathrm{z}\,,\, L_\pm\,]\; f + L_\pm\, L_\mathrm{z}\, f \overset{\text{Eq. (6.21)}}{=} \\ &=\; (\mu\,\pm\,\hbar\,)\, L_\pm\, f\;. \quad \text{q.e.d} \end{aligned} \tag{6.25}$$

We see that $L_\pm\, f$ is an eigenfunction of L_z with eigenvalue $(\mu\,\pm\,\hbar)$. Thus, starting from some eigenvalue μ the ladder operators "switch" between all possible eigenvalues of L_z, "climbing" or "descending" the eigenvalue ladder of L_+ and L_- respectively. Since the eigenvalue λ of $\vec{L}^2$ is the same for all the eigenfunctions produced by the action of the ladder operators we know that for a given value λ of $\vec{L}^2$ the z-component L_z must be bounded by the square root of λ. Thus there exists some function f_top, corresponding to the highest possible value of L_z, such that

$$L_+\, f_\text{top} \,=\, 0\,. \tag{6.26}$$

Let us further assume the eigenvalue of L_z at f_top is $\hbar l$. We then have

$$L_\mathrm{z}\, f_\text{top} \,=\, \hbar l\, f_\text{top} \quad , \quad \vec{L}^2\, f_\text{top} \,=\, \lambda\, f_\text{top}\,. \tag{6.27}$$

Before continuing we will establish an equality which will be useful in the following calculations

$$\begin{aligned} L_\pm\, L_\mp \;&=\; (L_\mathrm{x}\,\pm\, i\, L_\mathrm{y})(L_\mathrm{x}\,\mp\, i\, L_\mathrm{y}) \,=\, L_\mathrm{x}^2 \,+\, L_\mathrm{y}^2 \,\mp\, i\, \underbrace{[\, L_\mathrm{x}\,,\, L_\mathrm{y}\,]}_{i\,\hbar\, L_\mathrm{z}} \,= \\ &=\; \underbrace{L_\mathrm{x}^2 \,+\, L_\mathrm{y}^2 \,+\, L_\mathrm{z}^2}_{\vec{L}^2} - L_\mathrm{z}^2 \,\pm\, \hbar\, L_\mathrm{z} \\ \Rightarrow\;\; \vec{L}^2 \;&=\; L_\pm\, L_\mp \,+\, L_\mathrm{z}^2 \,\mp\, \hbar\, L_\mathrm{z}\;. \end{aligned} \tag{6.28}$$

We can then use Eq. (6.28) to calculate the eigenvalue of $\vec{L}^2$ in terms of the z-component eigenvalue $\hbar l$

$$\vec{L}^2\, f_\text{top} \,=\, (\, L_-\, L_+ \,+\, L_\mathrm{z}^2 \,+\, \hbar\, L_\mathrm{z})\, f_\text{top} \,=\, \hbar^2\, l(l\,+\,1)\, f_\text{top}\,, \tag{6.29}$$

where we used Eq. (6.26) and the left part of Eq. (6.27). Therefore we conclude that the angular momentum eigenvalue λ is given by

$$\lambda \,=\, \hbar^2\, l(l\,+\,1)\,, \tag{6.30}$$

where $\hbar l$ is the highest possible value of L_z.

Before we analyze this result, let us do the analogue computation for the eigenfunction f_{bottom}, corresponding to the lowest possible eigenvalue of L_{z} (for a fixed value of λ), which we assume to be $\hbar\bar{l}$. Then the eigenvalue equations are

$$L_{\text{z}}\, f_{\text{bottom}} \,=\, \hbar\bar{l}\, f_{\text{bottom}} \quad , \quad \vec{L}^2\, f_{\text{bottom}} \,=\, \lambda\, f_{\text{bottom}} \,. \tag{6.31}$$

Using again Eq. (6.28) we apply $\vec{L}^2$ to f_{bottom}

$$\vec{L}^2\, f_{\text{bottom}} \,=\, (\,L_+\,L_- \,+\, L_{\text{z}}^2 \,-\, \hbar\,L_{\text{z}})\, f_{\text{bottom}} \,=\, \hbar^2\,\bar{l}(\bar{l} \,-\, 1)\, f_{\text{bottom}} \,. \tag{6.32}$$

Since $\lambda = \hbar^2\,\bar{l}(\bar{l} \,-\, 1)$ we can equate this to the result of Eq. (6.30), which requires

$$\bar{l} \,=\, -\,l\,. \tag{6.33}$$

Result: The eigenvalues of L_{z} take on values between $-\hbar l$ and $+\hbar l$ in steps that are integer multiples of $\hbar$. Thus there is some integer N counting the steps from $-l$ to l such that

$$l \,=\, -l + N \quad \Leftrightarrow \quad l \,=\, \tfrac{N}{2}\,, \tag{6.34}$$

which implies for l to be either one of the integer values $0, 1, 2, \cdots$, or one of the half-integer values $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \cdots$. The number l appearing in the angular momentum eigenvalue equation (see Eq. (6.30)) is called the ***azimuthal quantum number***. In order not to confuse it with the possible eigenvalues for the angular momentum z-component, one denotes the eigenvalue of L_{z} by $\mu = \hbar m$, where

$$m \,=\, -l, -l+1, \cdots, 0, \cdots, l-1, +l \tag{6.35}$$

is called the ***magnetic quantum number***. For a given l there are $(2l+1)$ values for m such that $|m| \leq l$.

The eigenfunctions f turn out to be the so-called *spherical harmonics* Y_{lm}, where the index pair labels the corresponding eigenvalues of the angular momentum observables. Restating the eigenvalue equations in terms of the spherical harmonics we have

$$\vec{L}^2\,Y_{lm} \,=\, \hbar^2\,l(l\,+\,1)\,Y_{lm} \tag{6.36}$$

$$L_{\text{z}}\,Y_{lm} \,=\, \hbar\,m\;Y_{lm} \tag{6.37}$$

$$L_{\pm}\,Y_{lm} \,=\, \hbar\,\sqrt{l(l\,+\,1)\,-\,m\,(m\,\pm\,1))}\;Y_{l,m\pm1}\,, \tag{6.38}$$

where the eigenvalue of $L_{\pm}$ is obtained from the normalization condition.

Remark: If the system is in an eigenstate of L_{z} then the observables L_{x} and L_{y} have in general an uncertainty and the uncertainty relation can be easily computed since the expectation value of L_{z} is equal to its eigenvalue $\hbar m$

$$\Delta L_{\text{x}}\,\cdot\,\Delta L_{\text{y}} \,\geq\, \tfrac{\hbar}{2}\,|\,\langle\,L_{\text{z}}\,\rangle\,| \,=\, \tfrac{\hbar^2}{2}\,m\,. \tag{6.39}$$

6.3 Angular Momentum Eigenfunctions

In order to calculate the spherical harmonics, we will rewrite the angular momentum observable $\vec{L}$ in terms of spherical coordinates, see Fig. 6.2 ,

$$\begin{aligned} x &= r \sin\theta \cos\varphi \\ y &= r \sin\theta \sin\varphi \\ z &= r \cos\theta\,. \end{aligned} \tag{6.40}$$

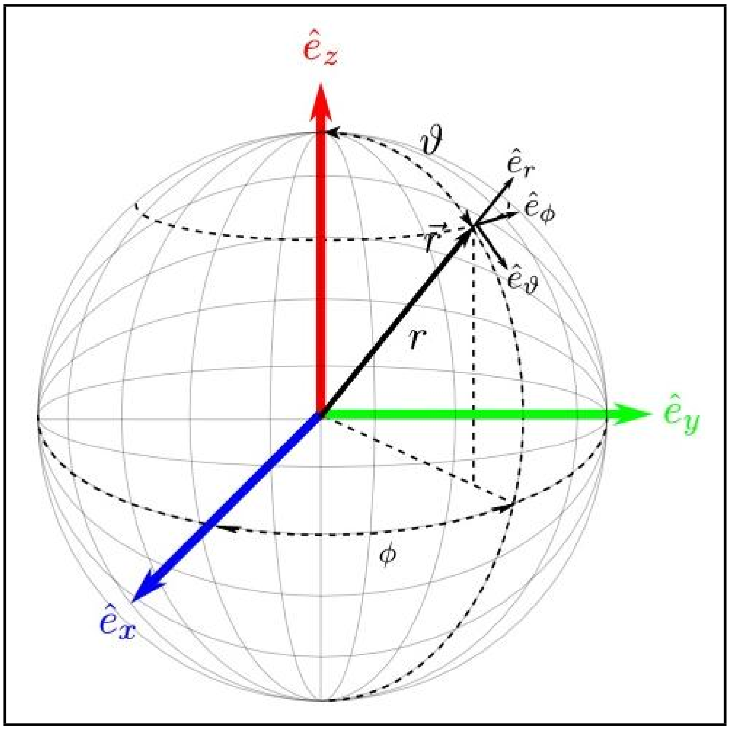

Figure 6.2: Spherical coordinates: figure from:
http://de.wikipedia.org/w/index.php?title=Bild:Kugelkoordinaten_svg.png

To this end we will need to transform the nabla-operator

$$\vec{\nabla} = \vec{e}_{\text{x}} \frac{\partial}{\partial x} + \vec{e}_{\text{y}} \frac{\partial}{\partial y} + \vec{e}_{\text{z}} \frac{\partial}{\partial z} \tag{6.41}$$

by transforming its components, the partial derivatives, as well as the basis vectors $\vec{e}_{\text{i}}$ into spherical coordinates. The partial derivatives are easily transformed using

$$\frac{\partial}{\partial x^{\text{i}}} = \frac{\partial y^{\text{j}}}{\partial x^{\text{i}}} \frac{\partial}{\partial y^{\text{j}}}\,, \tag{6.42}$$

so that we have

$$\begin{aligned}
\frac{\partial}{\partial x} &= \frac{\partial r}{\partial x}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial x}\frac{\partial}{\partial \theta} + \frac{\partial \varphi}{\partial x}\frac{\partial}{\partial \varphi} = \\
&= \sin\theta\,\cos\varphi\,\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\,\cos\varphi\,\frac{\partial}{\partial \theta} - \frac{1}{r}\frac{\sin\varphi}{\sin\theta}\frac{\partial}{\partial \varphi} \qquad (6.43)\\
\frac{\partial}{\partial y} &= \sin\theta\,\sin\varphi\,\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\,\sin\varphi\,\frac{\partial}{\partial \theta} + \frac{1}{r}\frac{\cos\varphi}{\sin\theta}\frac{\partial}{\partial \varphi} \qquad (6.44)\\
\frac{\partial}{\partial z} &= \cos\theta\,\frac{\partial}{\partial r} - \frac{1}{r}\sin\theta\,\frac{\partial}{\partial \theta}\,. \qquad (6.45)
\end{aligned}$$

For the transformation of the basis vectors, one first calculates the basis $\{\vec{e}_{\mathrm{i}}'\}$ of the new coordinate system with respect to the old basis $\{\vec{e}_{\mathrm{j}}\}$, e.g. $\vec{e}_1' = \vec{e}_{\mathrm{r}}$,

$$\vec{e}_{\mathrm{r}} = \left|\frac{\partial x^{\mathrm{i}}}{\partial r}\right|^{-1}\vec{e}_{\mathrm{j}}\,\frac{\partial x^{\mathrm{j}}}{\partial r} = \frac{\vec{e}_{\mathrm{x}}\,\frac{\partial x}{\partial r} + \vec{e}_{\mathrm{y}}\,\frac{\partial y}{\partial r} + \vec{e}_{\mathrm{z}}\,\frac{\partial z}{\partial r}}{\sqrt{(\frac{\partial x}{\partial r})^2 + (\frac{\partial y}{\partial r})^2 + (\frac{\partial z}{\partial r})^2}}\,, \tag{6.46}$$

and for the basis vectors of the spherical coordinate system with respect to the cartesian basis we get

$$\vec{e}_{\mathrm{r}} = \begin{pmatrix}\sin\theta\,\cos\varphi\\ \sin\theta\,\sin\varphi\\ \cos\theta\end{pmatrix}\quad,\quad \vec{e}_\theta = \begin{pmatrix}\cos\theta\,\cos\varphi\\ \cos\theta\,\sin\varphi\\ \sin\theta\end{pmatrix}\quad,\quad \vec{e}_\varphi = \begin{pmatrix}-\sin\varphi\\ \cos\varphi\\ 0\end{pmatrix}\,. \tag{6.47}$$

From above expressions we can easily compute the cartesian basis vectors with respect to the basis of the spherical coordinate system, yielding

$$\vec{e}_{\mathrm{x}} = \begin{pmatrix}\sin\theta\,\cos\varphi\\ \cos\theta\,\cos\varphi\\ -\sin\varphi\end{pmatrix}\quad,\quad \vec{e}_{\mathrm{y}} = \begin{pmatrix}\sin\theta\,\sin\varphi\\ \cos\theta\,\sin\varphi\\ \cos\varphi\end{pmatrix}\quad,\quad \vec{e}_{\mathrm{z}} = \begin{pmatrix}\cos\theta\\ -\sin\theta\\ 0\end{pmatrix}\,. \tag{6.48}$$

Eqs. (6.43) - (6.45) together with Eq. (6.48) are then inserted into our definition of the nabla-operator (Eq. (6.41)) to finally find

$$\vec{\nabla} = \vec{e}_{\mathrm{r}}\,\frac{\partial}{\partial r} + \vec{e}_\theta\,\frac{1}{r}\frac{\partial}{\partial \theta} + \vec{e}_\varphi\,\frac{1}{r\,\sin\theta}\frac{\partial}{\partial \varphi}\,. \tag{6.49}$$

We now want to calculate the angular momentum operator components with respect to the cartesian basis, i.e. $L_{\mathrm{x}}, L_{\mathrm{y}}$ and L_{z} , in terms of spherical coordinates. We recall that the angular momentum operator is defined as

$$\vec{L} = \vec{x}\times\vec{p} = \frac{\hbar}{i}\,\vec{x}\times\vec{\nabla}\,, \tag{6.50}$$

we use Eq. (6.49) and the cartesian representation of the r, θ, φ basis vectors (Eq. (6.47)), and additionally, by comparing Eq. (6.40) and the r-basis vector we note that the position vector $\vec{x}$ can be written as $\vec{x} = r\,\vec{e}_\mathrm{r}$. Finally[2] we find

$$\begin{aligned} L_\mathrm{x} &= \frac{\hbar}{i}\left(-\sin\varphi\,\frac{\partial}{\partial\theta} - \cos\varphi\,\cot\theta\,\frac{\partial}{\partial\varphi}\right) && (6.51)\\ L_\mathrm{y} &= \frac{\hbar}{i}\left(\cos\varphi\,\frac{\partial}{\partial\theta} - \sin\varphi\,\cot\theta\,\frac{\partial}{\partial\varphi}\right) && (6.52)\\ L_\mathrm{z} &= \frac{\hbar}{i}\,\frac{\partial}{\partial\varphi}\,. && (6.53) \end{aligned}$$

For the ladder operators (Definition 6.1) follows then in spherical coordinates

$$L_\pm \,=\, \hbar\, e^{\pm i\varphi}\left(\pm\frac{\partial}{\partial\theta} + i\,\cot\theta\,\frac{\partial}{\partial\varphi}\right)\,, \tag{6.54}$$

for the angular momentum squared

$$\vec{L}^2 \,=\, -\,\hbar^2\left(\frac{1}{\sin\theta}\,\frac{\partial}{\partial\theta}\,\sin\theta\,\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\,\frac{\partial^2}{\partial\varphi^2}\right)\,, \tag{6.55}$$

and for the eigenvalue equations (Eq. (6.36) and Eq. (6.37))

$$\left(\frac{1}{\sin\theta}\,\frac{\partial}{\partial\theta}\,\sin\theta\,\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\,\frac{\partial^2}{\partial\varphi^2}\right) Y_{lm} \,=\, -\,l(l+1)\,Y_{lm} \tag{6.56}$$

$$\frac{\partial}{\partial\varphi}\,Y_{lm} \,=\, i\,m\,Y_{lm}\,. \tag{6.57}$$

In order to solve this partial differential equations we use a **separation ansatz**, the same method as in Section 4.1, Eq. (4.2),

$$Y_{lm}(\theta,\varphi) \,=\, P(\theta)\,\Phi(\varphi)\,. \tag{6.58}$$

Inserting this ansatz into Eq. (6.57) we immediately find

$$\Phi(\varphi) \,=\, e^{i\,m\,\varphi}\,. \tag{6.59}$$

Furthermore, from the continuity of the wave function we conclude that $\Phi(\varphi + 2\pi) = \Phi(\varphi)$, which restricts the magnetic quantum number m (and thus also l) to be an integer number (eliminating the possibility for half-integer numbers), i.e.,

$$\begin{aligned} l &= 0, 1, 2, \cdots && (6.60)\\ m &= -l, -l+1, \cdots, 0, \cdots, l-1, +l\,. && (6.61) \end{aligned}$$

---

[2] We could have obtained this result much faster by just transforming the partial derivatives and expressing $\vec{x}$ by r, θ and φ, but we wanted to explicitly present the apparatus of spherical coordinates, especially the form of the nabla operator and how one can calculate it.

Thus the spherical harmonics we want to calculate from Eq. (6.56) are of the form

$$\Rightarrow \quad Y_{lm}(\theta,\varphi) \,=\, e^{i\,m\,\varphi}\, P(\theta)\ . \tag{6.62}$$

Inserting expression (6.62) into Eq. (6.56) we get

$$\left(\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \sin\theta \frac{\partial}{\partial\theta} \,-\, \frac{m^2}{\sin^2\theta} \,+\, l(l\,+\,1) \right) P(\theta) \,=\, 0\ , \tag{6.63}$$

and with a change of variables

$$\xi \,=\, \cos\theta \quad \Rightarrow \quad \frac{\partial}{\partial\theta} \,=\, \frac{\partial\xi}{\partial\theta}\frac{\partial}{\partial\xi} \,=\, -\sin\theta \frac{\partial}{\partial\xi}\ , \tag{6.64}$$

Eq. (6.63) thereupon becomes

$$\left(\left(1\,-\,\xi^2\right) \frac{d^2}{d\xi^2} \,-\, 2\,\xi\,\frac{d}{d\xi} \,+\, l(l\,+\,1) \,-\, \frac{m^2}{1\,-\,\xi^2} \right) P_{l\mathrm{m}}(\xi) \,=\, 0\ , \tag{6.65}$$

which is the ***general Legendre equation***. It is solved by the *associated Legendre functions*, also called associated Legendre polynomials[3], which can be calculated from the (ordinary) Legendre polynomials P_l via the formula

$$P_{l\mathrm{m}}(\xi) \,=\, \left(1\,-\,\xi^2\right)^{\frac{m}{2}} \frac{d^m}{d\xi^m} P_l\ . \tag{6.66}$$

The Legendre polynomials P_l, on the other hand, can be calculated from ***Rodrigues' formula***

$$P_l(\xi) \,=\, \frac{1}{2^l\, l!} \frac{d^l}{d\xi^l} \left(\xi^2\,-\,1\right)^l\ . \tag{6.67}$$

For $m = 0$ the associated functions are just the Legendre polynomials themselves. The P_l are polynomials of the order l, while the $P_{l\mathrm{m}}$ are of order $(l - m)$ in ξ, multiplied with a factor $\sqrt{(1-\xi^2)^m} = \sin^m\theta$, and have $(l - m)$ roots in the interval $(-1, 1)$. To give an impression about their form we will write down some examples for (associated) Legendre polynomials

$$\begin{aligned} &P_0 \,=\, 1\,, \quad P_1 \,=\, \cos\theta\,, \quad P_2 \,=\, \tfrac{1}{2}\,(3\,\cos^2\theta\,-\,1) \\ &P_{1,0} \,=\, \cos\theta\,, \quad P_{1,1} \,=\, \sin\theta\ . \end{aligned} \tag{6.68}$$

For the normalized (to 1) spherical harmonics in terms of θ and φ we then find the following expression

$$Y_{lm}(\theta,\varphi) \,=\, (-1)^{\frac{1}{2}(m\,+\,|m|)} \left[\frac{2l\,+\,1}{4\pi} \frac{(\,l\,-\,|m|\,)!}{(\,l\,+\,|m|\,)!} \right]^{\frac{1}{2}} e^{i\,m\,\varphi}\, P_{l|\mathrm{m}|}(\cos\theta)\ . \tag{6.69}$$

---

[3] Technically, they are only polynomials if m is even, but the term became accustomed and is also used for odd values of m.

These are the eigenfunctions of $\vec{L}^2$ and L_{z}, which satisfy

$$Y_{l,-m} \,=\, (-1)^m \, Y^*_{lm} \,, \tag{6.70}$$

as well as

$$Y_{l0}\,(\theta,\varphi) \,=\, \sqrt{\frac{2l+1}{4\pi}}\, P_l(\cos\theta) \,. \tag{6.71}$$

Some examples for spherical harmonics are

$$Y_{0\,0} \,=\, \left(\frac{1}{4\pi}\right)^{\frac{1}{2}} \,, \quad Y_{1\,0} \,=\, \left(\frac{3}{4\pi}\right)^{\frac{1}{2}} \cos\theta \,, \quad Y_{1\,\pm1} \,=\, \mp \left(\frac{3}{8\pi}\right)^{\frac{1}{2}} \sin\theta \; e^{\pm i\varphi} \,. \tag{6.72}$$

The spherical harmonics as well as the (associated) Legendre polynomials form orthogonal[4] systems; the (associated) Legendre polynomials on the interval $[-1,+1]$ and the spherical harmonics on the unit sphere. This means that all (square integrable) functions can be expanded in terms of these special functions in the respective regions. The orthogonality relations are written as

$$\int\limits_{4\pi} d\Omega\, Y^*_{lm}\,(\theta,\varphi)\, Y_{l'\,m'}\,(\theta,\varphi) \,=\, \delta_{ll'}\,\delta_{mm'} \,, \tag{6.73}$$

where the integration is carried out over the unit sphere, which is denoted by the value 4π of the full solid angle Ω, see also Eq. (6.91), and

$$\int\limits_{-1}^{+1} d\xi\, P_{l\mathrm{m}}(\xi)\, P_{l'\mathrm{m}}(\xi) \,=\, \frac{2}{2l+1}\, \frac{(l+m)!}{(l-m)!}\, \delta_{ll'} \,. \tag{6.74}$$

This last equation reduces to the orthogonality relation for the Legendre polynomials if we set $m=0$

$$\int\limits_{-1}^{+1} d\xi\, P_l(\xi)\, P_{l'}(\xi) \,=\, \frac{2}{2l+1}\, \delta_{ll'} \,. \tag{6.75}$$

The spherical harmonics additionally form a complete system with completeness relation

$$\sum_{m=-l}^{+l} \sum_{l=0}^{\infty} Y^*_{lm}\,(\theta,\varphi)\, Y_{lm}\,(\theta',\varphi') \,=\, \frac{1}{\sin\theta}\, \delta(\theta\,-\,\theta')\, \delta(\varphi\,-\,\varphi') \,. \tag{6.76}$$

6.4 The 3–dimensional Schrödinger Equation

With the knowledge about the (orbital) angular momentum operator $\vec{L}$ from the previous sections we now want to solve the time-independent Schrödinger equation in 3 dimensions for a potential V, which only depends on r, i.e. $V = V(r)$,

$$\left(-\,\frac{\hbar^2}{2m}\,\Delta \,+\, V(r)\right)\, \psi(\vec{x}) \,=\, E\,\psi(\vec{x}) \,. \tag{6.77}$$

---

[4] The spherical harmonics are also complete and normalized to one.

The operator Δ is the Laplacian (or Laplace-operator)

$$\Delta \,=\, \vec{\nabla}^2 \,=\, \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\,, \tag{6.78}$$

which we transform into spherical coordinates by inserting Eqs. (6.43) - (6.45) into expression (6.78) and differentiating according to Leibniz's law. This leads us to

$$\Delta \,=\, \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\,. \tag{6.79}$$

Comparing expression (6.79) to the squared angular momentum operator in spherical coordinates (Eq. (6.55)), we recognize that we can rewrite the kinetic part of the Hamiltonian of Eq. (6.77) as

$$-\frac{\hbar^2}{2m}\Delta \,=\, -\frac{\hbar^2}{2m}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\vec{L}^2}{2mr^2}\,. \tag{6.80}$$

The Schrödinger equation is then of the form

$$\left(-\frac{\hbar^2}{2m}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\vec{L}^2}{2mr^2} + V(r)\right)\psi(\vec{x}) \,=\, E\,\psi(\vec{x})\,, \tag{6.81}$$

which we solve by a separation ansatz into a function only depending on the radius r, the so called *radial wave function* $R(r)$, and a function comprising of the angle- dependencies of the wave function

$$\psi(r,\theta,\varphi) \,=\, R(r)\,Y_{lm}(\theta,\varphi)\,. \tag{6.82}$$

Since the spherical harmonics are eigenfunctions of the squared angular momentum operator, see Eq. (6.36), we can rewrite the Schrödinger equation as an equation just for its radial component $R(r)$

$$\left(-\frac{\hbar^2}{2m}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\hbar^2\,l(l+1)}{2mr^2} + V(r)\right)R_l(r) \,=\, E\,R_l(r)\,. \tag{6.83}$$

This equation does still depend on the azimuthal quantum number l therefore we assign a corresponding label to the radial wave function. To simplify the equation we follow three simple steps, the first one is to rewrite the differential operator on the left side

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) \,=\, \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\,. \tag{6.84}$$

Next we replace the function $R(r)$ by the so-called ***reduced wave function*** $u(r)$, which is defined by:

Definition 6.2 *The* ***reduced wave function*** $u_l(r)$ *is given by*

$$u_l \,:=\, r\,R_l(r)\,.$$

Differentiating the reduced wave function we get

$$u' \;=\; (r\,R)' \;=\; R + r\,R' \tag{6.85}$$

$$u'' \;=\; 2\,R' + r\,R'' \tag{6.86}$$

$$\frac{u''}{r} \;=\; R'' + \frac{2}{r}\,R'\;. \tag{6.87}$$

Comparing this result to the one of Eq. (6.84) the Schrödinger equation (6.83), when multiplied by r can be reformulated for the reduced wave function

$$\left(-\,\frac{\hbar^2}{2m}\,\frac{\partial^2}{\partial r^2} \,+\, \frac{\hbar^2\,l(l+1)}{2mr^2} \,+\, V(r)\right)\,u_l(r) \;=\; E\,u_l(r)\;. \tag{6.88}$$

The third step is to define an effective potential $V_{\text{eff.}}(r)$ to bring the equation to a more appealing form.

Definition 6.3 *The* ***effective Potential*** $V_{\text{eff.}}(r)$ *is given by*

$$V_{\text{eff.}}(r) \;=\; \tfrac{\hbar^2\,l(l+1)}{2mr^2} \,+\, V(r)\;.$$

The Schrödinger equation in 3 dimensions, a second order partial differential equation, has thus become a 1–dimensional ordinary differential equation for the reduced wave function, which is of the same form as the one-dimensional Schrödinger equation, except from the term $\hbar^2\,l(l+1)/2mr^2$ in the effective potential. This additional term is a repulsive potential, called *centrifugal term*, in analogy to the classical (pseudo) force

$$u_l''(r) \,+\, \frac{2m}{\hbar^2}\,\left(E_l \,-\, V_{\text{eff.}}(r)\right)\,u_l(r) \;=\; 0\;. \tag{6.89}$$

Normalization: Let us now study how the normalization condition of the total wave function influences the current situation. We can write the normalization condition as

$$\int\limits_{4\pi} d\Omega \int\limits_0^\infty dr\,r^2\,|\psi(r,\theta,\varphi)|^2 \;=\; 1\;, \tag{6.90}$$

where again the integration over the whole solid angle has to be understood as

$$\int\limits_{4\pi} d\Omega \;=\; \int\limits_0^{2\pi} d\varphi \int\limits_0^{\pi} d\theta\,\sin\,\theta\;. \tag{6.91}$$

Because the spherical harmonics are normalized, see Eq. (6.73), we can easily transform our normalization condition to

$$\int_0^\infty dr\, r^2\, |R_l(r)|^2 \;=\; 1\;, \tag{6.92}$$

which translates to the reduced wave function as

$$\int_0^\infty dr\, |u_l(r)|^2 \;=\; 1\;. \tag{6.93}$$

The normalizability of the radial wave function $R(r) = \frac{u(r)}{r}$ restricts it to be bounded at the origin, which subsequently requires the reduced wave function to vanish at the origin

$$u_l(r=0) \;=\; 0\,. \tag{6.94}$$

This constraint imposes the character of an odd function, sketched in Fig. 6.3, on the reduced wave function. The ground state of the 3–dimensional problem thus corresponds to the first excited state of the 1–dimensional problem, and is therefore not available for potentials of arbitrary strength, as already mentioned in Section 4.5.1 (see also Fig. 4.5 and the corresponding remark).

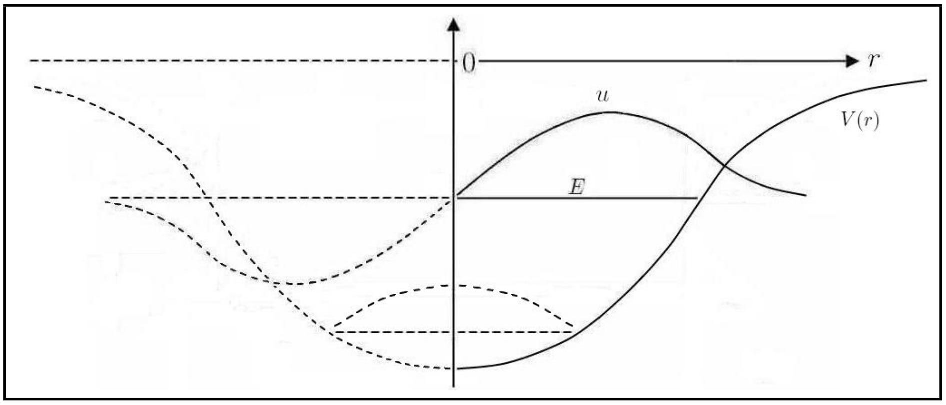

Figure 6.3: Bound states in 3 dimensions: The 3–dim. problem can be reduced to a 1–dim problem in r, but the reduced wave function u of the ground state then is an odd function, which does not allow for a bound state for arbitrarily weak potentials $V(r)$.

Theorem 6.2 *In a 3–dimensional problem, the potential must exhibit a minimal strength for the ground state to exist. (without proof)*

Furthermore, if there is no ground state for azimuthal quantum number $l = 0$, then there won't be a ground state (in this problem) for any other $l \neq 0$ either.

6.5 Coulomb Potential and H–Atom

We now want to use the three–dimensional Schrödinger equation to calculate the energy levels of the simplest atomic system, the hydrogen atom, consisting only of an electron of mass m_e and a proton, which is considered as infinitely massive in this approximation (though to introduce a reduced mass would be straightforward). Furthermore, the spin of the particles is not considered. In order to proceed we simply insert for the attraction of the electron the ***Coulomb potential*** into the term $V(r)$

$$V(r) \,=\, -\frac{q^2}{4\pi\epsilon_0 r} \,=\, -\frac{e^2}{r} \,. \tag{6.95}$$

We then introduce constants that are typical characteristics of such a system to further simplify and analyze the problem at hand. The first of which will be the *fine structure constant* α

$$\alpha \,=\, \frac{e^2}{\hbar c} \,=\, \frac{q^2}{4\pi\epsilon_0 \hbar c} \,\approx\, \frac{1}{137} \,. \tag{6.96}$$

The fine structure constant is a fundamental constant, whose numerical value does not depend on the choice of units, and which therefore can be displayed in numerous ways, e.g. in units of $\hbar = c = 1$, $\alpha \,=\, \frac{q^2}{4\pi\epsilon_0}$ and thus the Coulomb potential is $V(r) \,=\, -\,\alpha/r$. This is also the reason for its many different physical interpretations, the most common of which is as the coupling constant of the electromagnetic interaction, i.e. the strength at which photons couple to the electric charge. Another interpretation, relevant for our calculations, is that the fine structure constant can be seen as the ratio of the typical electron velocity in a bound state of the hydrogen atom to the speed of light. Viewing it that way, the smallness of α gives us the justification for neglecting effects predicted by the theory of special relativity.

The characteristic length unit of the hydrogen atom is the *Bohr radius* r_B

$$r_\mathrm{B} \,=\, \frac{\hbar^2}{m_\mathrm{e}\, e^2} \,=\, \frac{1}{\alpha}\,\frac{\hbar}{m_\mathrm{e}\, c} \,=\, \frac{1}{\alpha}\,\lambda_\mathrm{C} \,\approx\, 0,53\,\mathring{A} \,, \tag{6.97}$$

where λ_C is the Compton wavelength of the electron. Finally, we can construct a constant with dimension of energy out of the constants already used

$$E_\mathrm{I} \,=\, \frac{m_\mathrm{e}\, e^4}{2\hbar^2} \,=\, \tfrac{1}{2}\, m_\mathrm{e}\, c^2\, \alpha^2 \,\approx\, 13,6\, eV \,, \tag{6.98}$$

which is the *ionization energy* of the hydrogen atom. With this knowledge we now easily reformulate the Schrödinger equation in terms of dimensionless variables

$$\rho \,=\, \frac{r}{r_\mathrm{B}} \,, \quad \epsilon \,=\, -\,\frac{E}{E_\mathrm{I}} \,, \tag{6.99}$$

which yields the equation

$$\left(\frac{1}{\rho} \frac{d^2}{d\rho^2} \rho \, - \, \frac{l(l+1)}{\rho^2} \, + \frac{2}{\rho} \, - \, \epsilon \right) R_l(\rho) \; = \; 0 \; . \tag{6.100}$$

This differential equation can be solved by the ansatz

$$R_l(\rho) \; = \; e^{-\sqrt{\epsilon}\rho} \, \rho^{l+1} \, Q_{n',l}(\rho) \; . \tag{6.101}$$

It will provide us an equation for the yet unknown functions $Q_{n',l}(\rho)$, which we express as a power series. The series has to terminate at a finite degree for reasons of normalizability, resulting such in a polynomial of degree n', the *radial quantum number*. These polynomials are called the associated Laguerre polynomials. The normalizable solutions (Eq. (6.101)) correspond to particular values of ϵ

$$\epsilon \; = \; \frac{1}{(n' \, + \, l + 1)^2} \; = \; \frac{1}{n^2} \; . \tag{6.102}$$

By changing the notation to a new label n, the *principal quantum number*, we find that ϵ_{n} is an eigenvalue for all radial equations where $l < n$. We thus find that the solutions of the 3–dimensional Schrödinger equation are labeled by three integer quantum numbers: n, l and m, such that

$$n \; = \; 1, 2, \cdots \quad , \quad l \, = \, 0, 1, \cdots, n-1 \quad , \quad m \; = \; -l, \, \cdots, +l \, . \tag{6.103}$$

The energy solutions depend only[5] on the principal quantum number

$$E_{\mathrm{n}} \; = \; -\epsilon \, E_{\mathrm{I}} \; = \; - \, \frac{E_{\mathrm{I}}}{n^2} \; , \tag{6.104}$$

which, by reinserting our constant from Eq. (6.98), provides the so called *Bohr formula*

$$E_{\mathrm{n}} \; = \; - \, \frac{m_{\mathrm{e}} \, e^4}{2\hbar^2 n^2} \; . \tag{6.105}$$

To every energy level correspond different values of angular momentum, the degeneracy in terms of the principal quantum number is

$$\sum_{l=0}^{n-1} (2l \, + \, 1) \; = \; n^2 \; , \tag{6.106}$$

since for a given n there are $(n-1)$ possible values of l, each one allows for $(2l+1)$ different values of m. Finally, the solutions of the Schrödinger equation can be written in terms of the characteristic constants of our problem and the quantum numbers as

$$\psi_{\mathrm{n},l,\mathrm{m}}(r, \theta, \varphi) \; = \; Y_{lm} \, (\theta, \varphi) \, e^{-\, r / n \, r_{\mathrm{B}}} \left(\frac{r}{r_{\mathrm{B}}} \right)^l Q_{\mathrm{n}-l-1}(\frac{r}{r_{\mathrm{B}}}) \; . \tag{6.107}$$

---

[5] This is, of course, only true in our approximation, where we neglected any special relativistic influences and spin.

Remark I: The fact that there is degeneracy with respect to the angular momentum is an interesting property of the $1/r$ potential. It hints at an additional symmetry other than the rotational invariance. This symmetry also has a counterpart in the classical theory, where it is the Lenz vector that is a constant of motion in this kind of potential.

Remark II: The treatment of the hydrogen atom has been totally nonrelativistic in this section. When we use the correct relativistic description, such as the relativistic energy–momentum relation and the spin, we find an additional splitting of the now degenerate levels. Such effects include the 2 possible spin levels for each set of quantum numbers n, l and m, as well as interactions of the magnetic moment, caused by the motion of the electron, which are called spin-orbit-interactions and that lead to the *fine structure* of the hydrogen atom. When we furthermore include the interaction of the electron with the spin of the proton, we get the so-called *hyperfine structure*. All these effects can be better understood in terms of the Dirac equation, which is the relativistic counterpart of the Schrödinger equation, though they can be treated with perturbative methods to get satisfying results (see Chapter 8). With such perturbative methods we can also treat the influences of exterior fields which cause the Zeeman–effect for magnetic and the Stark–effect for electric fields.

Chapter 7

Spin and Spin–Addition

7.1 Stern-Gerlach Experiment – Electron Spin

In 1922, at a time, the hydrogen atom was thought to be understood completely in terms of Bohr's atom model, two assistants at the University of Frankfurt, Otto Stern and Walther Gerlach, performed an experiment which showed that the electrons carry some intrinsic angular momentum, the spin, which is quantized in two distinct levels. This was one of the most important experiments done in the twentieth century, as its consequences allowed for many interesting experimental and theoretical applications.

7.1.1 Electron in a Magnetic Field

To fully understand the concept of spin, we start by reviewing the some properties of a classical charged particle rotating about its own symmetry axis. The angular momentum due to this rotation, let us call it $\vec{S}$, will create a magnetic dipole moment $\vec{\mu}$, proportional to the angular momentum

$$\vec{\mu} \,=\, \gamma\,\vec{S} \,=\, g\,\frac{q}{2\,m\,c}\,\vec{S}\,, \tag{7.1}$$

where γ is the *gyromagnetic ratio* and g is just called *g-factor*. For the electron we have

$$\vec{\mu} \,=\, \gamma_{\rm e}\,\vec{S} \,=\, -\,g_{\rm e}\,\frac{e}{2\,m_{\rm e}\,c}\,\vec{S} \,=\, -\,g_{\rm e}\,\frac{\mu_{\rm B}}{\hbar}\,\vec{S}\,. \tag{7.2}$$

The g-factor of the electron equals two, $g_{\rm e} = 2$, although for a classical angular momentum, it should be equal to 1. The fact, that the spin of the electron contributes twice as strong to the magnetic moment as its orbital angular momentum, is called the *anomalous magnetic moment* of the electron. The constant $\mu_{\rm B}$ is known as *Bohr's magneton*[1].

If such a magnetic dipole is subject to an external (homogeneous) magnetic field $\vec{B}$ it starts to precess, due to the torque $\vec{\Gamma}$, known as the *Larmor torque*, exerted by the magnetic field

$$\vec{\Gamma} \,=\, \vec{\mu}\,\times\,\vec{B} \,=\, \gamma\,\vec{S}\,\times\,\vec{B}\,, \tag{7.3}$$

---

[1] Bohr's magneton is here given in Gaussian CGS units. In SI-units it looks the same, but the speed of light is removed, $\mu_{\rm B} = \frac{e\hbar}{2m_{\rm e}}$.

The potential energy corresponding to this torque is given by

$$H \,=\, -\,\vec{\mu}\,\vec{B}\,. \tag{7.4}$$

Thus the Hamiltonian for a particle with spin in an exterior magnetic field of strength $\vec{B}$ is of the form

$$H \,=\, -\,\gamma\,\vec{S}\,\vec{B}\,. \tag{7.5}$$

7.1.2 Stern-Gerlach Experiment

In the Stern-Gerlach experiment silver atoms, carrying no orbital angular momentum but with a single electron opening up a new s-orbital[2] ($l = 0$), were sent through a special magnet which generates an inhomogeneous magnetic field, see Fig. 7.1. The properties of the silver atom in this state are such that the atom takes over the intrinsic angular momentum, i.e. spin, of this outermost single electron.

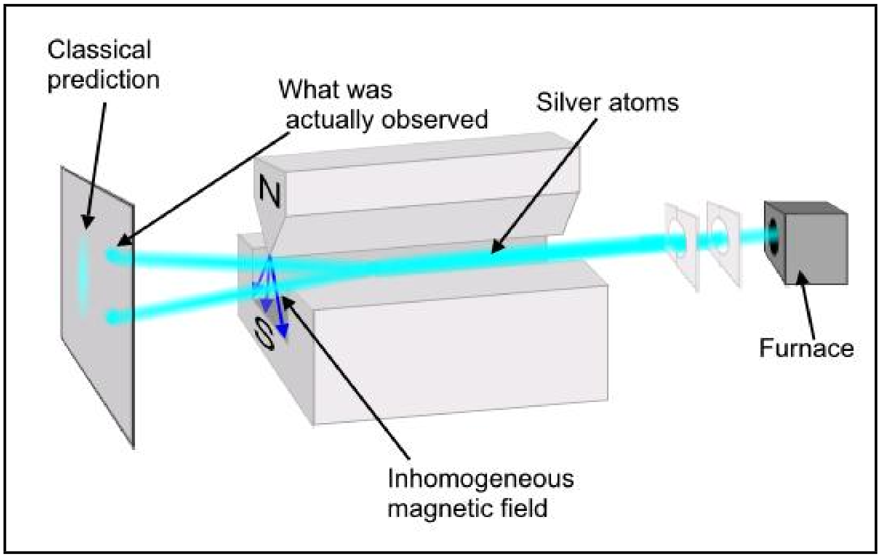

Figure 7.1: Stern-Gerlach Experiment: The inhomogeneous magnetic field exerts a force on the silver atoms, depending on the spin z-component. Classically a continuous distribution is expected, but the experiment reveals only two values of the spin z-component. Figure from: http://en.wikipedia.org/wiki/Image:Stern-Gerlach_experiment.PNG

The inhomogenity of the magnetic field causes a force $\vec{F}$ acting on the magnetic dipole in addition to the torque

$$\vec{F} \,=\, -\,\vec{\nabla}\,V \,=\, \vec{\nabla}\,(\vec{\mu}\,\vec{B})\,, \tag{7.6}$$

---

[2] This notation is often used in spectroscopy, where one labels the states of different angular momenta by s ("sharp", $l = 0$), p ("principal", $l = 1$), d ("diffuse", $l = 2$), f ("fundamental", $l = 3$) and alphabetically from there on, i.e. g,h,i,...; Every azimuthal quantum number is degenerate in the sense that it allows for $2 \cdot (2l + 1)$ bound electrons, which together are called an orbital.

where $V = H$ if the particle is at rest. Since the force depends on the value of the spin, it can be used to separate different spins. Classically, the prediction would be a continuous distribution, bounded by two values, representing spins parallel and antiparallel to the direction of the magnetic field. All the spins which are not perfectly (anti-)aligned with the magnetic field would be expected to have components in that direction that lie in between these maximal and minimal values. We can write the magnetic field as a homogeneous and an inhomogeneous part, such that it is oriented parallel to the z-axis, i.e. $B_{\mathrm{x}} = B_{\mathrm{y}=0}$,

$$\vec{B} \,=\, B_{\mathrm{z}}\,\vec{e}_{\mathrm{z}} \,=\, (B_{\mathrm{hom}} \,+\, \alpha\, z)\,\vec{e}_{\mathrm{z}}\,. \tag{7.7}$$

The force can then be expressed via the z-component of the spin $\vec{S}$, which in the quantum mechanical formalism will be an operator

$$F_{\mathrm{z}} \,=\, \alpha\,\gamma\,S_{\mathrm{z}}\,. \tag{7.8}$$

The separation of the particles with different spin then reveals experimentally the eigenvalues of this operator.

The result of the experiment shows that the particles are equally distributed among two possible values of the spin z-component, half of the particles end up at the upper spot ("spin up"), the other half at the lower spot ("spin down"). Spin is a angular momentum observable, where the degeneracy of a given eigenvalue l is $(2l+1)$. Since we observe two possible eigenvalues for the spin z-component (or any other direction chosen), see Fig. 7.2, we conclude the following value for s

$$2\,s \,+\, 1 \,=\, 2 \quad \Rightarrow \quad s \,=\, \frac{1}{2}\,. \tag{7.9}$$

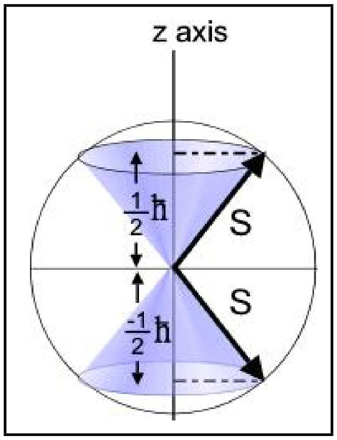

Figure 7.2: Spin $\frac{1}{2}$: The spin component in a given direction, usually the z-direction, of a spin $\frac{1}{2}$ particle is always found in either the eigenstate "↑" with eigenvalue $+\frac{1}{2}$ or "↓" with eigenvalue $-\frac{1}{2}$. Figure from: http://en.wikipedia.org/wiki/Image:Quantum_projection_of_S_onto_z_for_spin_half_particles.PNG

Result: Two additional quantum numbers are needed to characterize the nature of the electron, the *spin quantum number* s and the *magnetic spin quantum number* $m_{\mathrm{s}} = -s, \cdots, +s$. We conclude: ***spin is quantized*** and the eigenvalues of the corresponding observables are given by

$$S_{\mathrm{z}} \to \hbar\, m_{\mathrm{s}} = \pm \frac{\hbar}{2} \quad , \quad \vec{S}^2 \to \hbar^2\, s\,(s + 1) = \frac{3}{4}\,\hbar^2 \,. \tag{7.10}$$

The spin measurement is an example often used to describe a typical quantum mechanical measurement. Let us therefore elaborate this example in more detail. Consider a source emitting spin $\frac{1}{2}$ particles in an unknown spin state. The particles propagate along the y-axis and pass through a spin measurement apparatus, realized by a Stern-Gerlach magnet as described in Fig. 7.1, which is oriented along the z-axis, see Fig. 7.3.

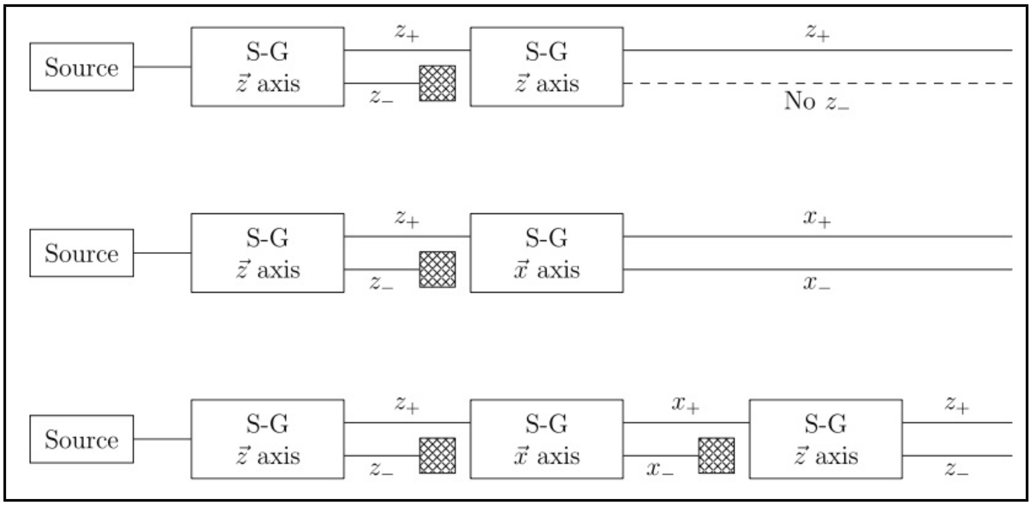

Figure 7.3: Spin $\frac{1}{2}$ measurement: Spin measurements change the state of the particles, if they are not in an eigenstate of the corresponding operator. Therefore subsequent measurements along perpendicular directions produce random results. Figure from: http://en.wikipedia.org/wiki/Image:Sg-seq.svg

All particles leaving the Stern-Gerlach apparatus are then in an eigenstate of the S_{z} operator, i.e., their spin is either “up” or “down” with respect to the z-direction. Let’s now concentrate on the “spin up” particles (in z-direction), that means we block up the “spin down” in some way, and perform another spin measurement on this part of the beam. If the second measurement is also aligned along the z-direction then all particles will provide the result “spin up”, since they are all already in an eigenstate of S_{z} (see the upper part of Fig. 7.3). The measurement of a particle being in an eigenstate of the corresponding operator leaves the state unchanged, therefore no particle will “flip” its spin.

If, however, we perform the spin measurement along a direction perpendicular to the z-axis, let’s choose the x-axis, then the results will be equally distributed among “spin up" or "spin down" in x-direction (see the middle part of Fig. 7.3). Thus, even though we knew the state of the particles beforehand, in this case the measurement resulted in a random spin flip in either of the measurement directions. Mathematically, this property is expressed by the nonvanishing of the commutator of the spin operators

$$[\, S_{\mathrm{z}} \, , \, S_{\mathrm{x}} \,] \, \neq \, 0 \, . \tag{7.11}$$

If we finally repeat the measurement along the z-direction the result will be random again (see the lower part of Fig. 7.3). We do not obtain precise information about the spin in different directions at the same time, due to the nonvanishing of the commutator (7.11) there holds an uncertainty relation for the spin observables.

Of course, we could also choose an orientation of the Stern-Gerlach magnet along some arbitrary direction. Let us assume we rotate the measurement apparatus by an angle θ (in the $z - x$ plane), then the probability P_+ to find the particle with "spin up" and P_- to find the particle with "spin down" (along this new direction) is given by

$$P_+ \, = \, \cos^2 \frac{\theta}{2} \quad \text{and} \quad P_- \, = \, \sin^2 \frac{\theta}{2} \ , \quad \text{such that} \quad P_+ \, + \, P_- \, = \, 1 \, . \tag{7.12}$$

7.2 Mathematical Formulation of Spin

Now we turn to the theoretical formulation of spin. We will describe spin by an operator, more specifically by a 2 × 2 matrix, since it has two degrees of freedom and we choose convenient matrices which are named after Wolfgang Pauli.

7.2.1 The Pauli–Matrices

The *spin observable* $\vec{S}$ is mathematically expressed by a vector whose components are matrices

$$\vec{S} \, = \, \frac{\hbar}{2} \, \vec{\sigma} \ , \tag{7.13}$$

where the vector $\vec{\sigma}$ contains the so-called ***Pauli matrices*** $\sigma_{\mathrm{x}}, \sigma_{\mathrm{y}}, \sigma_{\mathrm{z}}$:

$$\vec{\sigma} \, = \, \begin{pmatrix} \sigma_{\mathrm{x}} \\ \sigma_{\mathrm{y}} \\ \sigma_{\mathrm{z}} \end{pmatrix} \ , \ \sigma_{\mathrm{x}} \, = \, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ , \ \sigma_{\mathrm{y}} \, = \, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \ , \ \sigma_{\mathrm{z}} \, = \, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ . \tag{7.14}$$

Then the spin vector $\vec{S}$ (or the Pauli vector $\vec{\sigma}$) can be interpreted as the generator of rotations (remember Theorem 6.1) in the sense that there is a unitary operator $U(\theta)$

$$U(\theta) \, = \, e^{\frac{i}{\hbar} \vec{\theta} \vec{S}} \, = \, \mathbb{1} \, \cos \frac{\theta}{2} \, + \, i \, \hat{n} \, \vec{\sigma} \, \sin \frac{\theta}{2} \ , \tag{7.15}$$

generating rotations around the $\vec{\theta}$-axis by an angle $|\vec{\theta}|$ of the state vectors in Hilbert space. The scalar product $\vec{\theta} \, \vec{\sigma}$ is to be understood as a matrix

$$\vec{\theta} \, \vec{\sigma} \, = \, \theta_{\mathrm{x}} \, \sigma_{\mathrm{x}} \, + \, \theta_{\mathrm{y}} \, \sigma_{\mathrm{y}} \, + \, \theta_{\mathrm{z}} \, \sigma_{\mathrm{z}} \, . \tag{7.16}$$

What's very interesting to note here is the fact that a spin $\frac{1}{2}$ particle has to be rotated by $2 \times 2\pi = 4\pi$ (!) in order to become the same state, very much in contrast to our classical expectation. It is due to the factor $\frac{1}{2}$ in the exponent. This very interesting quantum feature has been experimentally verified by the group of Helmut Rauch [16] using neutron interferometry.

7.2.2 Spin Algebra

Since spin is some kind of angular momentum we just use again the Lie algebra [3], which we found for the angular momentum observables, and replace the operator $\vec{L}$ by $\vec{S}$

$$[\, S_i \,,\, S_j \,] \,=\, i\hbar\, \epsilon_{ijk}\, S_k \;. \tag{7.17}$$

The spin observable squared also commutes with all the spin components, as in Eq. (6.19)

$$\left[\, \vec{S}^{\,2} \,,\, S_i \,\right] \,=\, 0 \;. \tag{7.18}$$

Still in total analogy with Definition 6.1 we can construct ladder operators $S_\pm$

$$S_\pm \,:=\, S_{\mathrm{x}} \,\pm\, i\, S_{\mathrm{y}} \,, \tag{7.19}$$

which satisfy the analogous commutation relations as before (see Eqs. (6.21) and (6.23))

$$\begin{aligned} [\, S_{\mathrm{z}} \,,\, S_\pm \,] \;&=\; \pm\,\hbar\, S_\pm && (7.20)\\ [\, S_+ \,,\, S_- \,] \;&=\; 2\,\hbar\, S_{\mathrm{z}} \;. && (7.21) \end{aligned}$$

The operators now act on the space of (2 component) spinor states, a two–dimensional Hilbert space which is equipped with a basis of eigenstates $|\, s \,,\, m_{\mathrm{s}} \,\rangle$, labeled by their eigenvalues s and m_{s} of $\vec{S}^{\,2}$ and S_{z} respectively

$$\left|\, \tfrac{1}{2} \,,\, \tfrac{1}{2} \,\right\rangle \equiv |\!\uparrow\rangle \quad , \quad \left|\, \tfrac{1}{2} \,,\, -\tfrac{1}{2} \,\right\rangle \equiv |\!\downarrow\rangle \;. \tag{7.22}$$

These two states, we call them "up" and "down", are eigenstates of the σ_{z} Pauli matrix, which we can interpret as the spin observable in z-direction, with eigenvalues $+1$ and -1

$$\begin{aligned} \sigma_{\mathrm{z}}\, |\!\uparrow\rangle \;&=\; +\, |\!\uparrow\rangle && (7.23)\\ \sigma_{\mathrm{z}}\, |\!\downarrow\rangle \;&=\; -\, |\!\downarrow\rangle \;. && (7.24) \end{aligned}$$

---

[3] Remark for experts: The Lie algebra we studied earlier was that of the three–dimensional rotation group $SO(3)$, the group of orthogonal (hence the "O") 3×3 matrices with determinant 1 (which is indicated by "S" for "special"), while we here are studying the group $SU(2)$, the group of unitary 2×2 matrices with determinant 1. The fact that these two Lie algebras look identical is not a mere coincidence, but is due to the fact that $SU(2)$ is the so called universal covering group of $SO(3)$ relating those groups in a very close way.

The eigenstates are orthogonal and normalized, i.e.

$$\langle \uparrow | \downarrow \rangle = 0 \quad , \quad \langle \uparrow | \uparrow \rangle = \langle \downarrow | \downarrow \rangle = 1 \, . \tag{7.25}$$

Let's now gather the established facts to find a representation of the operators on the aforesaid 2–dimensional Hilbert space by noting that we are looking for a hermitian 2 × 2 matrix with eigenvalues ±1, which is trivially satisfied by choosing

$$\sigma_{\mathrm{z}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \, . \tag{7.26}$$

The eigenstates take the form

$$| \uparrow \rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad | \downarrow \rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \, . \tag{7.27}$$

We then construct ladder operators from the Pauli matrices, i.e.

$$\sigma_{\pm} = \sigma_{\mathrm{x}} \pm i \, \sigma_{\mathrm{y}} \, , \tag{7.28}$$

which satisfy the spin algebra (recall Eq. (7.17) or Eq. (7.20))

$$[\, \sigma_{\mathrm{z}} \, , \, \sigma_{\pm} \,] = \pm 2 \, \sigma_{\pm} \, . \tag{7.29}$$

Since the following relations hold true for the ladder operators

$$\sigma_{-} \, | \uparrow \rangle = 2 \, | \downarrow \rangle \tag{7.30}$$

$$\sigma_{+} \, | \downarrow \rangle = 2 \, | \uparrow \rangle \tag{7.31}$$

$$\sigma_{-} \, | \downarrow \rangle = \sigma_{+} \, | \uparrow \rangle = 0 \, , \tag{7.32}$$

we can represent them as

$$\sigma_{+} = 2 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad , \quad \sigma_{-} = 2 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \, . \tag{7.33}$$

We then easily get the representation of σ_{x} and σ_{y} from Eq. (7.28)

$$\sigma_{\mathrm{x}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \sigma_{\mathrm{y}} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \, . \tag{7.34}$$

Properties of the Pauli matrices:

The Pauli matrices satisfy a Lie algebra

$$[\, \sigma_{\mathrm{i}} \, , \, \sigma_{\mathrm{j}} \,] = 2 \, i \, \epsilon_{\mathrm{ijk}} \, \sigma_{\mathrm{k}} \, , \tag{7.35}$$

their square is the identity (in two dimensions) and they are traceless

$$\sigma_{\mathrm{i}}^2 \,=\, \mathbb{1} \quad , \quad \mathrm{Tr}\,\sigma_{\mathrm{i}} \,=\, 0 \,. \tag{7.36}$$

Furthermore, an interesting property arises if one considers not only the square of the Pauli matrices, but the following object

$$(\vec{\sigma}\,\vec{A})\,(\vec{\sigma}\,\vec{B}) \;=\; \vec{A}\,\vec{B} \,+\, i\,\vec{\sigma}\,(\vec{A}\,\times\,\vec{B}) \,. \tag{7.37}$$

If we e.g. set $\vec{A} = \vec{B} = \vec{P}$ in Eq. (7.37) the cross product on the right side vanishes and we can express the kinetic energy as

$$\frac{\vec{p}^{\,2}}{2m} \;=\; \frac{(\vec{\sigma}\vec{p})^2}{2m} \,. \tag{7.38}$$

It turns out to be actually the correct way to include spin into the kinetic part of the Schrödinger equation.

7.2.3 Spin Measurements

Let us now verify on the theoretical side the conclusions about the spin measurements which we presented in Fig. 7.3. So we assume to have already performed a measurement in the z-direction and thus obtained a particle in the "up" state $|\uparrow\rangle$. If we again perform a measurement in the same direction, which means we apply the operator σ_{z} on the state $|\uparrow\rangle$, we always get the result $+1$

$$\langle\uparrow|\,\underbrace{\sigma_{\mathrm{z}}\,|\uparrow\rangle}_{+1\,|\uparrow\rangle} \,=\, \langle\uparrow|\uparrow\rangle \,=\, 1 \,. \tag{7.39}$$

It is easily verified that we will always get the result -1 if we perform the same measurement in the state $|\downarrow\rangle$.

Let us now consider a spin measurement in x-direction on that "up" state

$$\langle\uparrow|\,\underbrace{\sigma_{\mathrm{x}}\,|\uparrow\rangle}_{|\downarrow\rangle} \,=\, \langle\uparrow|\downarrow\rangle \,=\, 0 \,. \tag{7.40}$$

What we find is that, though individual results give ±1, on average we find a zero result. The reason is that each individual result $+1$ or -1 occurs exactly to 50% of the total outcome. That the individual measurements give ±1 is obvious since these are the eigenvalues of σ_{x}. The corresponding eigenfunctions are given by

$$|+\rangle \,=\, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \quad , \quad |-\rangle \,=\, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix} \,. \tag{7.41}$$

7.3 Two Particles with Spin $\frac{1}{2}$

In this section we show how to describe a system of two particles with spin, how to combine the two spins. The addition of the two spins of the constituent particles works in the same way as the addition of any other angular momenta. To describe the states of the combined spin we need the ***tensor product*** operation "$\otimes$", which expresses a product between elements of different (Hilbert) spaces, e.g. $\mathcal{H}^{\mathrm{A}}$, $\mathcal{H}^{\mathrm{B}}$. These tensor products form an element of a larger (Hilbert) space $\mathcal{H}^{\mathrm{AB}}$, where

$$\mathcal{H}^{\mathrm{AB}} \,=\, \mathcal{H}^{\mathrm{A}} \otimes \mathcal{H}^{\mathrm{B}} . \tag{7.42}$$

To give this a little more substance, let us consider two spin $\frac{1}{2}$ particles that each can be in either of the states $\uparrow$ or $\downarrow$. Then we conclude that there are 4 different ways to combine these states: $\uparrow\uparrow$, $\uparrow\downarrow$, $\downarrow\uparrow$, and $\downarrow\downarrow$, which we construct by using the tensor product operation (in the notation the tensor product $\otimes$ is often omitted for brevity)

$$\begin{aligned}
|\!\uparrow\rangle \otimes |\!\uparrow\rangle &= |\!\uparrow\rangle\,|\!\uparrow\rangle = |\!\uparrow\uparrow\rangle && (7.43)\\
|\!\uparrow\rangle \otimes |\!\downarrow\rangle &= |\!\uparrow\rangle\,|\!\downarrow\rangle = |\!\uparrow\downarrow\rangle && (7.44)\\
|\!\downarrow\rangle \otimes |\!\uparrow\rangle &= |\!\downarrow\rangle\,|\!\uparrow\rangle = |\!\downarrow\uparrow\rangle && (7.45)\\
|\!\downarrow\rangle \otimes |\!\downarrow\rangle &= |\!\downarrow\rangle\,|\!\downarrow\rangle = |\!\downarrow\downarrow\rangle\ . && (7.46)
\end{aligned}$$

Let us now try to define the spin operators for the composite system by looking at the tensor product structure of the Hilbert space (see Eq. (7.42)). Since we know the individual spin operators $\vec{S}^{\,(\mathrm{A})}$ and $\vec{S}^{\,(\mathrm{B})}$ acting in $\mathcal{H}^{\mathrm{A}}$ or $\mathcal{H}^{\mathrm{B}}$ respectively, we construct the composite spin operator such that the individual operators are acting in their respective subspace, i.e. we set

$$\vec{S}^{\,(\mathrm{AB})} \,=\, \vec{S}^{\,(\mathrm{A})} \otimes \mathbb{1}^{\,(\mathrm{B})} + \mathbb{1}^{\,(\mathrm{A})} \otimes \vec{S}^{\,(\mathrm{B})} \,=\, \vec{S}^{\,(\mathrm{A})} + \vec{S}^{\,(\mathrm{B})} . \tag{7.47}$$

In total analogy we construct the operator for the spin component in z-direction as

$$S_{\mathrm{z}}^{(\mathrm{AB})} \,=\, S_{\mathrm{z}}^{(\mathrm{A})} \otimes \mathbb{1}^{\,(\mathrm{B})} + \mathbb{1}^{\,(\mathrm{A})} \otimes S_{\mathrm{z}}^{(\mathrm{B})} \,=\, S_{\mathrm{z}}^{(\mathrm{A})} + S_{\mathrm{z}}^{(\mathrm{B})} . \tag{7.48}$$

Next we calculate the spin z-component of the vector $\uparrow\uparrow$ (Eq. (7.43)) by using the relations discussed above, in particular Eqs. (7.23), (7.24) and Eq. (7.13)

$$\begin{aligned}
S_{\mathrm{z}}^{(\mathrm{AB})}\,|\!\uparrow\uparrow\rangle &= \big(S_{\mathrm{z}}^{(\mathrm{A})} + S_{\mathrm{z}}^{(\mathrm{B})}\big)\,|\!\uparrow\rangle \otimes |\!\uparrow\rangle = \\
&= (S_{\mathrm{z}}^{(\mathrm{A})}\,|\!\uparrow\rangle) \otimes |\!\uparrow\rangle + |\!\uparrow\rangle \otimes (S_{\mathrm{z}}^{(\mathrm{B})}\,|\!\uparrow\rangle) = \\
&= (\frac{\hbar}{2}\,|\!\uparrow\rangle) \otimes |\!\uparrow\rangle + |\!\uparrow\rangle \otimes (\frac{\hbar}{2}\,|\!\uparrow\rangle) = \\
&= (\frac{\hbar}{2} + \frac{\hbar}{2})\,|\!\uparrow\rangle \otimes |\!\uparrow\rangle = \hbar\,|\!\uparrow\uparrow\rangle\ . && (7.49)
\end{aligned}$$

Using the same methods, for the other combinations (Eqs. (7.44) - (7.46)), we find

$$\begin{aligned}
S_{\mathrm{z}}\,|\!\downarrow\downarrow\rangle &= -\,\hbar\,|\!\downarrow\downarrow\rangle && (7.50)\\
S_{\mathrm{z}}\,|\!\uparrow\downarrow\rangle &= S_{\mathrm{z}}\,|\!\downarrow\uparrow\rangle = 0\ , && (7.51)
\end{aligned}$$

where we dropped the label ${}^{(\mathrm{AB})}$ for ease of notation. If not indicated otherwise the operators now always act on the total space, analogously to Eqs. (7.47) and (7.48). Unluckily, these naive combinations of two spins, although eigenstates of the S_{z} operator, are not all of them simultaneous eigenstates of the squared spin operator $\vec{S}^{\,2}$. The operator $\vec{S}^{\,2}$ we now express in terms of tensor products

$$\begin{aligned} \vec{S}^{\,2} &= (\vec{S}^{\,(\mathrm{AB})})^2 = \Big(\vec{S}^{\,(\mathrm{A})} \otimes \mathbb{1}^{(\mathrm{B})} + \mathbb{1}^{(\mathrm{A})} \otimes \vec{S}^{\,(\mathrm{B})}\Big)^2 = \\ &= (\vec{S}^{\,(\mathrm{A})})^2 \otimes \mathbb{1}^{(\mathrm{B})} + 2\,\vec{S}^{\,(\mathrm{A})} \otimes \vec{S}^{\,(\mathrm{B})} + \mathbb{1}^{(\mathrm{A})} \otimes (\vec{S}^{\,(\mathrm{B})})^2 \,. \end{aligned} \tag{7.52}$$

We calculate $(\vec{S}^{\,(\mathrm{A})})^2$ by using a property of the Pauli matrices Eq. (7.36)

$$(\vec{S}^{\,(\mathrm{A})})^2 = \frac{\hbar^2}{4}\left(\sigma_{\mathrm{x}}^2 + \sigma_{\mathrm{y}}^2 + \sigma_{\mathrm{z}}^2\right) = \frac{\hbar^2}{4}\,3\,\mathbb{1} = \frac{3}{4}\,\hbar^2 \,. \tag{7.53}$$

Using this result we rewrite Eq. (7.52)

$$\begin{aligned} \vec{S}^{\,2} &= \frac{3}{4}\,\hbar^2\,\mathbb{1}^{(\mathrm{A})} \otimes \mathbb{1}^{(\mathrm{B})} + 2\,\vec{S}^{\,(\mathrm{A})} \otimes \vec{S}^{\,(\mathrm{B})} + \frac{3}{4}\,\hbar^2\,\mathbb{1}^{(\mathrm{A})} \otimes \mathbb{1}^{(\mathrm{B})} = \\ &= \frac{\hbar^2}{4}\,\left[6\,\mathbb{1} \otimes \mathbb{1} + 2\,(\sigma_{\mathrm{x}} \otimes \sigma_{\mathrm{x}} + \sigma_{\mathrm{y}} \otimes \sigma_{\mathrm{y}} + \sigma_{\mathrm{z}} \otimes \sigma_{\mathrm{z}})\right] \,. \end{aligned} \tag{7.54}$$

Computing the action of $\vec{S}^{\,2}$ on our spin states (Eqs. (7.43) - (7.46)) we find that, though $\uparrow\uparrow$ and $\downarrow\downarrow$ are indeed also eigenstates of $\vec{S}^{\,2}$ corresponding to a quantum number $s = 1$

$$\begin{aligned} \vec{S}^{\,2}\,|\!\uparrow\uparrow\rangle &= \frac{\hbar^2}{4}\,\left[6\,\mathbb{1} \otimes \mathbb{1} + 2\,(\sigma_{\mathrm{x}} \otimes \sigma_{\mathrm{x}} + \sigma_{\mathrm{y}} \otimes \sigma_{\mathrm{y}} + \sigma_{\mathrm{z}} \otimes \sigma_{\mathrm{z}})\right]|\!\uparrow\rangle \otimes |\!\uparrow\rangle = \\ &= \frac{\hbar^2}{4}\,\left[6\,|\!\uparrow\uparrow\rangle + 2\,|\!\downarrow\downarrow\rangle + 2\,(i)^2\,|\!\downarrow\downarrow\rangle + 2\,|\!\uparrow\uparrow\rangle\right] = \\ &= \frac{\hbar^2}{4}\,8\,|\!\uparrow\uparrow\rangle = 2\,\hbar^2\,|\!\uparrow\uparrow\rangle = \hbar^2\,1\,(1+1)\,|\!\uparrow\uparrow\rangle \;, \end{aligned} \tag{7.55}$$

$$\vec{S}^{\,2}\,|\!\downarrow\downarrow\rangle = 2\,\hbar^2\,|\!\downarrow\downarrow\rangle \;, \tag{7.56}$$

the states $\uparrow\downarrow$ and $\downarrow\uparrow$ are no eigenstates[4] of the squared spin operator

$$\vec{S}^{\,2}\,|\!\uparrow\downarrow\rangle = \vec{S}^{\,2}\,|\!\downarrow\uparrow\rangle = \hbar^2\,(|\!\uparrow\downarrow\rangle + |\!\downarrow\uparrow\rangle) \;. \tag{7.57}$$

However, we can form linear combinations of the states $\uparrow\downarrow$ and $\downarrow\uparrow$, where we choose the appropriate weights[5] $\frac{1}{\sqrt{2}}$ for normalization, which are eigenstates to the quantum numbers $s = 1$ and $s = 0$

$$\vec{S}^{\,2}\,\tfrac{1}{\sqrt{2}}\,(|\!\uparrow\downarrow\rangle + |\!\downarrow\uparrow\rangle) = 2\,\hbar^2\,\tfrac{1}{\sqrt{2}}\,(|\!\uparrow\downarrow\rangle + |\!\downarrow\uparrow\rangle) \tag{7.58}$$

---

[4] Generally, naively forming tensor products of simultaneous eigenstates of squared angular momentum and angular momentum z-component operators does not give an eigenstate of the corresponding operators on the tensor product space. There is, however, a simple way of finding these eigenstates, by calculating the so called Clebsch-Gordon coefficients, which though mathematically simple can be a quite tiresome procedure.

[5] They are the Clebsch-Gordon coefficients for the addition of two spin $\frac{1}{2}$'s.

$$\vec{S}^2 \, \tfrac{1}{\sqrt{2}} \left(|\uparrow\downarrow\rangle \, - \, |\downarrow\uparrow\rangle \right) \; = \; 0 \, . \tag{7.59}$$

So we find a ***triplet*** of states $|\, s, \, m_{\mathrm{s}} \,\rangle$, belonging to the spin quantum number $s = 1$, with magnetic spin quantum numbers $m_{\mathrm{s}} = -1, 0, 1$,

$$|\, 1, \, -1 \,\rangle \; = \; |\downarrow\downarrow\rangle \tag{7.60}$$

$$|\, 1, \, 0 \,\rangle \; = \; \tfrac{1}{\sqrt{2}} \left(|\uparrow\downarrow\rangle \, + \, |\downarrow\uparrow\rangle \right) \tag{7.61}$$

$$|\, 1, \, +1 \,\rangle \; = \; |\uparrow\uparrow\rangle \tag{7.62}$$

and a ***singlet*** with quantum numbers $s = 0$ and $m_{\mathrm{s}} = 0$

$$|\, 0, \, 0 \,\rangle \; = \; \tfrac{1}{\sqrt{2}} \left(|\uparrow\downarrow\rangle \, - \, |\downarrow\uparrow\rangle \right) \, . \tag{7.63}$$

Remark I: Of course, we could have also calculated the state (7.61) by applying the lowering operator S_{-} (see Eq. (7.19)) to the $|\uparrow\uparrow\rangle$ state

$$S_{-} \, |\uparrow\uparrow\rangle \; = \; \left(S_{-}^{\mathrm{A}} \, + \, S_{-}^{\mathrm{B}} \right) \, |\uparrow\uparrow\rangle \; = \; \hbar \left(|\uparrow\downarrow\rangle \, + \, |\downarrow\uparrow\rangle \right) \, . \tag{7.64}$$

Remark II: Statistics. We divide particles into two groups, those of integer spin values, called ***bosons***, and those of half-integer spin values, called ***fermions***. One can then conclude, that wave functions describing bosons must always be even/symmetric functions, while the ones, describing fermions must be odd/antisymmetric functions. This causes for both types of particles to be subjected to different statistical behavior, fermions are governed by the ***Fermi–Dirac statistics*** while bosons follow the ***Bose–Einstein statistics***.

One very important conclusion from this spin–statistic relation is the so called *Pauli exclusion principle.*

Proposition 7.3.1 (Pauli exclusion principle) *No two or more fermions may occupy the same quantum state simultaneously.*

Therefore, if we describe two (or more) electrons in the same spin state (symmetric spin states), they can not be located at the same position, since this would also be a symmetric state and thus the total wave function would be symmetric, which is impossible for fermions.

Chapter 8

Perturbation Theory, Zeeman Effect, Stark Effect

Unfortunately, apart from a few simple examples, the Schrödinger equation is generally not exactly solvable and we therefore have to rely upon approximative methods to deal with more realistic situations. Such methods include *perturbation theory*, the *variational method* and the *WKB*[1]*-approximation.* In our Scriptum we, however, just cope with perturbation theory in its simplest version. It is a systematic procedure for obtaining approximate solutions to the unperturbed problem which is assumed to be known exactly.

8.1 Time–Independent Perturbation Theory

The method of perturbation theory is that we deform slightly – perturb – a known Hamiltonian H_0 by a new Hamiltonian H_{I}, some potential responsible for an interaction of the system, and try to solve the Schrödinger equation approximately since, in general, we will be unable to solve it exactly. Thus we start with a Hamiltonian of the following form

$$H \;=\; H_0 \,+\, \lambda\, H_{\mathrm{I}} \qquad \text{where} \quad \lambda \,\in\, [0,1]\,, \tag{8.1}$$

where the Hamiltonian H_0 is perturbed by a smaller term, the interaction H_{I} with λ small. The unperturbed Hamiltonian is assumed to be solved and has well-known eigenfunctions and eigenvalues, i.e.

$$H_0 \left|\, n^0 \right\rangle \;=\; E_{\mathrm{n}}^{(0)} \left|\, n^0 \right\rangle \;, \tag{8.2}$$

where the eigenfunctions are chosen to be normalized

$$\left\langle\, m^0 \,\middle|\, n^0 \right\rangle \;=\; \delta_{\mathrm{mn}}\,. \tag{8.3}$$

The superscript " ^{0} " denotes the eigenfunctions and eigenvalues of the unperturbed system H_0 , and we furthermore require the unperturbed eigenvalues to be non-degenerate

$$E_{\mathrm{m}}^{(0)} \;\neq\; E_{\mathrm{n}}^{(0)}\,, \tag{8.4}$$

---

[1], named after Wentzel, Kramers and Brillouin.

otherwise we would use a different method leading to the so-called *degenerate perturbation theory*. What we are now going to investigate are the eigenvalues E_n and eigenfunctions $|\, n\,\rangle$ of the total Hamiltonian H

$$H\,|\,n\,\rangle \,=\, E_\mathrm{n}\,|\,n\,\rangle\;. \tag{8.5}$$

The basic idea of perturbation theory then is to expand the eigenvalues and eigenfunctions in a power series, without taking care of the corresponding convergence properties. In many cases, considering just the first few terms of the power series gives us reasonable results, even if the whole power series can be shown to be divergent.

Let us therefore start with a power series in the parameter λ, which is a good approximation for small λ

$$E_\mathrm{n} \,=\, E_\mathrm{n}^{(0)} \,+\, \lambda\, E_\mathrm{n}^{(1)} \,+\, \lambda^2\, E_\mathrm{n}^{(2)} \,+\, \cdots \tag{8.6}$$

$$|\,n\,\rangle \,=\, \left|\,n^{\,0}\,\right\rangle \,+\, \lambda\,\left|\,n^{\,1}\,\right\rangle \,+\, \lambda^2\,\left|\,n^{\,2}\,\right\rangle \,+\, \cdots\;, \tag{8.7}$$

where we have coefficient functions of lambda, the unperturbed term, first and second order perturbative corrections and so on. Inserting these expansions into the Schrödinger equation (8.5) together with Hamiltonian (8.1), we get

$$\begin{aligned}&(H_0 \,+\, \lambda\, H_\mathrm{I})\;(|\,n^{\,0}\,\rangle \,+\, \lambda\,|\,n^{\,1}\,\rangle \,+\, \lambda^2\,|\,n^{\,2}\,\rangle \,+\, \cdots) \,=\\ =\;&\Big(E_\mathrm{n}^{(0)} \,+\, \lambda\, E_\mathrm{n}^{(1)} \,+\, \lambda^2\, E_\mathrm{n}^{(2)} \,+\, \cdots\Big)\;(|\,n^{\,0}\,\rangle \,+\, \lambda\,|\,n^{\,1}\,\rangle \,+\, \lambda^2\,|\,n^{\,2}\,\rangle \,+\, \cdots)\;.\end{aligned} \tag{8.8}$$

Equating coefficients we obtain the following relations:

$$\lambda^0 : \qquad H_0\,|\,n^0\,\rangle \,=\, E_\mathrm{n}^{(0)}\,|\,n^0\,\rangle \tag{8.9}$$

$$\lambda^1 : \qquad H_0\,|\,n^1\,\rangle \,+\, H_\mathrm{I}\,|\,n^0\,\rangle \,=\, E_\mathrm{n}^{(0)}\,|\,n^1\,\rangle \,+\, E_\mathrm{n}^{(1)}\,|\,n^0\,\rangle \tag{8.10}$$

$$\lambda^2 : \quad H_0\,|\,n^2\,\rangle \,+\, H_\mathrm{I}\,|\,n^1\,\rangle \,=\, E_\mathrm{n}^{(0)}\,|\,n^2\,\rangle \,+\, E_\mathrm{n}^{(1)}\,|\,n^1\,\rangle \,+\, E_\mathrm{n}^{(2)}\,|\,n^0\,\rangle\;. \tag{8.11}$$

Additionally, it is convenient to normalize the following transition amplitude for the perturbed energy eigenstate $|\,n\,\rangle$

$$\langle\,n^0\,|\,n\,\rangle \,=\, 1 \tag{8.12}$$

$$\Rightarrow\quad \lambda\,\left\langle\,n^0\,\middle|\,n^1\,\right\rangle \,+\, \lambda^2\,\left\langle\,n^0\,\middle|\,n^2\,\right\rangle \,+\, \cdots \,=\, 0 \tag{8.13}$$

$$\Rightarrow\quad \left\langle\,n^0\,\middle|\,n^1\,\right\rangle \,=\, \left\langle\,n^0\,\middle|\,n^2\,\right\rangle \,=\, \cdots \,=\, 0\,. \tag{8.14}$$

8.1.1 First Order Corrections

To derive the first order corrections we multiply the first order coefficient equation (Eq. 8.10) with $\langle\,n^0\,|$ from the left and let the hermitian, unperturbed Hamiltonian $H_0^\dagger = H_0$ act to the left on $\langle\,n^0\,|$

$$E_\mathrm{n}^{(0)}\,\underbrace{\left\langle\,n^0\,\middle|\,n^1\,\right\rangle}_{0} + \left\langle\,n^0\,\middle|\;H_\mathrm{I}\,\middle|\,n^0\,\right\rangle \,=\, E_\mathrm{n}^{(0)}\,\underbrace{\left\langle\,n^0\,\middle|\,n^1\,\right\rangle}_{0} + E_\mathrm{n}^{(1)}\,\underbrace{\left\langle\,n^0\,\middle|\,n^0\,\right\rangle}_{1}\;. \tag{8.15}$$

This leads us directly to the following fundamental result, which is already a sufficient correction for many cases to reproduce physical observations:

Theorem 8.1 *The perturbative energy corrections of first order are given by the expectation value of the perturbing Hamiltonian in the unperturbed states*

$$E_{\mathrm{n}}^{(1)} = \langle n^0 | H_{\mathrm{I}} | n^0 \rangle$$

Although we now have the energy correction we are interested in, we are curious as to how the state itself is perturbed. We therefore expand the first order expansion state $| n^1 \rangle$ into a complete orthonormal system $\{| n^0 \rangle\}$

$$| n^1 \rangle = \sum_{m \neq n} | m^0 \rangle \langle m^0 | n^1 \rangle , \tag{8.16}$$

where all the terms for $m = n$ vanish due to our choice of normalization (Eq. 8.14). We are then interested in the expansion coefficient $\langle m^0 | n^1 \rangle$, which we calculate by using again Eq. (8.10). This time, however, we multiply the equation with $\langle m^0 |$

$$E_{\mathrm{m}}^{(0)} \langle m^0 | n^1 \rangle + \langle m^0 | H_{\mathrm{I}} | n^0 \rangle = E_{\mathrm{n}}^{(0)} \langle m^0 | n^1 \rangle + E_{\mathrm{n}}^{(1)} \langle m^0 | n^0 \rangle . \tag{8.17}$$

Though the last term on the right side of Eq. (8.17) would be nonzero if $m = n$, the state expansion excludes this case in the summation and therefore we can ignore this term, which allows us to express the expansion coefficient as

$$\langle m^0 | n^1 \rangle = \frac{\langle m^0 | H_{\mathrm{I}} | n^0 \rangle}{E_{\mathrm{n}}^{(0)} - E_{\mathrm{m}}^{(0)}} , \tag{8.18}$$

where the matrix element $\langle m^0 | H_{\mathrm{I}} | n^0 \rangle$ has to be calculated.

So the important result for the first order state correction is:

Theorem 8.2 *The perturbative state corrections of first order are given by the superposition of all unperturbed states with energies different from the considered state*

$$| n^1 \rangle = \sum_{m \neq n} \frac{\langle m^0 | H_{\mathrm{I}} | n^0 \rangle}{E_{\mathrm{n}}^{(0)} - E_{\mathrm{m}}^{(0)}} | m^0 \rangle$$

Remark: Due to the energy denominator the adjacent energy levels contribute in this approximation stronger to the corrections of the state (assuming nearly constant matrix elements) than the more remote ones, what is intuitively expected.

8.1.2 Second Order Corrections

To find the corrections of second order, we start similar as in the last subsection, by this time multiplying Eq. (8.11) by $\langle\, n^0 \,|$

$$E_{\mathrm{n}}^{(0)} \underbrace{\langle\, n^0 \,|\, n^2 \,\rangle}_{0} + \langle\, n^0 \,|\; H_{\mathrm{I}} \;|\, n^1 \,\rangle \;=\; E_{\mathrm{n}}^{(0)} \underbrace{\langle\, n^0 \,|\, n^2 \,\rangle}_{0} + E_{\mathrm{n}}^{(1)} \underbrace{\langle\, n^0 \,|\, n^1 \,\rangle}_{0} + E_{\mathrm{n}}^{(2)} \underbrace{\langle\, n^0 \,|\, n^0 \,\rangle}_{1} \;. \tag{8.19}$$

Further inserting our result of Section 8.1.1 (Theorem 8.2), we get

$$E_{\mathrm{n}}^{(2)} \;=\; \sum_{m \neq n} \langle\, n^0 \,|\; H_{\mathrm{I}} \;|\, m^0 \,\rangle \, \frac{\langle\, m^0 \,|\; H_{\mathrm{I}} \;|\, n^0 \,\rangle}{E_{\mathrm{n}}^{(0)} \;-\; E_{\mathrm{m}}^{(0)}} \;, \tag{8.20}$$

then we immediately find the result for the second order energy correction as:

Theorem 8.3 *The perturbative energy corrections of second order are determined by all unperturbed states with energy different from the considered state*

$$E_{\mathrm{n}}^{(2)} \;=\; \sum_{m \neq n} \frac{\left|\langle\, m^0 \,|\, H_{\mathrm{I}} \,|\, n^0 \,\rangle\right|^2}{E_{\mathrm{n}}^{(0)} - E_{\mathrm{m}}^{(0)}}$$

Remark I: Again, the adjacent energy levels contribute to the majority of the correction if the matrix elements are assumed to be of approximately the same order.

Remark II: For the ground state the energy correction of second order is always negative, i.e. $E_1^{(2)} < 0$, since $E_1^{(0)} < E_{\mathrm{m}}^{(0)}$.

8.2 Zeeman Effect

The energy levels of particles carrying a magnetic dipole moment, e.g. the H-atom, in an exterior (homogeneous) magnetic field are further split up, i.e., the degeneracy of the energy eigenvalues is removed.

The Hamiltonian of the electron in the H-atom therefore has an additional term, the ***Zeeman term*** H_{Z}, perturbing the already known term $H_{\mathrm{Coul.}}$ which contains the Coulomb potential (Eq. 6.95). We thus have the total Hamiltonian

$$H \;=\; H_{\mathrm{Coul.}} \,+\, H_{\mathrm{Z}} \,, \tag{8.21}$$

where the Zeeman term is the scalar product of the magnetic dipole moment $\vec{\mu}$ and the exterior magnetic field $\vec{B}$, see Eq. (7.4)

$$H_{\mathrm{Z}} \;=\; -\,\vec{\mu}\,\vec{B} \;=\; -\,\frac{q}{2\,m\,c}\,\left(\vec{L} \,+\, g\,\vec{S}\right)\,\vec{B} \;. \tag{8.22}$$

We now assume without loss of generality that the magnetic field is oriented along the z-direction, where it has the strength $B^{\,\mathrm{ext.}}$, and insert the electron mass m_{e}, charge $e = -|e|$ and g-factor $g_{\mathrm{e}} = 2$

$$H_{\mathrm{Z}} \,=\, \frac{|e|}{2\, m_{\mathrm{e}}\, c}\, (\, L_{\mathrm{z}} \,+\, 2\, S_{\mathrm{z}}\,)\; B^{\,\mathrm{ext.}}\; . \tag{8.23}$$

We further assume that the magnetic field is sufficiently strong, such that the orbital angular momentum $\vec{L}$ and the spin $\vec{S}$ decouple and the unperturbed states are product states of the form

$$|\, n,\, l,\, m_l \rangle\; |\!\uparrow\rangle \quad \text{and} \quad |\, n,\, l,\, m_l \rangle\; |\!\downarrow\rangle\ , \tag{8.24}$$

which is called the *Paschen–Back Effect*. The (first order) energy corrections $\Delta\, E$ to the pure Coulomb energy are then given by Theorem 8.1, where we immediately see that

$$\Delta\, E^{(1)}_{\mathrm{n},\, l,\, \mathrm{m}_l,\, \mathrm{m}_s} \,=\, \frac{|e|\,\hbar}{2\, m_{\mathrm{e}}\, c}\, (\, m_l \,+\, \underbrace{2\, m_s}_{\pm 1})\, B^{\,\mathrm{ext.}}\; . \tag{8.25}$$

Thus every, until now degenerate, l-level is split up by the energy correction term, depending on the spin orientation. When recalling the energy levels of the Coulomb potential (Eq. 6.104)

$$E^{\mathrm{Coul.}}_{\mathrm{n},\, l} \,=\, -\,\frac{E_{\mathrm{I}}}{n^2} \,=\, -\,\frac{13,6\, eV}{n^2}\, , \tag{8.26}$$

and Bohr's magneton μ_{B}

$$\mu_{\mathrm{B}} \,=\, \frac{|e|\,\hbar}{2\, m_{\mathrm{e}}\, c} \,=\, 5,788\,\times\,10^{-5}\, eV\, /\, \mathrm{Tesla}\ , \tag{8.27}$$

we see that the Zeeman energy term is a small correction with respect to the Coulomb energy, even for very strong magnetic fields ($B^{\,\mathrm{ext.}} \approx$ Tesla).

8.3 Stark Effect

The Stark effect is the electric analogue to the Zeeman effect, i.e., a particle carrying an electric dipole moment, like the H-atom, will get a splitting of its energy levels when subjected to an exterior electric field. The Hamiltonian of the H-atom thus has (another) additional term, the ***Stark term*** H_{Stark}, which is perturbing the Coulomb Hamiltonian

$$H_{\mathrm{Stark}} \,=\, -\,\vec{\mathcal{E}}\,\vec{\mu}_{\mathrm{el}} \,=\, -\,\underbrace{q}_{-|e|}\;\vec{\mathcal{E}}\,\vec{x} \,=\, |e|\,\mathcal{E}\, z\ , \tag{8.28}$$

where $\vec{\mathcal{E}}$ denotes the external electric field which, without loss of generality, we choose along the z-direction, and $\vec{\mu}_{\mathrm{el}}$ is the electrical dipole moment. For the simple case of an electron of charge $-|e|$ placed at a distance $\vec{x}$ of the oppositely charged proton, resting at the origin, the electric dipole moment reduces to $\vec{\mu}_{\mathrm{el}} = -|e|\vec{x}$.

To calculate the energy corrections of first (Theorem 8.1) and second (Theorem 8.3) order , we need to consider expectation values or scalar products of the Stark term in the Coulomb states of the form $|\, n,\, l,\, m_l \rangle$. This is, however, more complicated than the situation we studied before. To make the problem more precise, let us calculate the following scalar product, where we already know from Eq. (6.5) that L_z and z commute

$$\langle\, n,\, l,\, m_l \,|\, \underbrace{[\, L_z\,,\, z\,]}_{=0} \,|\, n',\, l',\, m_l' \,\rangle \;=\; (m_l \,-\, m_l')\; \langle\, n,\, l,\, m_l \,|\; z \;|\, n',\, l',\, m_l' \,\rangle \;=\; 0\;. \tag{8.29}$$

Since the commutator of L_z and z is zero we can conclude that the states $|\, n,\, l,\, m_l \rangle$ and $z\,|\, n',\, l',\, m_l' \,\rangle$ must be orthogonal for $m_l \neq m_l'$, and since we are only interested in nonzero matrix elements we choose

$$m_l \equiv m_l'\,. \tag{8.30}$$

At the same time, however, for the scalar product to be nonzero, we have to require that the azimuthal quantum numbers be unequal, i.e. $l \neq l'$. The reason for this is the parity of the z operator. We remember from Theorem 4.2 that a basis of states can be chosen for a symmetric potential, such as the Coulomb potential (Eq. (6.95)), that consists entirely of even and odd functions, which in this case are the Legendre polynomials $P_l(\cos\theta)$. These are polynomials of order l in $\cos\theta = \frac{z}{r}$ and thus alternately even and odd functions.

Therefore, if the functions $\psi_{\mathrm{n},\, l,\, \mathrm{m}_l}$ and $\psi_{\mathrm{n}',\, l',\, \mathrm{m}_l}$ have the same azimuthal quantum number, i.e. parity, their product has parity +1. The z operator however changes its sign under a spatial inversion and the integration of the scalar product vanishes, because

$$\int\limits_{-a}^{+a} dx\; \underbrace{f(x)}_{\text{even}}\; \underbrace{g(x)}_{\text{odd}} \;\equiv\; 0\;. \tag{8.31}$$

We can then assume that

$$l' \;=\; l \pm 1\,, \tag{8.32}$$

which is a reasonable choice since the electric dipole interacts with the electric field by exchanging photons of spin 1, so the angular momentum of the electron must decrease or increase by 1. The Stark effect will therefore only occur for matrix elements of the form

$$\langle\, n,\, l,\, m_l \,|\; z \;|\, n',\, (l \pm 1),\, m_l \rangle \;\neq\; 0\,. \tag{8.33}$$

Let us now consider the energy correction for the ground state of the hydrogen atom, the first order correction is

$$E^{(1)}_{1,\,0,\,0} \;=\; |e|\,\mathcal{E}\;\langle\, 1,\, 0,\, 0 \,|\; z \;|\, 1,\, 0,\, 0 \,\rangle \;=\; 0\;. \tag{8.34}$$

It vanishes due to the requirement of Eq. (8.32) and we thus see that there is no linear Stark effect for the hydrogen ground state. The effect of second order is of the form

$$E^{(2)}_{1,\,0,\,0} \;=\; e^2\,\mathcal{E}^2 \sum_{n=2}^{\infty} \frac{|\,\langle\, n,\, 1,\, 0 \,|\; z \;|\, 1,\, 0,\, 0 \,\rangle\,|^2}{E^{(0)}_1 \;-\; E^{(0)}_n}\;, \tag{8.35}$$

where the unperturbed energy levels $E_n^{(0)}$ are given by Eq. (6.105). This expression can be calculated using perturbative methods, which leads to the result

$$E^{(2)}_{1,\,0,\,0} \;=\; -\,\frac{9}{4}\,\frac{\hbar^2}{m_{\mathrm{e}}\,e^2}\,\mathcal{E}^2 \;=\; -\,\frac{9}{4}\,r_{\mathrm{B}}\,\mathcal{E}^2\;, \tag{8.36}$$

where r_{B} is the Bohr radius (Eq. (6.97)).

The first excited state, in contrast to the ground state, is four-fold degenerate, i.e. there are 4 states $|\,2,\,0,\,0\,\rangle$, $|\,2,\,1,\,1\,\rangle$, $|\,2,\,1,\,0\,\rangle$ and $|\,2,\,1,\,-1\,\rangle$, that belong to the principal quantum number $n = 2$. In this case one has to consider degenerate perturbation theory, which leads to calculating the eigenvalues of a non-diagonal matrix.

Because of the requirement of Eq. (8.30) we only consider elements with equal magnetic quantum number, which is only the case for the states $|\,2,\,0,\,0\,\rangle$ and $|\,2,\,1,\,0\,\rangle$. It leaves us a 2×2 matrix to diagonalize. The second requirement (Eq. (8.32)) rids us of the diagonal elements of the matrix. The remaining nonzero off-diagonal elements are

$$\langle\,2,\,1,\,0\,|\;z\;|\,2,\,0,\,0\,\rangle \;=\; \langle\,2,\,0,\,0\,|\;z\;|\,2,\,1,\,0\,\rangle\;. \tag{8.37}$$

The nonzero eigenvalues can thus be easily calculated and we conclude that the excited state has a linear Stark effect correction with the result

$$E^{(1)}_{2,\,l,\,\mathrm{m}_l} \;=\; \pm\,|e|\,r_{\mathrm{B}}\,\mathcal{E}\;. \tag{8.38}$$

Further splitting of the excited hydrogen energy levels can be seen in Fig. 8.1.

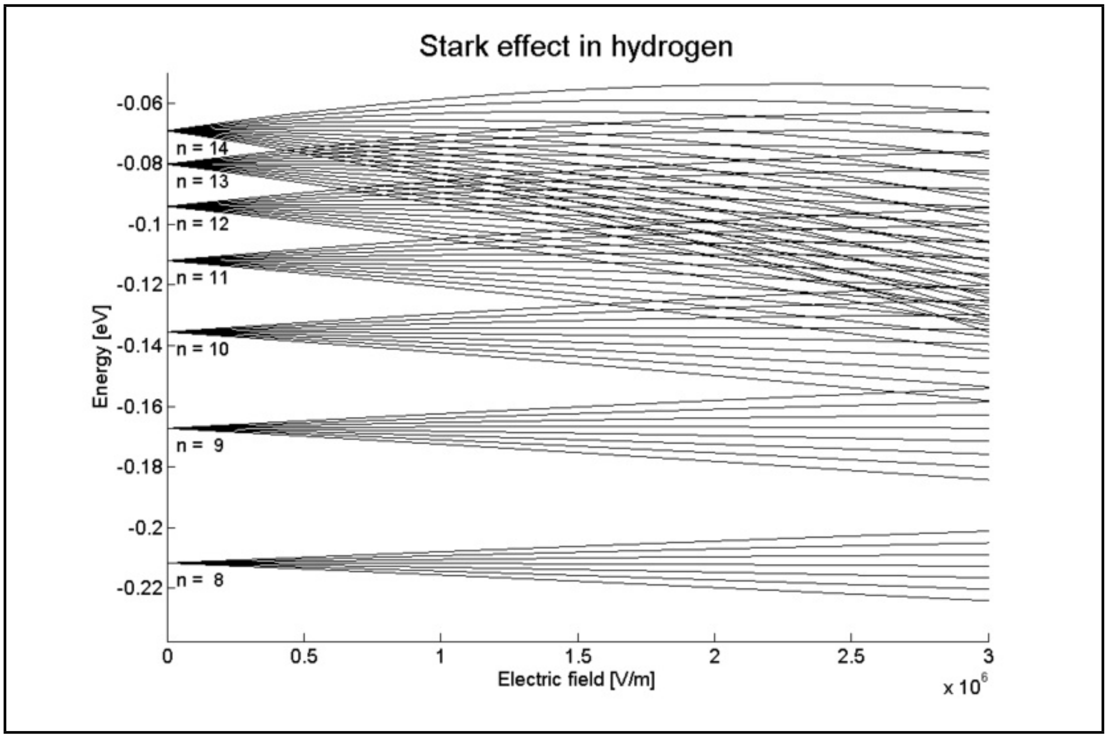

Figure 8.1: Stark Effect in Hydrogen: The until then degenerate excited energy levels are split up if an exterior electric field is applied. Figure from: http://en.wikipedia.org/wiki/Image:Stark_splitting.png

Chapter 9

Density Matrices

In this chapter we want to introduce density matrices, also called density operators, which conceptually take the role of the state vectors discussed so far, as they encode all the (accessible) information about a quantum mechanical system. It turns out that the "pure" states, described by state vectors[1] $|\,\psi\,\rangle$ on Hilbert space, are idealized descriptions that cannot characterize statistical (incoherent) mixtures, which often occur in the experiment, i.e. in Nature. These objects are very important for the theory of quantum information and quantum communication. More detailed information about the density matrix formalism can be found in [17].

9.1 General Properties of Density Matrices

Consider an observable A in the "pure" state $|\,\psi\,\rangle$ with the expectation value given by

$$\langle\, A\,\rangle_\psi \,=\, \langle\,\psi\,|\, A\,|\,\psi\,\rangle\;, \tag{9.1}$$

then the following definition is obvious:

Definition 9.1 *The **density matrix** ρ for the pure state $|\,\psi\,\rangle$ is given by*

$$\rho \,:=\, |\,\psi\,\rangle\,\langle\,\psi\,|$$

This density matrix has the following properties:

$$\text{I)}\quad \rho^2 \,=\, \rho \quad \text{projector} \tag{9.2}$$

$$\text{II)}\quad \rho^\dagger \,=\, \rho \quad \text{hermiticity} \tag{9.3}$$

$$\text{III)}\quad \mathrm{Tr}\,\rho \,=\, 1 \quad \text{normalization} \tag{9.4}$$

$$\text{IV)}\quad \rho \,\geq\, 0 \quad \text{positivity} \tag{9.5}$$

---

[1] Remark for experts: It is possible to find a vector representation for every given quantum mechanical state, even those represented by a density matrix. This can be done via the so-called GNS (Gelfand-Neumark-Segal) construction. This vector representation need not, however, be of any practical form and the concept of the density matrix is therefore inevitable.

The first two properties follow immediately from Definition 9.1 and property III) can be verified using the definition of the trace operation for an arbitrary operator D:

Definition 9.2 *The* ***trace*** *of an operator D is given by*

$$\mathrm{Tr}\, D \;:=\; \sum_{\mathrm{n}} \langle\, n \,|\; D \;|\, n \,\rangle$$

where $\{\,|\, n\,\rangle\,\}$ is an arbitrary CONS.

Lets e.g. take the operator $D = |\,\psi\,\rangle\,\langle\,\phi\,|$ and calculate its trace

$$\mathrm{Tr}\, D \;=\; \sum_{\mathrm{n}} \langle\, n \,|\, \psi \,\rangle \langle\, \phi \,|\, n \,\rangle \;=\; \sum_{\mathrm{n}} \langle\, \phi \underbrace{|\, n \,\rangle \langle\, n \,|}_{\mathbb{1}} \psi \,\rangle \;=\; \langle\, \phi \,|\, \psi \,\rangle \;. \tag{9.6}$$

Property IV) means that the eigenvalues of ρ are greater or equal to zero, which can also be expressed as

$$\langle\, \varphi \,|\; \rho \;|\, \varphi \,\rangle \;=\; \langle\, \varphi \,|\, \psi \,\rangle \langle\, \psi \,|\, \varphi \,\rangle \;=\; |\,\langle\, \varphi \,|\, \psi \,\rangle\,|^2 \;\geq\; 0 \;, \tag{9.7}$$

which is an important property because probabilities are always greater or equal to zero. What we still have to ensure is that the expectation value of an observable in the state $|\,\psi\,\rangle$ can be reproduced, which we will formulate in the following theorem:

Theorem 9.1 *The expectation value of an observable A in a state, represented by a density matrix ρ, is given by*

$$\langle\, A \,\rangle_{\rho} \;=\; \mathrm{Tr}\,(\rho\, A)$$

Proof:

$$\begin{aligned} \mathrm{Tr}\,(\rho\, A) \;&=\; \mathrm{Tr}\;(|\,\psi\,\rangle \langle\, \psi \,|\; A) \;=\; \sum_{\mathrm{n}} \langle\, n \,|\, \psi \,\rangle \langle\, \psi \,|\; A \;|\, n \,\rangle \;= \\ &=\; \sum_{\mathrm{n}} \langle\, \psi \,|\; A \underbrace{|\, n \,\rangle \;\langle\, n \,|}_{\mathbb{1}} \psi \,\rangle \;=\; \langle\, \psi \,|\; A \;|\, \psi \,\rangle \;=\; \langle\, A \,\rangle \;. \quad \text{q.e.d.} \end{aligned} \tag{9.8}$$

9.2 Pure and Mixed States

Now we can introduce a broader class of states represented by density matrices, the so-called ***mixed states*** in contrast to the states we have considered until now, the so-called ***pure states***.

9.2.1 Pure States

Let's begin with the pure states. Consider an ensemble of given objects in the states $\{\,|\,\psi_{\mathrm{i}}\,\rangle\,\}$. If all the objects are in the same state, the ensemble is represented by a ***pure state***. To make probabilistic statements the whole ensemble of identically prepared systems must be considered.

Let the system be, e.g., in the state $|\,\psi\,\rangle$ which we can expand with respect to the eigenstates of an (hermitian) operator A

$$|\,\psi\,\rangle \;=\; \sum_{\mathrm{n}} c_{\mathrm{n}}\;|\,n\,\rangle\;, \quad \text{where} \quad A\,|\,n\,\rangle \;=\; a_{\mathrm{n}}\;|\,n\,\rangle\;. \tag{9.9}$$

The expectation value is then given by

$$\langle\,A\,\rangle_{\psi} \;=\; \sum_{\mathrm{n}} |c_{\mathrm{n}}|^2\,a_{\mathrm{n}} \;=\; \sum_{\mathrm{n}} \frac{N_{\mathrm{n}}}{N}\,a_{\mathrm{n}}\;, \tag{9.10}$$

where $|c_{\mathrm{n}}|^2$ is the probability to measure the eigenvalue a_{n}. It corresponds to the fraction N_{n}/N , the incidence the eigenvalue a_{n} occurs, where N_{n} is the number of times this eigenvalue has been measured out of an ensemble of N objects.

The state is characterized by a density matrix of the form of Definition 9.1, with the properties I) - Ⅳ) (Eqs. (9.2) - (9.5)), where we can combine property I) and Ⅲ) to conclude

$$\mathrm{Tr}\,\rho^2 \;=\; 1\,. \tag{9.11}$$

9.2.2 Mixed States

Let us next study the situation where not all of the N systems (objects) of the ensemble are in the same state, i.e. N_{i} systems are in the state $|\,\psi_{\mathrm{i}}\,\rangle$ respectively, such that $\sum N_{\mathrm{i}} = N$. The probability p_{i} to find an individual system of the ensemble described by the state $|\,\psi_{\mathrm{i}}\,\rangle$ is then given by

$$p_{\mathrm{i}} \;=\; \frac{N_{\mathrm{i}}}{N}\;, \quad \text{where} \quad \sum_{\mathrm{i}} p_{\mathrm{i}} \;=\; 1\;. \tag{9.12}$$

We can thus write down the ***mixed state*** as a convex sum, i.e. a weighted sum with $\sum_{\mathrm{i}} p_{\mathrm{i}} \;=\; 1$, of pure state density matrices

$$\rho_{\mathrm{mix}} \;=\; \sum_{\mathrm{i}} p_{\mathrm{i}}\,\rho_{\mathrm{i}}^{\mathrm{pure}} \;=\; \sum_{\mathrm{i}} p_{\mathrm{i}}\;|\,\psi_{\mathrm{i}}\,\rangle\,\langle\,\psi_{\mathrm{i}}\,|\;\;. \tag{9.13}$$

The expectation value is again given by Theorem 9.1, i.e.

$$\langle\,A\,\rangle_{\rho_{\mathrm{mix}}} \;=\; \mathrm{Tr}\,(\rho_{\mathrm{mix}}\,\mathrm{A})\;, \tag{9.14}$$

where we can express the expectation value of the mixed state as a convex sum of expectation values of its constituent pure states, i.e.

$$\langle\, A\,\rangle_{\rho_{\mathrm{mix}}} \,=\, \sum_{\mathrm{i}} p_{\mathrm{i}}\, \langle\, \psi_{\mathrm{i}}\,|\, A\,|\, \psi_{\mathrm{i}}\,\rangle\ . \tag{9.15}$$

Proof:

$$\begin{aligned}
\mathrm{Tr}\,(\rho_{\mathrm{mix}}\, A) \;&=\; \mathrm{Tr}\left(\sum_{\mathrm{i}} p_{\mathrm{i}}\,|\,\psi_{\mathrm{i}}\,\rangle\,\langle\,\psi_{\mathrm{i}}\,|\,A\right) \,= \\
&=\; \sum_{\mathrm{n}}\sum_{\mathrm{i}} p_{\mathrm{i}}\,\langle\, n\,|\,\psi_{\mathrm{i}}\,\rangle\,\langle\,\psi_{\mathrm{i}}\,|\,A\,|\,n\,\rangle \,= \\
&=\; \sum_{\mathrm{i}} p_{\mathrm{i}}\,\langle\,\psi_{\mathrm{i}}\,|\,A\,\underbrace{\sum_{\mathrm{n}}|\,n\,\rangle\,\langle\, n\,|}_{\mathbb{1}}\,\psi_{\mathrm{i}}\,\rangle \,= \\
&=\; \sum_{\mathrm{i}} p_{\mathrm{i}}\,\langle\,\psi_{\mathrm{i}}\,|\,A\,|\,\psi_{\mathrm{i}}\,\rangle\ . \quad \text{q.e.d.}
\end{aligned} \tag{9.16}$$

Properties II) - IV) (Eqs. (9.3) - (9.5)) are still valid for mixed states, but property I) does no longer hold

$$\rho^2_{\mathrm{mix}} \,=\, \sum_{\mathrm{i}}\sum_{\mathrm{j}} p_{\mathrm{i}}\,p_{\mathrm{j}}\,|\,\psi_{\mathrm{i}}\,\rangle\,\underbrace{\langle\,\psi_{\mathrm{i}}\,|\,\psi_{\mathrm{j}}\,\rangle}_{\delta_{\mathrm{ij}}}\,\langle\,\psi_{\mathrm{j}}\,| \,=\, \sum_{\mathrm{i}} p^2_{\mathrm{i}}\,|\,\psi_{\mathrm{i}}\,\rangle\,\langle\,\psi_{\mathrm{i}}\,| \,\neq\, \rho_{\mathrm{mix}}\ , \tag{9.17}$$

where we, w.l.o.g., assumed that $|\,\psi_{\mathrm{i}}\,\rangle$ and $|\,\psi_{\mathrm{j}}\,\rangle$ are orthonormal. We can then calculate the trace of ρ^2, which, in contrast to pure states, is no longer equal to 1 but smaller

$$\begin{aligned}
\mathrm{Tr}\,\rho^2_{\mathrm{mix}} \;&=\; \sum_{\mathrm{n}}\langle\, n\,|\sum_{\mathrm{i}}\sum_{\mathrm{j}} p_{\mathrm{i}}\,p_{\mathrm{j}}\,|\,\psi_{\mathrm{i}}\,\rangle\,\langle\,\psi_{\mathrm{i}}\,|\,\psi_{\mathrm{j}}\,\rangle\,\langle\,\psi_{\mathrm{j}}\,|\,n\,\rangle \,= \\
&=\; \sum_{\mathrm{i}}\sum_{\mathrm{j}} p_{\mathrm{i}}\,p_{\mathrm{j}}\,\langle\,\psi_{\mathrm{i}}\,|\,\psi_{\mathrm{j}}\,\rangle\;\langle\,\psi_{\mathrm{j}}\,|\sum_{\mathrm{n}}|\,n\,\rangle\;\langle\, n\,|\,\psi_{\mathrm{i}}\,\rangle \,= \\
&=\; \sum_{\mathrm{i}}\sum_{\mathrm{j}} p_{\mathrm{i}}\,p_{\mathrm{j}}\,|\,\langle\,\psi_{\mathrm{i}}\,|\,\psi_{\mathrm{j}}\,\rangle\,|^2 \,= \\
&=\; \sum_{\mathrm{i}} p^2_{\mathrm{i}} \,<\, \sum_{\mathrm{i}} p_{\mathrm{i}} \,=\, 1\ .
\end{aligned} \tag{9.18}$$

The last step in this calculation is obvious, since $0 \leq p_{\mathrm{i}} \leq 1$ and therefore $p^2_{\mathrm{i}} \leq p_{\mathrm{i}}$. We conclude that the trace of ρ^2 is a good measure for the mixedness of a density matrix, since it is equal to 1 for pure states and strictly smaller than 1 for mixed states. For a ***maximally mixed state*** we have for a given dimension d of the system

$$\mathrm{Tr}\,\rho^2_{\mathrm{mix}} \,=\, \tfrac{1}{d} \,>\, 0\,. \tag{9.19}$$

9.3 Time Evolution of Density Matrices

We now want to find the equation of motion for the density matrix. We start from the time dependent Schrödinger equation and its hermitian conjugate

$$i\hbar \frac{\partial}{\partial t} \,|\, \psi \,\rangle \;=\; H \,|\, \psi \,\rangle \quad \xrightarrow{\dagger} \quad -i\hbar \frac{\partial}{\partial t} \,\langle\, \psi \,| \;=\; \langle\, \psi \,|\, H \;. \tag{9.20}$$

Then we differentiate the density matrix of a mixed state (Eq. (9.13)) with respect to time, we multiply it by $i\hbar$ and combine this with Eq. (9.20)

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \rho \;&=\; i\hbar \sum_{\mathrm{i}} p_{\mathrm{i}} \,(\underbrace{|\, \dot{\psi}_{\mathrm{i}} \,\rangle}_{-\frac{i}{\hbar} H \,|\, \psi_{\mathrm{i}} \,\rangle} \;\langle\, \psi_{\mathrm{i}} \,| + |\, \psi_{\mathrm{i}} \,\rangle \underbrace{\langle\, \dot{\psi}_{\mathrm{i}} \,|}_{\frac{i}{\hbar} \langle\, \psi_{\mathrm{i}} \,|\, H} \;) \;= \\ &=\; \sum_{\mathrm{i}} p_{\mathrm{i}} \,(\, H \, \rho_{\mathrm{i}}^{\mathrm{pure}} \,-\, \rho_{\mathrm{i}}^{\mathrm{pure}} \, H \,) \;= \\ &=\; [\, H \,,\, \rho \,] \;. \end{aligned} \tag{9.21}$$

Theorem 9.2 *Density matrices satisfy the* ***von Neumann Equation***

$$i\hbar \tfrac{\partial}{\partial t} \rho \;=\; [\, H \,,\, \rho \,]$$

The von Neumann equation is the quantum mechanical analogue to the classical Liouville equation, recall the substitution (2.85).

The time evolution of the density matrix we can also describe by applying an unitary operator, the ***time shift operator*** $U(t, t_0)$, also called ***propagator***

$$U(t, t_0) \;=\; e^{-\frac{i}{\hbar} H \,(t \,-\, t_0)} \;. \tag{9.22}$$

It allows us to relate the density matrix at a later time t to the density matrix at some earlier time t_0

$$\rho(t) \;=\; U(t, t_0) \, \rho(t_0) \, U^{\dagger}(t, t_0) \;. \tag{9.23}$$

Furthermore, it helps us to prove, for instance, that the *mixedness* $\mathrm{Tr}\, \rho^2$ *of a density matrix is time independent*

$$\mathrm{Tr}\, \rho^2(t) \;=\; \mathrm{Tr}\, (U \, \rho(t_0) \underbrace{U^{\dagger} \, U}_{\mathbb{1}} \, \rho(t_0) \, U^{\dagger}) \;=\; \mathrm{Tr}\, (\rho(t_0) \, \rho(t_0) \underbrace{U^{\dagger} \, U}_{\mathbb{1}}) \;=\; \mathrm{Tr}\, \rho^2(t_0) \;, \tag{9.24}$$

where we used the cyclicity of the trace operation.

Example: Density matrix for spin $\frac{1}{2}$

Generally, this will be a 2×2 matrix that can be written as linear combination of the identity $\mathbb{1}$ and the Pauli matrices $\sigma_\text{x}, \sigma_\text{y}$ and σ_z , as

$$\rho \,=\, \tfrac{1}{2}\,(\mathbb{1} \,+\, \vec{a}\,\vec{\sigma})\ . \tag{9.25}$$

The coefficient $\vec{a}$ is named the ***Bloch vector*** and can be calculated as the expectation value of the Pauli matrices

$$\vec{a} \,=\, \text{Tr}\,(\rho\,\vec{\sigma}) \,=\, \langle\,\vec{\sigma}\,\rangle\ . \tag{9.26}$$

All spin $\frac{1}{2}$ density matrices lie on or within the so-called ***Bloch sphere*** (with radius $\vec{a} = 1$) and are determined by the Bloch vector $\vec{a}$. The length of the Bloch vector thus tells us something about the mixedness, the polarization of an ensemble, i.e. of a beam of spin $\frac{1}{2}$ particles, e.g. electrons or neutrons. We say the beam is polarized if $a_\text{i} = 1$ and completely unpolarized if $a_\text{i} = 0$, for all i . This means that pure and mixed states can be characterized via the Bloch vector in the following way

$$\text{pure state} \quad \rho^2 \,=\, \rho \qquad \Rightarrow \qquad |\vec{a}| \,=\, 1 \tag{9.27}$$

$$\text{mixed state} \quad \rho^2 \,\neq\, \rho \qquad \Rightarrow \qquad |\vec{a}| \,<\, 1\ . \tag{9.28}$$

A totally mixed state ($a_\text{i} = 0$ for all i) can then be written as

$$\rho_\text{mix} \,=\, \frac{1}{2}\,(\,|\!\uparrow\rangle\,\langle\uparrow\!|\,+\,|\!\downarrow\rangle\,\langle\downarrow\!|\,) \,=\, \frac{1}{2}\,\mathbb{1}\ , \tag{9.29}$$

such that

$$\text{Tr}\,\rho_\text{mix} \,=\, 1 \quad \text{and} \quad \text{Tr}\,\rho^2_\text{mix} \,=\, \frac{1}{2}\ . \tag{9.30}$$

Remark: Note, the decomposition (9.29) into up $|\!\uparrow\rangle\,\langle\uparrow\!|$ and down $|\!\downarrow\rangle\,\langle\downarrow\!|$ states is by no means unique, we can achieve the totally mixed state $\rho_\text{mix} \,=\, \frac{1}{2}\,\mathbb{1}$ in many different ways.

Chapter 10

Entanglement of Quantum Systems

At the end we want to turn to an area which arose quite recently, in the late 80ties and 90ties, which received much attention since it opens a new kind of communication which wasn't possible before. This is quantum information, quantum communication and quantum computation. The foundation came from John Bell and his famous inequalities which focused on a certain feature of the quantum states, on the so-called nonlocality or entanglement of the states. Entanglement is the basic ingredient of quantum information theory and became nowadays the most used term in quantum physics, a kind of "magic word". In this chapter we will give an introduction into the entanglement concept, into Bell inequalities and discuss finally its most exciting consequence, namely quantum teleportation.

10.1 Entanglement and Separability

The term "entanglement" ("Verschränkung" in the original German phrasing) was introduced by Erwin Schrödinger in 1935 in order to describe an intrinsic feature of quantum mechanics, that arises from the structure of Hilbert space, i.e. the superposition principle of the states therein. He stated that:

> A composite quantum system, whose subsystems are distant from each other, is in an **"*entangled state*"** if the total system is in a well defined state, but the subsystems themselves are not.

To formulate this statement more rigorously let us look at systems of two components, e.g. two spin $\frac{1}{2}$ particles or two photons, and construct the four possible product states (eigenstates of σ_z), already known from Eqs. (7.43) - (7.46)

$$|\uparrow\rangle \otimes |\uparrow\rangle \quad , \; |\uparrow\rangle \otimes |\downarrow\rangle \quad , \; |\downarrow\rangle \otimes |\uparrow\rangle \quad , \; |\downarrow\rangle \otimes |\downarrow\rangle \; . \tag{10.1}$$

These states are of the class of the so-called "separable" states, or just "product" states. In quantum information theory the left subsystem is often called ***Alice*** and the right one ***Bob***. For photons the arrow notation of the σ_{z} eigenstates is usually replaced by the letters H and V, standing for horizontal and vertical polarization respectively

$$|\,H\,\rangle \otimes |\,H\,\rangle \quad , |\,H\,\rangle \otimes |\,V\,\rangle \quad , |\,V\,\rangle \otimes |\,H\,\rangle \quad , |\,V\,\rangle \otimes |\,V\,\rangle \;. \tag{10.2}$$

Generally, instead of the states of Eq. (10.1) and Eq. (10.2), one could use any quantum system consisting of two subsystems, i.e. a bipartite system with each two degrees of freedom, called a 2-qubit state. But whichever of the above product states you choose, the states of the individual subsystems Alice and Bob are well defined. For example, for the state $|\uparrow\rangle \otimes |\downarrow\rangle$, when measuring the spin along the z-direction, the outcome of the measurements of Alice will be "up", i.e. determined by the state $|\uparrow\rangle$, while Bob's result will be always "down", i.e. determined by the state $|\downarrow\rangle$.

From these product states we can form linear combinations, e.g.

$$|\,\psi^{\pm}\,\rangle \;=\; \frac{1}{\sqrt{2}}\,(\,|\uparrow\rangle\,|\downarrow\rangle \pm |\downarrow\rangle\,|\uparrow\rangle\,) \tag{10.3}$$

$$|\,\phi^{\pm}\,\rangle \;=\; \frac{1}{\sqrt{2}}\,(\,|\uparrow\rangle\,|\uparrow\rangle \pm |\downarrow\rangle\,|\downarrow\rangle\,)\;. \tag{10.4}$$

These are the four (maximally) entangled 2-qubit states, called the ***Bell states***. We also recognize that the Bell states are not only artificial constructions but we have already encountered them when studying composite quantum systems with spin, where the $|\,\psi^{\pm}\,\rangle$ were found to be simultaneous eigenstates of $\vec{S}$ and S_{z}, see Eqs. (7.61) - (7.63).

Let us now try to grasp what entanglement in terms of e.g. the Bell state $|\,\psi^{\pm}\,\rangle$ means. Consider two observers, Alice and Bob, who are far apart, both performing measurements on their part of the entangled state. They will both get the results ↑ or ↓ equally often, i.e. with a probability of $\frac{1}{2}$, which means that neither of the subsystems is in a well defined state. But on the other hand, everytime Alice measures ↑, Bob will with certainty measure the result ↓ and vice versa, i.e. the measurement results are perfectly correlated.

This has some far reaching consequences for the structure of the theory, i.e. it leads to a loss of locality and/or realism of the theory, which is backed up by experiments and it creates interesting possibilities for information processing, such as quantum cryptography, quantum teleportation and many more.

10.2 EPR Paradox

The EPR (Albert Einstein, Boris Podolsky and Nathan Rosen) paradox was formulated in 1935 (see Ref. [18]) as a gedankenexperiment to strengthen the claim that quantum mechanics in its probabilistic character was somehow incomplete. The argumentation was later on reformulated by David Bohm in 1952 (see Ref. [19]) for the simple quantum mechanical system of two spin $\frac{1}{2}$ particles, which is the way we will present the problem here.

The main argument of EPR is based on three requirements (see, e.g. Ref. [20]):

1. ***Completeness:*** Every element of physical reality must have a counterpart in the physical theory in order for the theory to be complete.
2. ***Realism:*** If the value of a physical quantity can be predicted with certainty, i.e. probability 1, without disturbing the system, then the quantity has physical reality.
3. ***Locality:*** There is no action at a distance. Measurements on a (sub)system do not affect measurements on (sub)systems that are far away.

EPR then conclude that under certain circumstances quantum mechanics is *not* a complete theory.

To understand this claim, let us consider two spin $\frac{1}{2}$ particles in the spin singlet state $|\,\psi^-\,\rangle$, the antisymmetric Bell state, and let them propagate in opposite direction along the x-axis from their source. Let then two observers, Alice and Bob, perform spin measurement along the z-direction.

Quantum mechanics tells us that for the state

$$|\,\psi^-\,\rangle \;=\; \frac{1}{\sqrt{2}}\,(\,|\!\uparrow\rangle\,|\!\downarrow\rangle \,-\, |\!\downarrow\rangle\,|\!\uparrow\rangle\,) \tag{10.5}$$

the result measured by Alice will be undetermined, i.e. either ↑ or ↓, but if Alice measures ↑, then Bob will measure ↓ with certainty and vice versa, which assigns physical reality to the spin of Bob's particle in the sense of EPR. Since there is no disturbance or action at a distance, EPR conclude, quantum mechanics does not contain any information about predetermined measurement outcomes and is therefore incomplete.

To account for the missing information, there must be some inaccessible parameter, a hidden variable, to determine which spin eigenvalue is realized in the measurement. EPR thus demand a hidden variable theory (HVT) to explain this problem.

In the same year as the EPR paper was published, Bohr replied (using the same title for his paper as EPR, i.e. "Can quantum-mechanical description of physical reality be considered complete ?", see Ref. [21]) criticizing their perception of reality and sweeping away their arguments without really answering the interesting paradox presented in the EPR paper. But due to Bohr's great authority the physical community followed his view that quantum mechanics *is* complete and the case rested for nearly 30 years, until John Bell published his famous article in 1964, presenting a way to solve the debate.

10.3 Bell Inequalities

In 1964 John S. Bell published an article (see Ref. [22]), proposing a solution to the (until then purely philosophical) Bohr-Einstein debate, which made it possible to determine experimentally, whether or not the requirements of the EPR paradox are fulfilled in Nature. The essence of Bell's proposal can be formulated in the following theorem:

Theorem 10.1 (Bell's theorem) *In certain experiments all local realistic theories (LRT) are incompatible with quantum mechanics.*

To formulate this mathematically, let us go back to the set-up discussed in Section 10.2, Alice and Bob performing spin measurements on an entangled 2-qubit state. For the schematic set-up of such an experiment see Fig. 10.1. The spin measurement along an arbitrary direction $\vec{a}$ is represented by the operator $\vec{\sigma}\vec{a}$ in quantum mechanics, but let us stay more general for now and assume that both Alice and Bob measure some observable $A^{\text{obs.}}(\vec{a})$ and $B^{\text{obs.}}(\vec{b})$ respectively, where the possible outcomes both also depend on some internal hidden parameter λ. The reality assumption of the EPR paradox then requires, that the parameter λ is the same for both subsystems

$$\begin{array}{rlcl} \text{Alice:} & A^{\text{obs.}}(\vec{a}) & \longrightarrow & A(\vec{a},\, \lambda) \;=\; \pm 1 \;(\uparrow \;\text{ or }\; \downarrow) \text{ or } 0 \text{ (no detection)} \\[1em] \text{Bob:} & B^{\text{obs.}}(\vec{b}) & \longrightarrow & B(\vec{b},\, \lambda) \;=\; \pm 1 \;(\uparrow \;\text{ or }\; \downarrow) \text{ or } 0 \text{ (no detection)} \end{array}$$

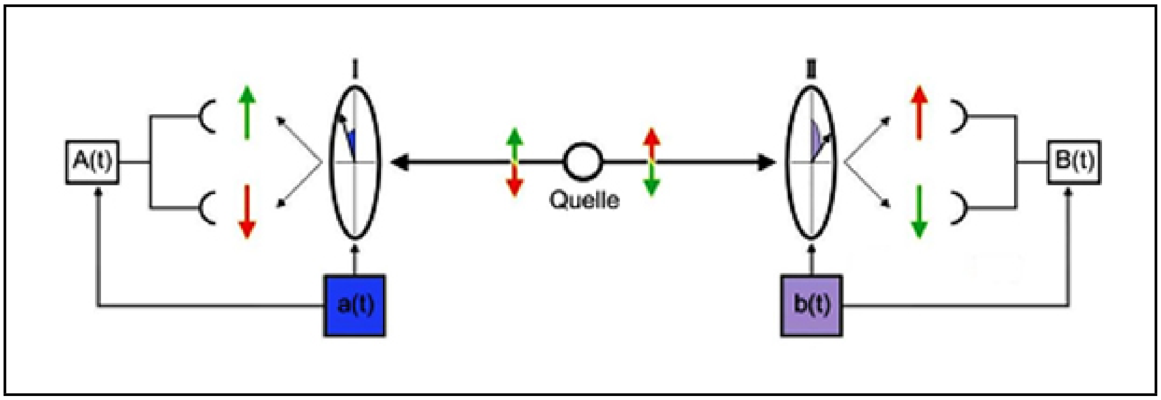

Figure 10.1: Set-up of a Bell-type experiment: Two entangled spin $\frac{1}{2}$ particles in the spin singlet state, the antisymmetric Bell state, are produced at a source and propagate in opposite direction along the x-axis. Two observers, Alice and Bob, perform a spin measurement along the z-direction. Whenever Alice measures $\uparrow$, Bob will measure $\downarrow$ with certainty, and vice versa. The detector orientations $a(t)$ and $b(t)$ will be fixed at random at each time t and registered by Alice and Bob together with the joint measurement results $A(t)$ and $B(t)$.

Bell's locality assumption requires that neither result depends on the measuring direction of the other side, i.e.

$$A(\vec{a},\, \cancel{\vec{b}},\, \lambda)\ , \quad B(\cancel{\vec{a}},\, \vec{b},\, \lambda)\ . \tag{10.6}$$

Furthermore we can assume that there is some normalized distribution function $\rho(\lambda)$,

$$\int d\lambda\, \rho(\lambda) \;=\; 1\ , \tag{10.7}$$

which determines the outcome for a given lambda. The expectation value of the combined measurement on both Alice's and Bob's side is then

$$E(\vec{a},\,\vec{b}) \;=\; \int d\lambda\,\rho(\lambda)\,A(\vec{a},\,\lambda)\,B(\vec{b},\,\lambda)\;. \tag{10.8}$$

Let us next consider a certain combination of such expectation values for measurements in different directions, but in the same state, i.e. for the same λ

$$\begin{aligned} E(\vec{a},\,\vec{b}) \,-\, E(\vec{a},\,\vec{b}') \;&=\; \int d\lambda\,\rho(\lambda)\,\Big(A(\vec{a},\,\lambda)\,B(\vec{b},\,\lambda) \,-\, A(\vec{a},\,\lambda)\,B(\vec{b}',\,\lambda)\Big) \\ &=\; \int d\lambda\,\rho(\lambda)\,\underbrace{A(\vec{a},\,\lambda)\,B(\vec{b},\,\lambda)}_{|\;|\leq 1}\Big(1 \,\pm\, A(\vec{a}',\,\lambda)\,B(\vec{b}',\,\lambda)\Big) \\ &-\; \int d\lambda\,\rho(\lambda)\,\underbrace{A(\vec{a},\,\lambda)\,B(\vec{b}',\,\lambda)}_{|\;|\leq 1}\Big(1 \,\pm\, A(\vec{a}',\,\lambda)\,B(\vec{b},\,\lambda)\Big)\;, \end{aligned} \tag{10.9}$$

where we just added and subtracted the same term to rewrite the equation. By then noting that the absolute values of the products of A and B on the right side must be smaller or equal to 1 we can make the following estimation:

$$\begin{aligned} \Big|\,E(\vec{a},\,\vec{b}) \,-\, E(\vec{a},\,\vec{b}')\,\Big| \;&\leq\; \Big|\int d\lambda\,\rho(\lambda)\,\Big(1 \,-\, A(\vec{a}',\,\lambda)\,B(\vec{b}',\,\lambda)\Big)\Big| \\ &+\; \Big|\int d\lambda\,\rho(\lambda)\,\Big(1 \,-\, A(\vec{a}',\,\lambda)\,B(\vec{b},\,\lambda)\Big)\Big|\;. \end{aligned} \tag{10.10}$$

Using Eq. (10.7) we get

$$\Big|\,E(\vec{a},\,\vec{b}) \,-\, E(\vec{a},\,\vec{b}')\,\Big| \,\leq\, 2 \,-\, \Big|\,E(\vec{a}',\,\vec{b}') \,+\, E(\vec{a}',\,\vec{b})\,\Big|\;, \tag{10.11}$$

which we rewrite in terms of the so-called *Bell parameter function* $S(\vec{a},\,\vec{a}',\,\vec{b},\,\vec{b}')$ as

$$|\,S(\vec{a},\,\vec{a}',\,\vec{b},\,\vec{b}')\,| = \Big|\,E(\vec{a},\,\vec{b}) \,-\, E(\vec{a},\,\vec{b}')\,\Big| \,+\, \Big|\,E(\vec{a}',\,\vec{b}') \,+\, E(\vec{a}',\,\vec{b})\,\Big| \,\leq\, 2\;. \tag{10.12}$$

Equation (10.12) is called the ***CHSH form of the Bell inequality***, named after Clauser, Horne, Shimony and Holt, who derived it separately from Bell. This inequality must be satisfied by any *local realistic theory*! (Clearly, satisfying Eq. (10.8).

The interesting question, of course, is whether or not quantum mechanics satisfies the inequality. Let us do the calculation. We first have to calculate the expectation value $E(\vec{a},\,\vec{b})$ for the state $|\,\psi^-\,\rangle$

$$\begin{aligned} E(\vec{a},\,\vec{b}) \;&=\; \langle\,\psi^-\,|\,\vec{\sigma}\vec{a}\otimes\vec{\sigma}\vec{b}\,|\,\psi^-\,\rangle \;= \\ &=\; \langle\,\psi^-\,|\begin{pmatrix} a_{\mathrm{z}} & a_{\mathrm{x}} \,-\, ia_{\mathrm{y}} \\ a_{\mathrm{x}} \,+\, ia_{\mathrm{y}} & -a_{\mathrm{z}} \end{pmatrix} \otimes \begin{pmatrix} b_{\mathrm{z}} & b_{\mathrm{x}} \,-\, ib_{\mathrm{y}} \\ b_{\mathrm{x}} \,+\, ib_{\mathrm{y}} & -b_{\mathrm{z}} \end{pmatrix}|\,\psi^-\,\rangle \;= \\ &=\; \frac{1}{2}\,(\,-\,2a_{\mathrm{x}}b_{\mathrm{x}} \,-\, 2a_{\mathrm{y}}b_{\mathrm{y}} \,-\, 2a_{\mathrm{z}}b_{\mathrm{z}}\,) \;=\; -\,\vec{a}\,\vec{b}\;, \end{aligned} \tag{10.13}$$

where we used the representation of Eq. (10.5) and Eq. (7.27). Furthermore we work with normalized vectors $|\vec{a}| = |\vec{b}| = 1$ and obtain

$$E(\vec{a}, \vec{b}) \,=\, -\,\vec{a}\,\vec{b} \,=\, -\,\cos(\alpha\,-\,\beta)\;, \tag{10.14}$$

where $(\alpha\,-\,\beta)$ is the angle between the directions $\vec{a}$ and $\vec{b}$. Choosing the directions in the plane perpendicular to the propagation axis in steps of $45°^{1}$, i.e. such that

$$\alpha-\beta \,=\, \alpha'\,-\,\beta' \,=\, \alpha'\,-\,\beta \,=\, \frac{\pi}{4}\;,\;\; \alpha\,-\,\beta' \,=\, 3\,\frac{\pi}{4}\;, \tag{10.15}$$

we get the Bell parameter function as predicted by quantum mechanics

$$\begin{aligned} S_{\mathrm{QM}} \;&=\; |-\,\cos(\alpha-\beta) + \cos(\alpha\,-\,\beta')| \,+\, |-\,\cos(\alpha'\,-\,\beta') \,-\, \cos(\alpha'\,-\,\beta)| \,= \\ &=\; |-\,\frac{\sqrt{2}}{2} \,-\, \frac{\sqrt{2}}{2}| \,+\, |-\,\frac{\sqrt{2}}{2} \,-\, \frac{\sqrt{2}}{2}| \,=\, 2\sqrt{2} \,>\, 2\;. \end{aligned} \tag{10.16}$$

This is the maximal possible violation of the Bell inequality, showing that all local realistic theories are *incompatible* with quantum mechanics which in turn must be a nonlocal and/or nonrealistic theory.

Experiments: In the past many experiments by several have been carried out, testing Bell inequalities. The first ones were Clauser and Freedman [23] and Fry and Thompson [24] in the mid 70ties. They used a radiative cascade of calcium or mercury to create the desired entangled states and had static analyzers. The next step forward happened in the early 80ties, when A. Aspect and his group [25] included an optical switch mechanism in the set-up of their instruments in order to achieve a spacelike separation of the event intervals. The experimental outcome of all experiments was in accordance with quantum mechanics.

The best experiment up to now concerning tests of Bell inequalities has been performed by A. Zeilinger and his group [26] in the late 90ties. They used a revolutionary new source, a BBO crystal, for producing the entangled states and were able to fabricate a truly random and ultrafast electro-optical switch mechanism for the analyzers, see Fig. 10.2. In this way Alice couldn't get any information from Bob with velocities less than the speed of light, which means that the strict Einstein locality conditions have been fulfilled. The experimental value for the Bell parameter function in this experiment was determined to be

$$S_{\mathrm{exp}} \,=\, 2,73\,\pm\,0,02\;, \tag{10.17}$$

in perfect agreement with quantum mechanics, which implies that LRT for describing Nature are ruled out, i.e. Nature contains a kind of nonlocality in the sense described above.

---

[1] Funny enough, these angles were called *awkward Irish angels* by John Bell himself.

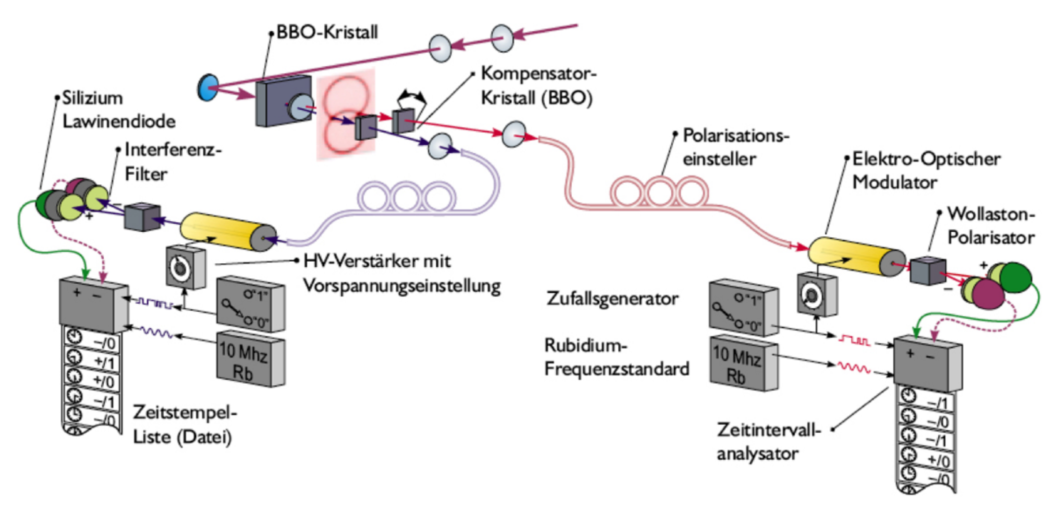

Figure 10.2: Experimental set-up of the Bell-type experiment of G. Weihs, T. Jennewein, C. Simon, H. Weinfurter and A. Zeilinger [26]: A BBO crystal is pumped by a laser and produces vertically V and horizontally H polarized light on two separate cones which overlap. In the overlap regions the photons (not "knowing" whether they are V or H) are entangled and are caught by two fibres which lead to the electro-optical modulators of the two observer stations Alice and Bob, where the data registration is carried out. The modulators rotate the polarization of the photons at random and ultrafast that no signal can travel from Alice to Bob with velocities less than the speed of light. In this way the group obtained a violation of Bell's inequality $S_{\text{exp}} = 2,73 \pm 0,02$ under strict Einstein locality conditions. The picture is taken from the PhD Thesis of Gregor Weihs, see http://www.uibk.ac.at/exphys/photonik/people/weihs.html

10.4 Quantum Teleportation

Quantum teleportation, invented by Charles Bennett and coauthors in 1993 [27], is one of the most exciting issues in quantum information and quantum communication. It's a special information transfer from Alice to Bob and often called "*beaming*" in the popular literature, but actually it's not such a thing, there is *no* transfer of matter or energy involved. Alice's particle has not been physically moved to Bob, only the particle's quantum state has been transferred. Let's see how it works.

10.4.1 Qubits

The information stored in a classical bit with value 0 or 1 is encoded in a physical two level system, which definitely is in one of two possible states. It can be realized by spin $\frac{1}{2}$ particles being in either spin $\uparrow$ or $\downarrow$ state, or by photons having polarization state H or V, to name just a few examples.

Quantum mechanically, however, superpositions of such definite states, e.g. ↑ and ↓, are possible, i.e.

$$\begin{aligned} |\,\psi\,\rangle &= a\,|\!\uparrow\rangle + b\,|\!\downarrow\rangle \\ &= a\,|\,H\,\rangle + b\,|\,V\,\rangle \\ &= a\,|\,0\,\rangle + b\,|\,1\,\rangle\;, \end{aligned} \tag{10.18}$$

where $a, b \in \mathbb{C}$ and $|a|^2 + |b|^2 = 1$. Such a state is called a ***qubit***. When measured, the qubit reveals 1 bit of information, either 0 or 1 with probabilities $|a|^2$ and $|b|^2$ respectively, which immediately follows when applying the corresponding projectors

$$|\,0\,\rangle\langle\,0\,|\,\psi\,\rangle = a\,|\,0\,\rangle \tag{10.19}$$

$$|\,1\,\rangle\langle\,1\,|\,\psi\,\rangle = b\,|\,1\,\rangle\;. \tag{10.20}$$

We see that a quantum mechanical system contains 1 qubit of information as long as it is not measured, which is more information than encoded in a classical bit. This fact is, for example, utilized in quantum communication or in developing a quantum computer. Another interesting application is quantum teleportation, what we want to discuss now in more detail.

10.4.2 Quantum Teleportation Set-Up

The idea of quantum teleportation is the transmission of an (even unknown) quantum state – 1 qubit – from one system at Alice's location onto another system at Bob's location. This transmission occurs instantaneously[2] by EPR correlations. In order for Bob to get any meaningful results he needs additional information from Alice, that is transmitted by a classical channel. This is also the reason why special relativity is not violated, since the classical information can not travel faster than the speed of light.

Let us now try to phrase this procedure in terms of quantum mechanics. Suppose Alice wants to teleport the state $|\,\psi\,\rangle$ of the incoming particle 1, given by

$$|\,\psi\,\rangle = a\,|\,H\,\rangle_1 + b\,|\,V\,\rangle_1 \qquad 1 \text{ qubit}\;. \tag{10.21}$$

An EPR source creates a pair of entangled particles, shared by Alice and Bob, which are in the Bell state

$$\left|\,\psi^-\,\right\rangle_{23} = \frac{1}{\sqrt{2}}\left(\,|\,H\,\rangle_2\,|\,V\,\rangle_3 - |\,V\,\rangle_2\,|\,H\,\rangle_3\,\right)\;. \tag{10.22}$$

Alice, on her side, performs a Bell-measurement, i.e. she entangles particles 1 and 2,

---

[2] Depending on the choice of interpretation of the quantum state, i.e. the wave function, one might come to different conclusions as to when the transmission of the state actually takes place. The transmission of the classical information, encoded in the qubit, however, is agreed upon by everybody, since this has to be supplemented by classical communication.

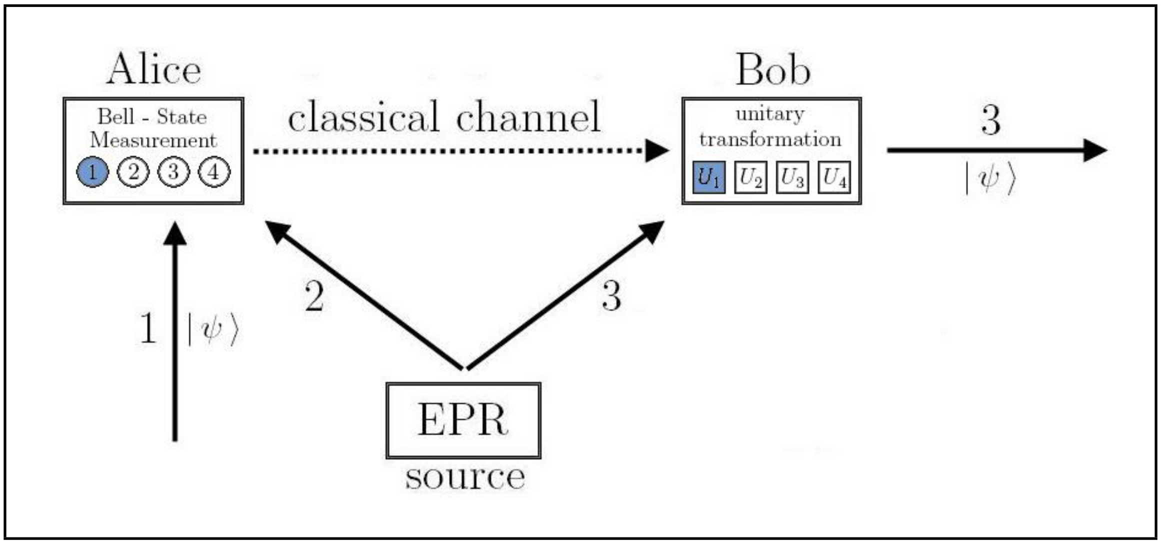

Figure 10.3: Quantum Teleportation Scheme: The state $|\,\psi\,\rangle$, carried at first by particle 1 at Alice's location is transferred to particle 3 at Bob's location via the EPR – entangled – pair of particles 2 & 3. This is done by entangling particles 1 and 2 via a Bell measurement and supplying information in addition to the EPR correlations via a classical channel.

which means that the state of particle 1 and 2 is projected onto one of the four Bell states

$$\begin{aligned} |\,\psi^{\pm}\,\rangle_{12} &= \frac{1}{\sqrt{2}}\,(\,|\,H\,\rangle_1\,|\,V\,\rangle_2 \pm |\,V\,\rangle_1\,|\,H\,\rangle_2\,) \qquad (10.23)\\ |\,\phi^{\pm}\,\rangle_{12} &= \frac{1}{\sqrt{2}}\,(\,|\,H\,\rangle_1\,|\,H\,\rangle_2 \pm |\,V\,\rangle_1\,|\,V\,\rangle_2\,)\;. \qquad (10.24)\end{aligned}$$

Thus before Alice's measurement the state of all three particles is expressed by

$$\begin{aligned} |\,\psi\,\rangle_{123} &= \frac{a}{\sqrt{2}}\,(\,|\,H\,\rangle_1\,|\,H\,\rangle_2\,|\,V\,\rangle_3 - |\,H\,\rangle_1\,|\,V\,\rangle_2\,|\,H\,\rangle_3\,)\\ &+ \frac{b}{\sqrt{2}}\,(\,|\,V\,\rangle_1\,|\,H\,\rangle_2\,|\,V\,\rangle_3 - |\,V\,\rangle_1\,|\,V\,\rangle_2\,|\,H\,\rangle_3\,)\;. \qquad (10.25)\end{aligned}$$

Next we utilize the fact, that the Bell states (Eqs. (10.23) - (10.24)) form a CONS for the 2-qubit states, which allows for any state of the form $|\ \rangle_1\,|\ \rangle_2$ to be expanded in terms of the Bell states. Rewriting the total state (10.25) in such an expansion yields

$$\begin{aligned} |\,\psi\,\rangle_{123} = \frac{1}{2}\,\{\ & |\,\psi^-\,\rangle_{12}\ \ (\,-a\,|\,H\,\rangle_3 - b\,|\,V\,\rangle_3\,) + \\ +\ & |\,\psi^+\,\rangle_{12}\ \ (\,-a\,|\,H\,\rangle_3 + b\,|\,V\,\rangle_3\,) + \\ +\ & |\,\phi^-\,\rangle_{12}\ \ (\,+a\,|\,H\,\rangle_3 + b\,|\,V\,\rangle_3\,) + \\ +\ & |\,\phi^+\,\rangle_{12}\ \ (\,+a\,|\,H\,\rangle_3 - b\,|\,V\,\rangle_3\,)\ \}\;. \qquad (10.26)\end{aligned}$$

Now Alice performs a Bell measurement, i.e. she projects the state of Eq. (10.26) onto one of the four Bell states, e.g. onto $|\,\psi^-\,\rangle_{12}$

$$|\,\psi^-\,\rangle_{12}\,\langle\,\psi^-\,|\,\psi\,\rangle_{123} \;=\; \frac{1}{2}\;|\,\psi^-\,\rangle_{12}\,(\,-\,a\;|\,H\,\rangle_3 \;-\; b\;|\,V\,\rangle_3\,) \;=\; -\,\frac{1}{2}\;|\,\psi^-\,\rangle_{12}\,|\,\psi\,\rangle_3\;, \quad (10.27)$$

then particle 3 on Bob's side is instantaneously in the state initially carried by particle 1, apart from some unitary transformation, i.e. the expansion from Eq. (10.26) can be rewritten as

$$\begin{aligned} |\,\psi\,\rangle_{123} \;=\; \frac{1}{2}\,\{ \quad & |\,\psi^-\,\rangle_{12} \quad U_1\;|\,\psi\,\rangle_3 \;+ \\ +\quad & |\,\psi^+\,\rangle_{12} \quad U_2\;|\,\psi\,\rangle_3 \;+ \\ +\quad & |\,\phi^-\,\rangle_{12} \quad U_3\;|\,\psi\,\rangle_3 \;+ \\ +\quad & |\,\phi^+\,\rangle_{12} \quad U_4\;|\,\psi\,\rangle_3 \;\}\;. \end{aligned} \quad (10.28)$$

However, since Bob does not know onto which of the four Bell states the total state was projected by Alice, because this happens randomly, he needs additional classical information from Alice (via mobile phone, e.g.) to decide, which of the four unitary operations U_1, U_2, U_3 or U_4 he has to perform. But once given this information, he will find particle 3 in exactly the same polarization state as particle 1 was before. The unitary transformations can be represented by the matrices

$$U_1 \;=\; -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\;,\quad U_2 \;=\; \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\;,\quad U_3 \;=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\;,\quad U_4 \;=\; \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\;. \quad (10.29)$$

All these transformations can be experimentally easily realized.

Remarks:

I) The state of particle 1, to be teleported, can be an unknown quantum state which even need not be well-defined (see the phenomenon of "*entanglement swapping*").

II) The initial state of particle 1 is destroyed by the teleportation process, i.e. one can not duplicate a given quantum state, which is expressed in the so called "*no cloning theorem*".

III) The teleportation process does not underly any distance restrictions, in principle, it can occur over any distances.

IV) Experiments on quantum teleportation have been repeatedly performed with great success, first of all, by Anton Zeilinger [28] and his group in Innsbruck in 1997, and since then by several other groups. A very appealing variant was performed by Rupert Ursin [29] in Vienna, see Fig. 10.4.

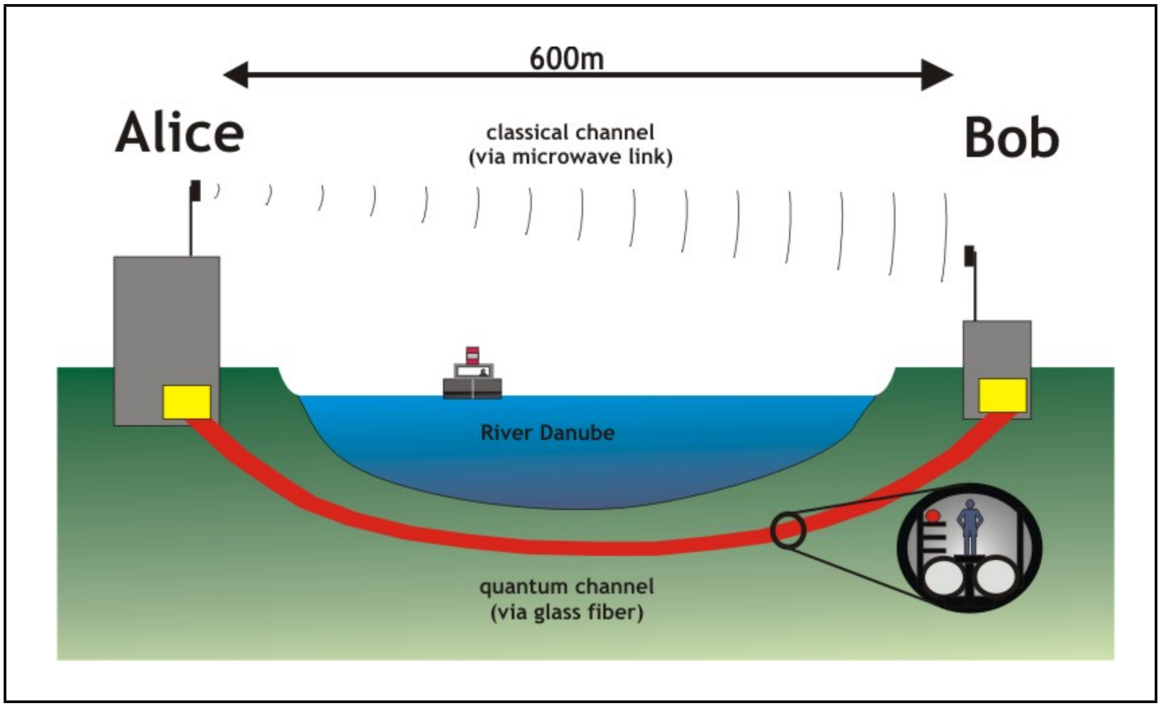

Figure 10.4: Quantum Teleportation under the Danube: The experiment was performed by R. Ursin of the Zeilinger group in 2004, where EPR pairs, created at Alice's location on one side of the Danube, were shared via a glass fiber to teleport quantum states across the river at a distance of ca. 600 m. Figure from: http://de.wikipedia.org/w/index.php?title=Bild:Experimental_overview.jpg

Bibliography

Literature

S. Arroyo Camejo *Skurrile Quantenwelt* · Springer · 2006

J.-L. Basdevant *Lectures on Quantum Mechanics* · Springer · 2007

J.-L. Basdevant & J. Dalibard *Quantum Mechanics* · Springer · 2002

R.P. Feynman *The Feynman lectures on physics, Vol.3* · Addison-Wesley · 1971

D.J. Griffiths *Introduction to Quantum Mechanics* · Prentice Hall · 1995

K.C. Hannabuss *An Introduction to Quantum Theory* · Clarendon Press · 1997

M. Le Bellac *Quantum Physics* · Camebridge University Press · 2006

A. Peres *Quantum Theory - Concepts and Methods* · Kluwer Academic · 2002

H. Pietschmann *Quantenmechanik verstehen* · Springer · 2003

J.J. Sakurai *Modern Quantum Mechanics* · Addison-Wesley · 1994

F. Schwabl *Quantenmechanik* · Springer · 1998

A. Zeilinger *Einsteins Spuk* · Goldmann · 2007

Articles and additional literature

[1] **M. Brandl** *Project Physnet - m219 - The Compton Effect*· 12.03.08
Michigan State Universtiy · http://www.physnet.org/modules/pdfmodules/m219.pdf

[2] **E. Hecht** *Optics*· Addison-Wesley · 1974

[3] **R.P. Feynman** *Lectures on Physics*· Vol.III Addison-Wesley · 1965

[4] **G. Möllenstedt & C. Jönsson** *Elektronen-Mehrfachinterferenzen an regelmäßig hergestellten Feinspalten* · Zeitschrift für Physik A, Vol. 155, Nr. 4, Pages 472-474 · 1959

[5] **H. Rauch & H. Kurz** *Beugung thermischer Neutronen an einem Strichgitter* · Zeitschrift für Physik A, Vol. 220, Nr. 5, Pages 419-426 · 1969

[6] **C. G. Shull** *Single-Slit Diffraction of Neutrons* · Phys. Rev., Vol. 179, Nr. 3, Pages 752-754 · 1969

[7] **A. Zeilinger et al.** *Single- and double-slit diffraction of neutrons* · Rev. Mod. Phys. 60, 1067-1073 · 1988

[8] **M. Arndt et al.** *Wave-particle duality of C60 molecules* · Nature 401, 680-682 · 1999

[9] **P. Holland** *Quantum Theory of Motion* · Camebridge University Press · 1993

[10] **O. Passon** *What you always wanted to know about Bohmian mechanics but were afraid to ask* · arXiv:quant-ph/0611032v1 · 2006

[11] **B. D'Espagnat** *Conceptual Foundations of Quantum Mechanics* · Perseus · 1999

[12] **O. Nairz, M. Arndt & A. Zeilinger** *Quantum Interference Experiments with Large Molecules* · Am. J. Phys. 71, 319 & 1084 · 2003

[13] **M. Abramowitz & I. Stegun** *Handbook of Mathematical Functions* · Dover Publications · 1972

[14] **H. Hüffel** *Theoretische Methoden der Physik 2* · script to the lecture · 20.5.08 University of Vienna · http://homepage.univie.ac.at/helmuth.hueffel/SS08/M2.pdf

[15] **P. Thun-Hohenstein** *Quantum Phase and Uncertainty* · Diploma Thesis · 2010 · University of Vienna

[16] **H. Rauch, A. Zeilinger, G. Badurek, A. Wilfing, W. Bauspiess & U. Bonse**
Verification of coherent spinor rotation of fermions
Physics Letters 54A, 425-427 · 1975

[17] **Tanja Traxler**
Decoherence and the Physics of Open Quantum Systems
script to the lecture of R. A. Bertlmann · 13.10.08 · University of Vienna
https://elearning.mat.univie.ac.at/physikwiki/images/9/9b/Decoscript.pdf

[18] **A. Einstein, B. Podolsky & N. Rosen**
Can quantum-mechanical description of physical reality be considered complete?
Physical Review 47, 777-780 · 1935

[19] **D. Bohm**
Discussion of experimental proof for the paradox of Einstein, Rosen and Podolsky
Physical Review 108, 1070-1076 · 1957

[20] **R. A. Bertlmann**
Bell's theorem and the nature of reality
Foundation of Physics 20, 1191-1212 · 1990

[21] **N. Bohr**
Can quantum-mechanical description of physical reality be considered complete?
Physical Review 48, 696-702 · 1935

[22] **J. S. Bell**
On the Einstein Podolsky Rosen paradox
Physics 1, 195-200 · 1964

[23] **J. S. Freedman & J. F. Clauser**
Experimental test of local hidden-variable theories
Physical Review Letters 28, 938-941 · 1972

[24] **E. S. Fry & R. C. Thompson**
Experimental test of local hidden-variable theories
Physical Review Letters 37, 465-468 · 1976

[25] **A. Aspect, J. Dalibard & G. Roger**
Expereimental test of Bell's inequalities using time-varying analyzers
Physical Review Letters 49, 1804-1807 · 1982

[26] **G. Weihs, T. Jennewein, C. Simon, H. Weinfurter & A. Zeilinger**
Violation of Bell's inequality under strict Einstein Locality Conditions
Physical Review Letters 81, 5039-5043 · 1998

[27] **C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres & W. K. Wootters**
Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels
Physical Review Letters 70, 1895-1899 · 1993

[28] **D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter & A. Zeilinger**
Experimental quantum teleportation
Nature 390, 575-579 · 1997

[29] **R. Ursin, T. Jennewein, M. Aspelmeyer, R. Kaltenbaek, M. Lindenthal, P. Walther & A. Zeilinger** *Quantum teleportation link across the Danube* ·
Nature 430, 849 · 2004